170
101

Effects of Color CRT Misconvergence and Font Type on Text Readability and Subjective Preference

by

Angelia L. Sebok

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Industrial and Systems Engineering

APPROVED:

R.J. Beaton, Chairman

H.L. Snyder

P.T. Kemmerling

April, 1991

Effects of Color CRT Misconvergence and Font Type on Text Readability and Subjective Preference

(ABSTRACT)

by

Angelia L. Sebok

Committee Chairman: Robert J. Beaton

Industrial and Systems Engineering

Our information-oriented society relies on the widespread use of color CRT displays. Misconvergence of the primary colors of a shadow-mask CRT is a problem with this technology that deserves human factors engineering consideration. The purpose of this research was as follows: (1) to determine the effects of misconvergence type and amount and font type on reading performance (time required and errors made), (2) to determine the effects of misconvergence type and amount and font type on subjective image quality ratings, (3) to determine the role of luminance and chrominance contrasts in predicting performance or subjective ratings.

Ten participants performed a simple reading task and rated the image quality of the text they had just seen using a nine-point scale. The text was presented on a shadow-mask CRT. Different misconvergence types and amounts and different font types were presented.

Neither font type, misconvergence type, nor misconvergence amount affected the time required to perform the reading task. Only misconvergence type affected the rate at which errors occurred, with blue misconvergence of a white character resulting in the most errors and cyan misconvergence resulting in the fewest errors. Font type, misconvergence type, misconvergence amount, and the misconvergence type and amount interaction all affected subjective ratings, with 1 to 2 arcmin being the largest acceptable misconvergence amounts.

The Yu′v′ chrominance contrast between the stationary misconvergence fringe and the background was positively correlated with subjective preference ratings.

ACKNOWLEDGEMENTS

I would like to thank to Dr. Robert Beaton for his many wise words of encouragment, advice, and guidance, and his understanding. I also thank Dr. Snyder and Professor Kemmerling for their encouragement and helpful suggestions.

Dr. Dryden, who managed to find subject money in times of budget cuts; Woody Farley, who created the misconvergence generator; and Deborah Cox, who offered much administrative guidance, all deserve special recognition and thanks. Also, thanks to Dr. Snyder, Dr. Greene, and Dr. Torgersen, who kept me employed throughout my graduate student career. Under the direction of Dr. Torgersen, the preparation of this paper was funded by the U.S. Department of Energy (DOE) grant number DE-FG05-86DP70033.

Thanks to my friends, as always, and thanks to Scott for his humor and adventurous sense of life which helped keep me sane and happy throughout this work.

Finally, I thank my dad, mom, brother, and sisters for their determination, willingness to face challenges, and successes, which were a constant inspiration throughout this work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

In our society, large amounts of information are transmitted and processed every minute. One popular method of presenting this information is by visual display. The most commonly used visual display device is the cathode-ray tube (CRT). This display device is used in most computer systems, televisions, aircraft cockpits, medical imaging systems, and head-up displays.

CRT technology is not new, having first been implemented in early electronic test equipment and, then, in television in the 1920s. As shown in Figure 1, a monochrome CRT display consists of a vacuum tube which houses an electron gun at one end and a phosphor-coated glass screen at the other end. In the electron gun structure, a cathode generates electrons which are sent through the tube toward the anode at the phosphor screen. As electrons strike the phosphor screen, light is emitted. The electron gun produces a beam of electrons which is deflected horizontally and vertically to illuminate the entire phosphor screen. By interrupting the electron stream so that only certain areas of the phosphor screen are illuminated, images can be created for viewing by an observer.

A color CRT is similar to the monochrome CRT, as shown in Figure 1. It consists of a cathode-ray tube and relies

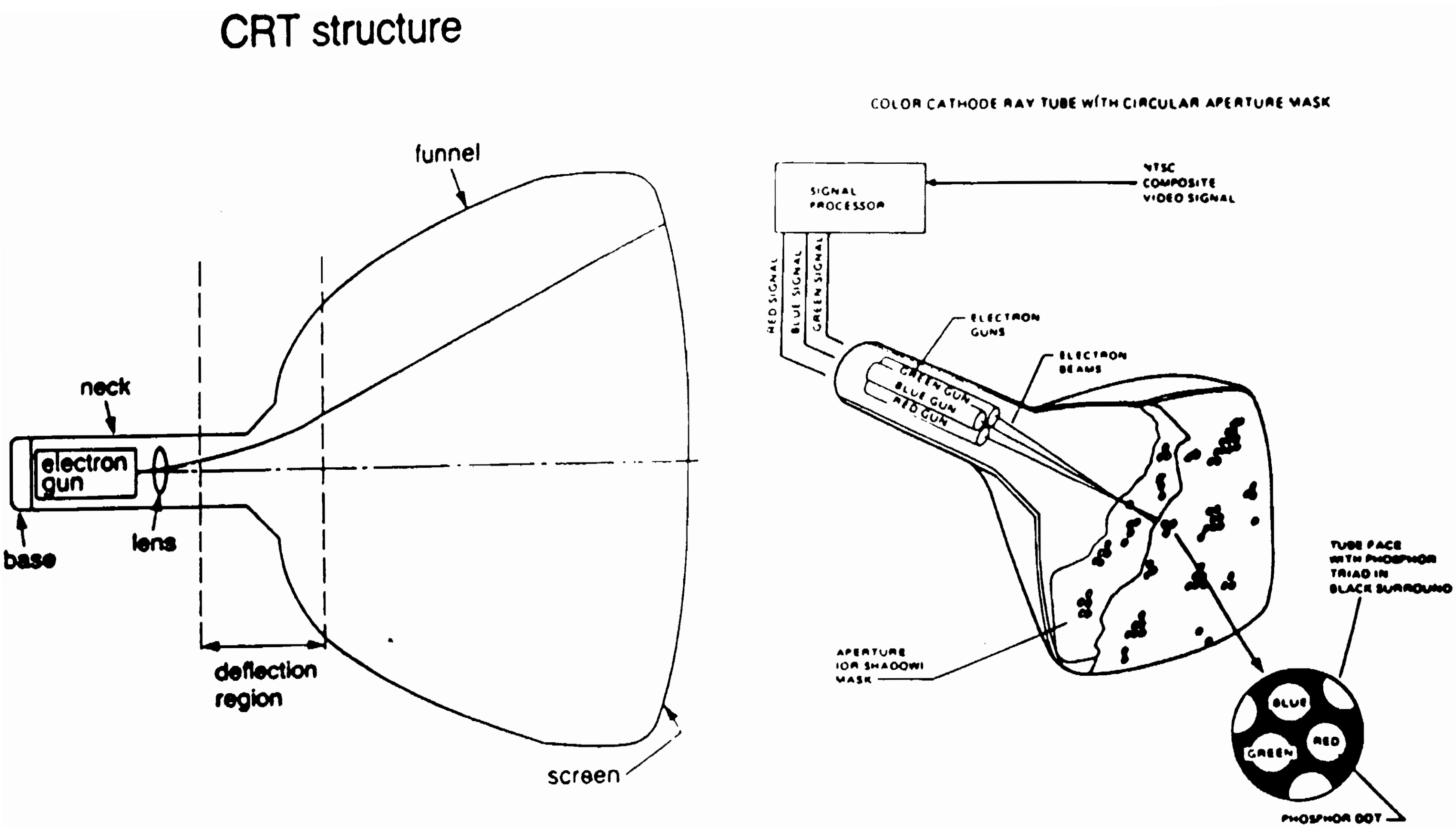


Figure 1: Monochrome and color CRTs
(Lehrer, 1985)

on the light-emitting properties of phosphors to produce images. Unlike the monochrome CRT, however, a color CRT has three colors of phosphors and three electron guns. Each gun activates only one of the three colors of phosphors. The three phosphors emit different colors of light: red, blue, and green. These three colors are the primaries in an additive color system. Various colors are produced by adding together combinations of the primary colors.

To produce a color image, the electron guns send electrons to the appropriate phosphors. A metal plate containing holes or slots, called the shadow mask, directs the electron beams to the appropriate phosphors. The three colors of phosphors are arranged on the screen in groupings called "triads." When the electron beam hits the phosphor, the phosphor emits light of the appropriate color. To obtain a color other than the three primary colors, electrons activate two or three colors of phosphor. The viewer's eye integrates these points of color into a color mixture. For example, to obtain the color yellow, the red and green phosphors are activated. The viewer sees these dots as their additive color, yellow.

Color displays offer several advantages over monochromatic displays. In terms of aesthetic advantages, users generally prefer color displays over monochrome displays. In terms of improved user performance, color displays offer information coding capabilities and increased flexibility in information coding schemes. Furthermore, color reduces the required search time, a benefit especially important in complex images. Finally, the use of color increases symbol visibility and may reduce the required display luminance (Silverstein, 1987).

However, color usage in CRTs can present some problems. One problem related to the use of color is "aesthetic overindulgence": the unnecessary or overextended use of color based on artistic preference rather than performance benefits. Too many colors may be used in a coding system, thus confusing or annoying the viewer (Hale and Billmayer, 1988).

Another set of problems involves the viewer's perception. Chromostereopsis, a three-dimensional visual effect, can result when pure, highly saturated colors (in particular, red and blue) are placed next to each other (Snyder, 1988). Chromatic aberration or peripheral color vision also can create problems (Hale and Billmayer, 1988). While a color may be defined physically, the perceptual interpretation of that color depends on viewing conditions and other visual stimuli; therefore, the use of color is more complex than specifying color coordinates (Walraven, 1984).

Some problems in color usage may be the result of the hardware. Misconvergence occurs when the three electron beams are not aligned correctly. One or more of the electron beams may be off-target, resulting in distorted images. The electrons from misconverged beams do not hit the phosphors in the correct triads. The result is a distorted image, which may have blurred edges, color fringing, or color and image separation (Robertson and Jones, 1984).

Blurred edges result from minimal misconvergence. In the case of color fringing, the misconverged primary results in a fringe of that primary color on one side of the secondary color image. A fringe of the remaining component colors appears on the other side of the secondary color image, as shown in Figure 2. The more severe the misconvergence, the more distorted the image. In extreme cases, misconvergence can result in two images of two separate distinct colors, neither of which is the intended color. These two distinct colors are the misconverged primary and the other color component of the intended secondary color.

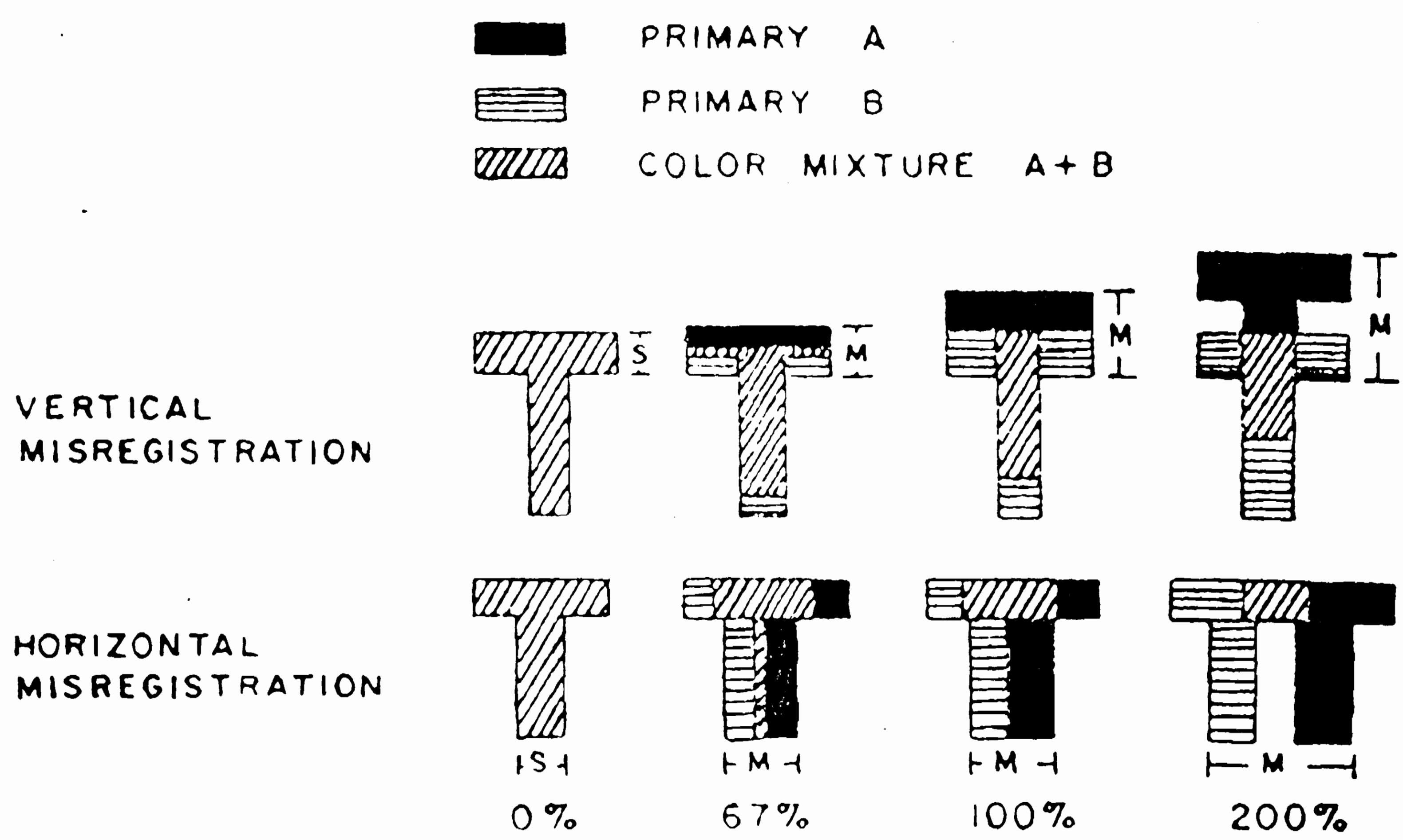


Figure 2: Misconvergence: Color fringing (Snadowsky, Rizy, and Elias, 1966)

Misconvergence can be minimized, but often the adjustments are costly. Given the prevalent use of CRT displays in our society and the increasing demand for color displays, research on the effects of misconvergence on human performance is necessary to determine the amount of misconvergence which can be tolerated in displays.

LITERATURE REVIEW

Misconvergence

Colors on a CRT screen are created by the actuation of the primary phosphor colors. If two or three electron beams are used to create a color mixture and the beams are correctly registered, the resulting pixel is converged. However, if the beams are misregistered, the resulting image is misconverged. The amount of misconvergence is the center-to-center misregistration among the electron beams (Merrifield, 1987). The amount of misregistration affects the quality of an image and, thus, may affect the performance of the observer. To minimize potential performance degradation, the Society of Automotive Engineers recommends misconvergence to be no greater than the minimum half-amplitude line width of the display, as shown in Figure 3.

Previous research on misconvergence is scarce. Much of the existing research has focused on the technological aspects of misconvergence rather than the user performance aspects. Research is needed to explore the effects of misconvergence on the performance of tasks as well as the subjective assessment of the quality of the appearance of the image.

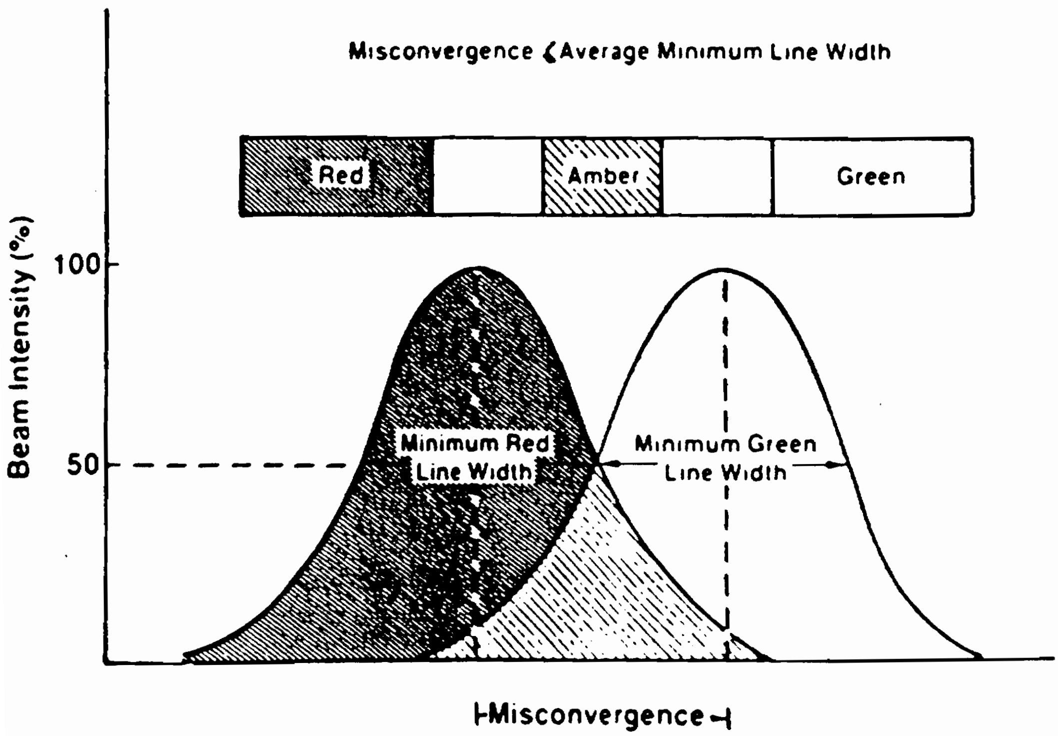


Figure 3: SAE misconvergence standard (Merrifield, 1987)

The studies of the effects of misconvergence on user performance can be divided into three general categories: those concerned with the detection of misconvergence, those concerned with subjective ratings of misconverged images, and those concerned with objective task performance under misconvergence conditions. Each of these areas and the studies performed in them will be discussed separately.

Detection.

Merrifield, Haakenstad, Ruggiero, and Lee, (1979). One purpose of this study was to determine the level of misconvergence detectable by viewers. Subjects were seated 81.3 cm from a computer screen and shown a "DH" symbol on the screen. The symbol was either yellow or magenta. In the yellow condition, either the red or the green primary was misconverged; in the magenta condition, the red or the blue was misconverged. The amount of misconvergence varied across tasks. Subjects were asked to identify the type and direction of misconvergence (for example, to say "red above" when the misconvergence caused a red fringe to appear above the secondary-color symbol). The detection threshold was defined as that level of misconvergence at which 50 percent of the subject responses were correct. In this study, the detection threshold was 0.13 mm (0.5 arcmin) of misconvergence. Also, a yellow symbol with red or green misconvergence was found to have a lower detection threshold than a magenta symbol with red or blue misconvergence.

Silverstein and Lepkowski (1986). In this study, subjects were seated 81.3 cm from a screen and presented a secondary color line on a Sperry high-resolution shadow-mask color CRT. Subjects were asked to vary the spot size of one of the primary colors in the line. These size variations occurred in 0.0254-mm (6-arcsec) increments. The subjects were asked to stop varying spot size as soon as they noticed a difference in the secondary-color line. The subjects then were asked to classify the noticed difference in the line: line width change, color fringing, line color change, or line brightness change. The secondary lines used in the study were magenta, yellow, and cyan. Subjects were tested in two conditions: (1) high ambient illumination (8600 lux) and high background luminance (42.5 cd/m^2) and (2) low ambient illumination (2.15 lux) and low background luminace (0.005 cd/m^2).

The findings of this study are shown in Table 1. Briefly summarized, changes in the blue primary spot size were far less noticable than changes in the spot size of the red or green primaries. Also, the detection threshold for the red and green primaries was constant across illumination and luminance conditions, whereas the detection threshold

TABLE 1: Misconvergence Detection Thresholds as found by Silverstein and Lepkowski (1986)

Lighting Conditions	Primary Misconverged Color	Secondary Color	Detection Threshold mm	arcmin
High	Green	Yellow	0.70	3.0
		Cyan	0.70	3.0
Low	Green	Yellow	0.70	3.0
		Cyan	0.70	3.0
High	Red	Yellow	0.80	3.4
		Magenta	0.60	2.5
Low	Red	Yellow	0.80	3.4
		Magenta	0.60	2.5
High	Blue	Magenta	1.07	4.5
		Cyan	1.07	4.5
Low	Blue	Magenta	1.40	5.9
		Cyan	1.40	5.9

for the blue primary increased in the low illumination and low luminance conditions.

Another finding, not shown in Table 1, was that the perceived type of change in the line varied with color. For changes in the red and green primaries, the first-noticed change in the line was a line width change. For changes in the blue primary, the first-noticed change was a blue fringe around the line. For observers to notice this fringe, a greater change in spot size was required than for the first-noticed change in line width under the red and green varying conditions.

DeVilbiss (1990). This study investigated the threshold of misconvergence detection, subjective reactions to misconvergence, and reading performance under misconvergence conditions. All three of these investigations were performed under the same general conditions. Twelve college-age subjects (six males and six females) with 20/25 or better visual acuity, normal lateral and vertical phoria, and normal color vision participated in the study. The subjects were seated 60 cm from the display. The ambient illumination was approximately 100 lux at the display.

For the detection threshold task, the subjects were presented a vertical line segment in the center of the display. This segment was of a secondary color. Nine levels of misconvergence were presented, ranging from 0.0 mm to 0.48 mm (0 to 2.75 arcmin). Subjects were shown these conditions in a random order and were asked to answer "yes" or "no" in response to the question "Do you see any misconvergence?"

This study showed that red or blue misconvergence on a magenta line was detected by 50% of the subjects at 0.09 mm (0.52 arcmin). Red or green misconvergence on a yellow line had to be 0.36 mm (2.06 arcmin) before it was detected by 50% of the subjects. Green or blue misconvergence on a cyan line was detectable at 0.24 mm (1.38 arcmin), and misconvergence of a white line was detected at approximately 0.20 mm (1.15 arcmin).

Summary of detection. In general, these studies suggest that misconvergence is detected at a separation between primaries of approximately 1 to 2 arcmin. The higher values (2.5 to 5.9 arcmin) found in the Silverstein and Lepkowski (1986) study are explained by the stimulus used in that study. Instead of actual misconvergence, as was used in the Merrifield et al. (1979) and DeVilbiss (1990) studies, the Silverstein and Lepkowski experiment used a widened spot profile. The primary colors continued to overlap, rather than separate. Thus, the change in spot size must be greater than the change in color separation for misconvergence to be detected (DeVilbiss, 1990).

Another finding of these studies is that blue misconvergence must be greater than red or green in order to be detected. In particular, the red or green misconvergence of a yellow line is generally the most readily detectable misconvergence. DeVilbiss (1990) found that the red or blue misconvergence of a magenta line was detected more readily than the red or green misconvergence of a yellow line. However, this difference may have been due to the stimuli used. The red and green beams were aligned as shown in Figure 4, so that for the first couple of misconvergence steps of the green beam, the green beam actually moved closer to the red beam (DeVilbiss, 1990). Silverstein and Lepkowski (1986) found red misconvergence of a magenta line to have the lowest detection threshold. While this particular finding agrees with DeVilbiss (1990), Silverstein and Lepkowski found blue misconvergence of the magenta line had the highest detection threshold.

Subjective ratings. These studies investigated how objectionable users found different levels of misconvergence. Subjective ratings are as important as objective measures, since users of commercial display products may reject a CRT with low perceived image quality.

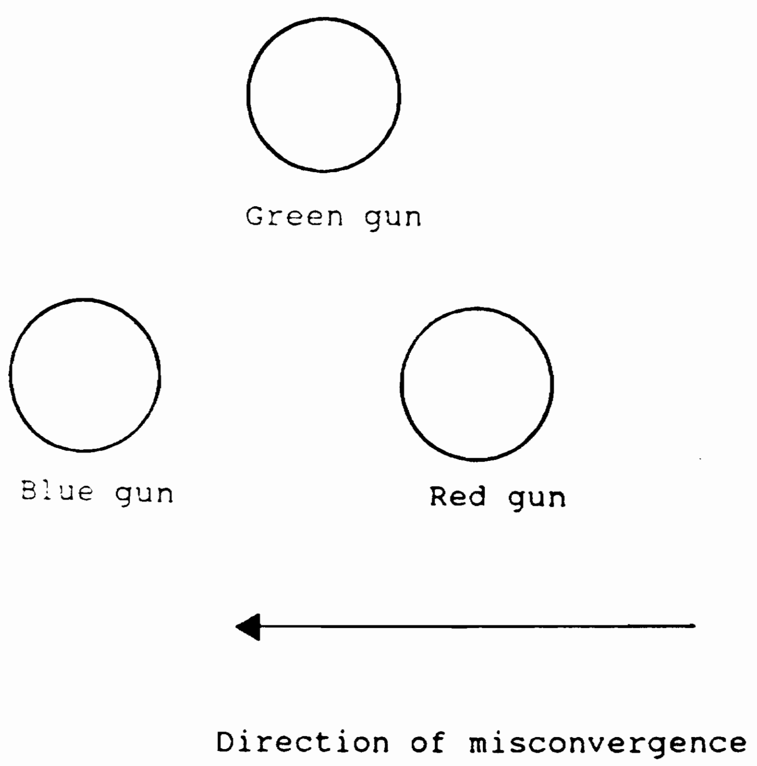


Figure 4: Red, blue, and green gun alignment used in the DeVilbiss (1990) study.

A consumer will not purchase a misconverged CRT, even though the misconvergence may not impair readability (Robertson and Jones, 1984).

Merrifield, Haakenstad, Ruggiero, and Lee (1979). In this study, discussed previously, subjects were seated 81.3 cm from a computer screen. They were presented a "DH" symbol and were asked to detect when misconvergence occurred. Also, they were asked to rate the level of objection they felt using a six-point scale ranging from "no objection" to "highly objectionable."

Qualitatively, subjects found misconvergence on the yellow symbol (red-green misconvergence) to be more objectionable than misconvergence of the magenta symbol (red-blue misconvergence). Subjects objected to higher levels of misconvergence such that the objection rating scale "tracked the misconvergence levels almost perfectly." (Merrifield et al., 1979). For the 34-mm spot size used in the study, 50% of the subjects detected misconvergence of 0.13 mm but did not find this level objectionable. However, 88% of the subjects objected to 0.33 mm of misconvergence.

Robertson and Jones (1984). This study used two experiments to investigate subjective opinions of different levels of misconvergence. Both experiments were run under the same environmental conditions -- normal office lighting and furniture and subjects were seated at their preferred distance from the computer screen. Viewing distance was not controlled in this study. In both conditions, subjects were shown seven colors and seven levels of misconvergence. They also were shown a screen of either text or graphics and asked to rate the quality of the screen along a seven-point scale ("very bad" to "excellent").

One experiment used a small area of a computer screen for text -- only one sixteenth of the screen was filled with text. The other condition used the whole screen for either text or graphics presentation. The image on the large screen was the same size as that shown on the small screen, but it was repeated sixteen times in order to fill the whole screen. The subjects were given practice trials in each condition to acquaint themselves with the images and the rating scale.

This study found that subjects had no objection to misconvergence up to 0.25 mm, but more than 50% objected to 0.55 mm of misconvergence. Based on these findings and the criterion that a misconvergence level was acceptable if 90% of the subjects rated it as "adequate" or better, the researchers recommended that the maximum allowable misconvergence on a CRT screen should be between 0.20 and 0.30 mm.

Another finding in this study was that no difference in viewer objection ratings was found between the text and graphics screens. Further, no differences were found between the objections to the various types of misconvergence, except for cyan misconvergence being less annoying on the small area screen than the other misconvergence types. No difference in subjective objections was found between the large and small area display screens.

DeVilbiss (1990). For the assessment of subjective ratings, subjects were presented a standard block of text containing eight lines and approximately 10 words per line. The text was presented under all misconvergence conditions mentioned above. Subjects were asked to rate image quality using a nine-point scale ("best imaginable" to "worst imaginable").

This study found that subjects objected to greater levels of misconvergence. As misconvergence increased, the subjective image quality rating decreased. For each type of misconvergence, the levels at which the text was considered "passable" (the middle of the scale) are shown in Table 2. Among the types of misconvergence examined, red or blue misconvergence of magenta text was found to be the most objectionable. As misconvergence increased, the subjective

TABLE 2: Passable Misconvergence Levels as found by DeVilbiss (1990)

Text Color	Misconverged Primary	Misconvergence mm	arcmin
Cyan	B and G	0.84	4.81
White	B	0.76	4.35
White	R	0.72	4.13
Yellow	R and G	0.60	3.44
Magenta	R and B	0.48	2.75

ratings for magenta text dropped faster and to lower levels than did the other types of misconvergence. Green and red misconvergence of yellow text was the second most objectionable condition. For white text, red misconvergence was slightly more objectionable than blue misconvergence. The least objectionable misconvergence was blue or green components of cyan text. The levels of misconvergence required to cause unfavorable ratings were not sufficient to cause impaired legibility.

Herb, 1990. In this study, 10 college-age subjects (5 males, 5 females) with normal vision participated in a visual search task. Six misconvergence types and five misconvergence amounts (0 - 5.5 arcmin) were tested. Following each target-selection task, the participants were asked to rate the quality of the image they had just seen using a nine-point scale. This study found BG-R misconvergence to be rated more favorably than the other types of misconvergence. The difference between the BG-R misconvergence and G-R and RB-G misconvergence was not significant. Further, the G-R and RB-G were not significantly different from the B-R or GR-B misconvergence types. Also, G-B misconvergence, the most objectionable in this study, was not significantly different from GR-B or B-R.

Ansley (1991). This study consisted of a reading task in which subjects were presented four lines of text with approximately 10 words per line. Ten college-age subjects (5 males and 5 females) with normal vision were asked to rate the image quality of the text using the nine-point scale of the DeVilbiss and Herb studies. Ansley tested the same misconvergence types as Herb, including three additional misconvergence types. Further, three screen luminance conditions were tested. Of the misconvergence types (and backgrounds) that were identical to those used in this study and the Herb (1990) study, two significantly different groups were found based on subjective ratings. The higher-rated group consisted of the BG-R, the G-B, and the RG-B misconvergence conditions, and the lower-rated group consisted of the R-G, the RB-G, and the R-B misconvergence.

Summary of subjective ratings. Generally, these studies suggest that misconvergence becomes objectionable before legibility or usability is impaired. Merrifield et al. (1979) found that misconvergence of approximately 100% of the spot width at half the maximum luminance was acceptable. The unacceptable misconvergence was 0.33 mm (1.4 arcmin) and the spot width was 0.34 mm. Robertson and Jones (1984) found 0.2 to 0.3 mm to be the maximum acceptable misconvergence. The spot width used in this study was 0.8 mm measured at the 5% maximum luminance boundaries.

To compare these two studies, the spot sizes must be measured in the same manner. Assuming the spot in the Robertson and Jones (1984) study to be Gaussian shaped (a reasonable assumption for CRT devices), the spot width at 50% of the maximum luminance is 0.38 mm (DeVilbiss, 1990). Therefore, the maximum allowable misconvergence of 0.3 mm is approximately 79% of the spot width. These results are in good agreement, since the Merrifield et al. (1979) study examined the minimum objectionable misconvergence and the Robertson and Jones (1984) study examined the maximum acceptable misconvergence.

The other studies show different results. DeVilbiss found misconvergence amounts receiving "passable" ratings varied between 0.48 mm and 0.84 mm (2.75 and 4.81 arcmin). Herb (1990) and Ansley (1991) found subjective ratings for misconverged characters and symbols decreased significantly from converged-image ratings at misconvergence amounts of 2.75 arcmin.

The differences in these experimental findings was probably due to the nature of the experiments. In the Robertson and Jones (1984) and the Merrifield et al. (1979) studies, subjects were asked to rate the level of objection experienced with misconvergence: they were not asked to perform a task. In the DeVilbiss (1990), Herb (1990), and Ansley (1991) studies, subjects first performed a task, then subjectively rated the misconvergence. Since misconvergence did not affect performance in those studies, perhaps subjects were less critical in their subjective rating evaluations. After realizing the misconvergence they just saw did not make the task too difficult to complete, subjects may have been less likely to rate the misconvergence as low as they might have otherwise.

The types of misconvergence found most objectionable varied among the studies. Merrifield et al. and Robertson and Jones found blue misconvergence to be less objectionable than red or green misconvergence. DeVilbiss (1990) found magenta (red and blue) misconvergence to be the most objectionable, which may be due to the arrangement of the electron beams, as described earlier. This arrangement required the red beam to move closer to the green beam for the first couple of steps of the misconvergence increments. The second most objectionable misconvergence was yellow (red and green), which agrees with the earlier studies. Ansley (1991) found the same ordering of acceptable misconvergence types as DeVilbiss (1990).

Herb (1990) found G-B and RG-B misconvergence types to be rated as highly objectionable. In the other studies discussed, these types were highly rated. Further, the R-G and the RB-G types were preferred in the Herb study, but were rated as highly objectionable in the other studies. The two other types of misconvergence, BG-R and R-B, were rated similarly between the Herb research and the DeVilbiss (1990) and Ansley (1991) research studies.

Performance.

Snadowski, Rizy, and Elias (1966). This study investigated the effects of misconvergence on the performance of a simple reading task. Six male college students served as subjects. They were screened for normal or corrected to normal 20/20 visual acuity and normal color vision. A Colorvision additive color projector, not a CRT, was used to present the stimuli. The subjects were seated 5.6 m from the screen and were presented an 18 x 14 matrix of capital letters and single digits. These characters subtended 27 arcmin of visual angle and were one of seven colors and in one of seven misregistration levels from 0 to 200 percent. The subjects were asked to read, left to right and top to bottom, all symbols of one specified color. The time required to complete the task as well as errors of identification and omission were recorded. Due to limitations in the equipment, red misconvergence was always presented for the first third of the experiment, then blue for the second third, and green for the final third.

This study found that as misconvergence increased, the time to complete the task and the errors increased. The decline in performance was statistically significant only when misconvergence was 100% of the character stroke width. At the 33 and 67 percent levels, the decline was not statistically different from performance at the 0% level.

Among the types of misconvergence used, blue had the least deleterious effect on performance. For misconvergence levels less than 100%, green misconvergence most negatively affected performance. For levels greater than 100%, red misconvergence was the most distracting.

The primary colors, which could not be misconverged, showed the least degradation in performance across all misconvergence conditions. However, they did show a performance degradation, likely due to the interference of other primary-color fringes on the misconverged secondary-color symbols. For the secondary colors, magenta resulted in the best performance. Yellow resulted in the second-best performance. For misconvergence less than 100%, white characters yielded the third-best performance, while cyan resulted in the worst performance. For misconvergence greater than 100%, this trend was reversed, with cyan characters resulting in better performance than white characters.

DeVilbiss (1990). In this study, subjects performed a simple reading task. They were timed for how long it took them to read the passage. The text was presented under nine levels of misconvergence (ranging from 0.0 mm to 0.96 mm or 0.0 to 5.5 arcmin) and four colors of text. The findings were that performance was not affected by misconvergence amount or type: none of the reading times were significantly different from the others.

Herb (1990). This study used a visual search task in which participants were shown a target icon and then presented a screen containing several symbols, including the target icon. They were instructed to search for the target icon and, then, to select it using a mouse input device. Six misconvergence types and five misconvergence amounts, the same as were used in this study, were tested. The results of this research also indicated that the time required to find the target icon was not significantly affected by either misconvergence type or amount.

Ansley (1991). Like DeVilbiss (1990), Ansley used a Tinker Speed of Reading test to measure the effects of misconvergence type and amount. The misconvergence types tested in this study include the six that were tested by

DeVilbiss (1990), Herb (1990), and in this research, as well as three additional conditions in which the background color was changed. The misconvergence amounts are the same as were tested by DeVilbiss and Herb, and in the present research. Ansley (1991) also found that neither misconvergence type nor amount significantly affected the time required to complete the reading task.

Summary of performance. Of the misconvergence studies conducted, all except one show that display misconvergence does not affect user performance. The Snadowski et al. (1966) study found that performance was degraded under misconvergence conditions, with more severe degradations occurring with more severe misconvergence. The DeVilbiss (1990), Herb (1990), and Ansley (1991) studies found that misconvergence did not affect performance. These differences may be due to the nature of the tasks used in the various experiments.

In the DeVilbiss (1990) and Ansley (1991) studies, the task involved reading a text passage and selecting a word. The Herb (1990) study consisted of selecting a particular icon from a set of symbols. The Snadowski et al. (1966) task involved verbally selecting all characters of a particular color.

The tasks used in the DeVilbiss and Ansley studies allowed the redundancy in the English language to be used so that difficult-to-read words could be identified by contextual cues. The Snadowski et al. (1966) task did not allow the use of redundancy. However, a comparison of the findings of the Herb (1990) study with the Snadowski et al. (1966) study does not support the advantage of language redundancy: the Herb study, like the Snadowski et al. study, did not allow the use of English language redundancy to aid in task completion, and misconvergence type or amount did not significantly affect the time required for task completion.

All of the studies which found misconvergence to have no significant effect on task performance time relied on character or symbol recognition. The study which found misconvergence to significantly affect performance time required subjects to identify alphanumeric characters based on their color. Misconvergence does not change the actual symbol shape; it may add fringes to the shape or, in severe cases, may result in two shapes being present. Misconvergence does change the amount of the color of the symbol, with increasing misconvergence resulting in less of the original color being present. In severe cases, none of the original character color is present. The effects of misconvergence are such that the symbol shape is distorted less dramatically than is the symbol color. Thus, studies that investigate the identification of symbol shapes show no increase in task completion time, but the study that investigated the identification of symbol colors did show an increase in time for task completion.

Other Issues in Visual Display Legibility

Besides misconvergence, other factors influence the legibility of text on a visual display. These factors include, but are certainly not limited to, the font type, the luminance and chrominance contrasts of the characters, and misconvergence fringes.

Font type. While many types of fonts exist, they can be divided into two general classes: serif and sans serif. Serif fonts have letters with accents or "little flourishes or embellishments" (Sanders and McCormick, 1987). Sans serif fonts lack these embellishments and are "modern" in appearance. Sanders and McCormick (1987) recommend using sans serif fonts when legibility is critical, implying that serif fonts are somewhat less legible.

No misconvergence studies have examined font types. The sans serif font should result in better performance than the serif font, since the extra "cues" of the serif fonts will become extra "visual noise" in misconvergence conditions.

Luminance and chrominance contrast. Several studies have investigated the effects of luminance and chrominance contrasts on legibility (Lippert, 1985; Rogers and Gutmann, 1983; Sayer, Sebok, and Snyder, 1990; Snyder and Taylor, 1979). While luminance contrast ratios of 7:1 to 10:1 generally are recommended (Human Factors Society, 1988; Shurtleff, 1980), this standard varies with character height. For larger characters, lower luminance contrasts are required (Rogers and Gutmann, 1983; Sanders and McCormick, 1987; Snyder and Taylor, 1979).

In misconvergence studies, no attempt to control luminance contrast has been made. When different colors were presented, the luminances of these colors were allowed to vary as they do in the operation of the display. Since blue has the lowest luminance and, therefore, the lowest luminance contrast, the finding that blue misconvergence is least distracting may not be surprising.

Rogers and Gutmann (1983) found that for luminance contrasts of 2:1, the character height should be at least 16.9 arcmin for comfortable reading by 95% of participants. In the Snadowski et al. (1966) study, the luminance contrast ratio between the blue characters and the background was 6:1, and the character height was 27 arcmin. Thus, the luminance of the characters did not interfere with legibility, even for the low-luminance blue characters. Since the minimum height of the characters in the present study were 28.7 arcmin and the minimum luminance contrast was greater than 2:1, the subjects should not have experienced problems reading the text because of insufficient luminance contrast.

Chrominance contrast of characters relative to their background also affects legibility (Lippert, 1985; Sayer et al., 1990). The closer the two chrominances, the more legibility is impaired. Under misconvergence conditions, the immediate background of characters is changed. For a perfectly converged character, the background is black. For a misconverged character, the immediate background is composed of the component colors of the character. For example, if the misconverged character is cyan, the immediate background is green and blue.

The immediate background may adversely affect legibility if it is similar in chrominance and luminance to the character. However, the opposite effect also may occur: similar colors may make a letter appear blurred, but still readable, where vastly different colors may result in visual chaos (for example, red and green fringes on a yellow character). Also, these two effects may both be important, each playing a role at different levels of misconvergence.

These effects may explain an unusual finding in some of the misconvergence studies. Some misconvergence colors are more distracting at lower misconvergence levels than other colors. For example, Snadowski et al. (1966) found that green misconvergence most adversely affected performance for misconvergence conditions less than 100%, whereas red had the greatest effect in more severe conditions.

The color metrics used to quantify the differences between the misconverged character and the fringes also need to be considered. Four types of color metrics, the CIE (Commission Internationale de l'Eclairage) L*u*v*, the CIE L*a*b*, the Yu'v', and a weighted version of the L*u*v*, have been used to quantify the color differences of images presented on a CRT, with previous research suggesting the Yu'v' and the weighted L*u*v* to be superior in predicting image quality (Lippert, 1985; Sayer et al., 1990) of a CRT-presented image. The CIE L*u*v* and L*a*b* metrics were developed to quantify the difference between colors on a reflective or hardcopy display (Post, 1983). The Yu'v' and the weighted L*u*v* metrics were developed for emissive displays (Lippert, 1985). The formulae for these metrics are presented in Table 3. This research also examines the color metrics to determine which most closely correlate with perceived image quality.

TABLE 3: Color Difference Formulae

$$\text{Delta E } (L^*a^*b^*) = [(L^*_t - L^*_b)^2 + (a^*_t - a^*_b)^2 + (b^*_t - b^*_b)^2]^{0.5}$$

(CIE, 1959)

$$\text{Delta E } (Yu'v') = [((155/Y_M)(Y_t - Y_b))^2 + (367(u'_t - u'_b))^2 + (167(v'_t - v'_b))^2]^{0.5}$$

(Lippert, 1985)

$$\text{Delta E } (L^*u^*v^*) = [(L^*_t - L^*_b)^2 + (u^*_t - u^*_b)^2 + (v^*_t - v^*_b)^2]^{0.5}$$

(CIE, 1959)

$$\text{Delta E (weighted } L^*u^*v^*) = [((595/L^*_M)(L^*_t - L^*_b))^2 + (0.75(u^*_t - u^*_b))^2 + (0.36(v^*_t - v^*_b))^2]^{0.5}$$

(Lippert, 1985)

Variables

L, Y = Luminance, in the respective color spaces

a,b,u',v',u*,v* = color coordinates, in their respective color spaces

Subscripts

t - indicates "target" or "character"

b - indicates "background"

Constants

Y_M = greatest screen luminance (91.5 cd/m^2, measured or 100 by convention)

L_M = greater of the background and character luminance

Summary of Literature Review

Misconvergence in color CRTs may degrade the user's opinion of display quality and may reduce task performance. Further research is needed to determine how legibility is affected by misconvergence.

Besides misconvergence, other factors such as font type, luminance contrast, and chrominance contrast can influence CRT legibility. Since these factors vary in the use of CRT displays, their influence on legibility should be examined.

Research Goals

DeVilbiss (1990), Herb (1990), and Ansley (1991) studied the effects of misconvergence on several subjective and objective tasks. This thesis expands upon that work. The conditions used in this research are the same as those used in these three studies to allow for easy comparisons among the studies.

The major objectives of this research are to determine the effect of misconvergence on readability under several conditions:

(1) varying font type,

(2) varying misconvergence amount, and

(3) varying misconvergence type.

Further, luminance and chrominance contrasts are examined for their correlation with subjective image ratings. The four color metrics are compared to determine the best predictors of image preference.

METHOD

In this experiment, participants were asked to read textual paragraphs and then to select the word from each passage which was out of context. These paragraphs appeared on a shadow-mask color CRT screen and varied in font type, color, and amount of misconvergence. The participants were timed on their task completion and were asked to rate the image quality of the text.

Subjects

Ten individuals (five males, five females) from the university population volunteered and were paid to participate in this study. Before being asked to participate, each individual was screened for normal visual capabilities. A Bausch and Lomb Ortho-Rater, a set of Dvorine Pseudo-Isochromatic Plates, and the Vistech vision contrast test system were used to measure near and far acuity, color vision, and contrast sensitivity, respectively.

Equipment

In the experiment, a Zenith 286-LP personal computer was used to control the conditions presented to the subjects, as well as to record the data collected from each subject. The type and level of misconvergence were controlled by a custom-designed misconvergence generator used by DeVilbiss (1990). This generator created misconvergence by introducing a delay in one or more of the video signal lines.

Because of the nature of the misconvergence generator, the misconvergence used in this experiment was along only the horizontal direction. Since the electron beams sweep horizontally across the screen, a delay in the signal channel caused that beam to "fall behind" the other beams. The image on the screen will be misconverged, since the guns that receive the signal without a delay will be converged and the "lagging" beam will be "off target." The amount of delay in the signal is proportional to the misconvergence on the screen. The generator provided four channels to which any primary beam may be assigned. These channels were as follows:

Channel 0 - no delay in the signal

Channel 1 - 0 to 63 nanoseconds of delay in the signal

Channel 2 - 1 to 126 nanoseconds of delay in the signal

Channel 3 - signal removed.

Channels 1 and 2 allow programmable levels of misconvergence, while channels 0 and 3 are fixed controls.

The computer display was interfaced with a VGA video card which provided 640 by 480 addressable pixels, at a resolution of 0.28 mm full-width half-maximum amplitude

(DeVilbiss, 1990). The luminance of the primary colors presented on this display 17.5 cd/m^2 for green, 10.5 cd/m^2 for red, and 8.3 cd/m^2 for blue.

Experimental Conditions

Each participant received all of the experimental conditions. Participants were seated 60 cm from a Zenith computer screen. The computer screen and mouse input device were placed on the table in front of the participant. A forehead rest attached to the table allowed the participants to maintain a constant viewing distance. The ambient illumination was approximately 100 lux at the screen.

Task Description

Subjects were asked to perform two tasks in the experiment. One was a reading task and the other was a subjective rating task. Five practice trials were given to each participant to acquaint them with the Tinker Reading Passages, the tasks, and the concept of misconvergence. Prior to the practice trials, the participants were instructed on how to use the mouse input device.

The Tinker Speed of Reading Test was developed by M.A. Tinker in 1947 to study the effects of typographical and illumination variations on reading speed (Feldt, 1959). This test was later modified by Carver (Feldman, 1972). The test provides 500 short passages to measure reading time.

Each passage contains one word, the "spoiler," that is clearly out of context in the passage. Subjects are required to read the paragraph and identify the spoiler word as quickly as possible. The time required to find the spoiler indexes reading speed. The following paragraph is an example of a Tinker passage (the spoiler word in this instance is "car"):

> Marge was on the lake when the lightning started flashing, and she rowed back home as fast as she could because she was frightened that the car would hit her.

The Tinker passages were selected randomly and displayed in the center of the screen. The text was presented in mixed upper and lower case 9 X 14 serif and sans serif fonts, examples of which are shown in Appendix A.

Once the out-of-context word was located, the participant pressed the right button of the mouse device. A paragraph of outline blocks, each block corresponding to a word of the Tinker paragraph, then appeared on the screen. Using the mouse device, the subject selected the block which corresponded to the out-of-context word. When the participant depressed the mouse button after identifying the target word, the "block paragraph" was replaced by a screen displaying the following nine adjectives:

Best Imaginable (9)

Excellent (8)

Good (7)

OK (6)

Passable (5)

Marginal (4)

Poor (3)

Awful (2)

Worst Imaginable (1)

The nine-point scale was used by DeVilbiss (1990). Each adjective corresponds to a number between 1 and 9 with 1 being "worst imaginable" and 9 being "best imaginable." To choose one of these adjectives, a subject used the mouse device to position a cursor next to the desired word and pressed the left button.

After the subjective rating task, a screen displaying the word "READY" was presented. When the subject wished to begin the next trial, he or she pressed the left mouse button. Subjects were instructed that they may become fatigued during the experiment and if they wished to rest, they should do so upon arriving at a "READY" screen. This instruction prevented subjects from taking a break after beginning a timed trial (when the Tinker passage was displayed) and prevented erroneously large response times.

The experimental tasks were the same as described for the practice sessions. The participants read the Tinker passage, selected the target word, and rated the image quality. The time required to find the spoiler, the correctness of the response, and the rating of the image quality were recorded by the computer. The participants' instructions are presented in Appendix B.

Experimental Design

This study consisted of a three-factor within-subjects design, as shown in Figure 5. All factorial combinations were presented randomly to each participant, with each condition being replicated three times for each participant.

Dependent variables. The time required to identify the spoiler was measured in hundredths of a second. The computer measured this time, which was signalled by two depressions of the mouse button (at the start of a trial and upon finding the spoiler).

Based on the correctness of the subject's selection of the box representing the spoiler word, an error measure was recorded. If the response was correct, a "0" was recorded by the computer. If the response was incorrect, a "1" was recorded.

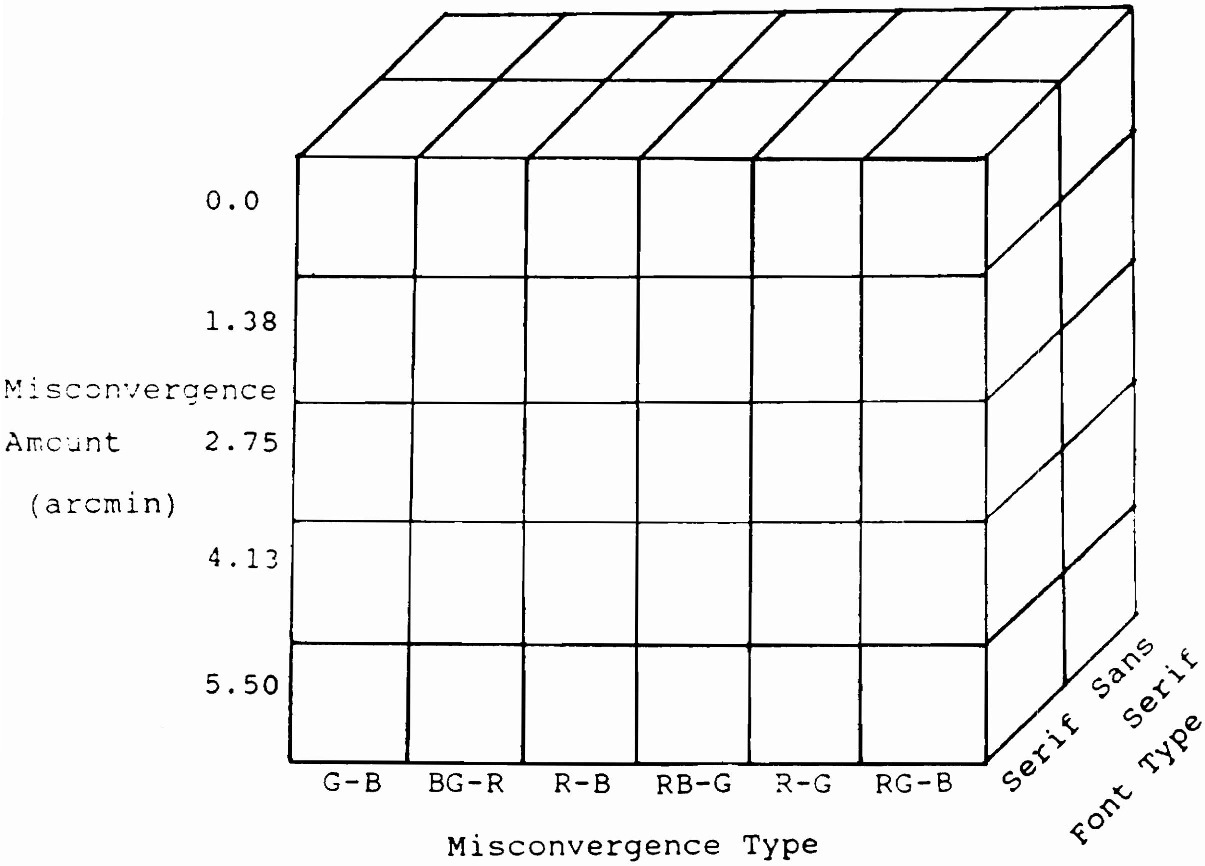


Figure 5: Experimental design.

The subjective quality of the image was the third dependent variable in this experiment. This rating was recorded as the number associated with the descriptive adjective from the previously discussed list.

Independent variables. Three within-subjects factors were used in this experiment: the misconvergence type (MT), the misconvergence amount (MAMT), and the font type (FT).

Misconvergence type (MT). The positions of the primary color beams as well as the arrangement of the electron beams define the color presented on the display. Six types of misconvergence, or six beam positions, were presented to each participant. These arrangements are shown in Figure 6. The electron beam positions are defined in terms of one or two reference beams and one varying beam. The letters R, B, and G represent the red, blue, and green beams, respectively. In the notation in Figure 6, the reference beam(s) precede(s) the hyphen and the varying beam follows it. Any primary not mentioned in the code is not used in creating the color: the corresponding beam is in the "off" state. The first three primary beam positions shown in Figure 6 are the primary pairs which create the secondary colors of text presented in this experiment: green and blue create cyan, red and green create yellow, and red and blue

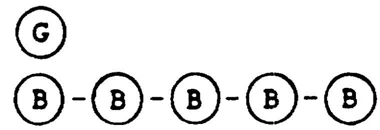
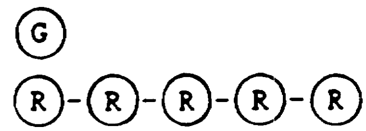
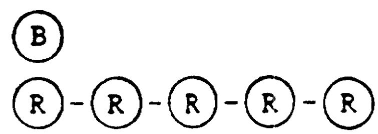
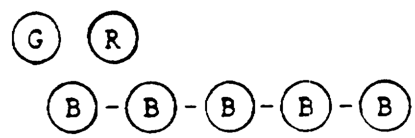
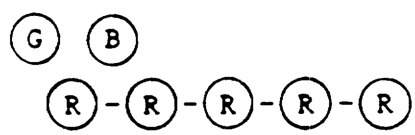
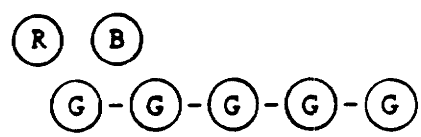
POSITION NUMBER	CODE	REFERENCE GUN(S)	VARYING GUN	
1	G-B	G	B	
2	G-R	G	R	
3	B-R	B	R	
4	GR-B	G R	B	
5	GB-R	G B	R	
6	RB-G	R B	G	

Figure 6: Misconvergence types: Primary beam positions

create magenta. The last three primary beam positions use all three beams, so the resulting text color is white.

Misconvergence amount (MAMT). Five amounts of misconvergence were tested in this experiment. All amounts of misconvergence were horizontal and were assigned as increments of the distance (P_H) between same color phosphor dots in adjacent triads. P_H for this display is 480 micrometers, and the electron beam travels across the screen at 10 microns per nanosecond. Thus, a 24-nanosecond delay in a beam results in a delay of one P_H, 0.48 mm (1.38 arcmin). The five misconvergence levels used in this experiment were between 0 and 2 P_H, as shown in Table 4.

Font type (FT). Two fonts were presented to all subjects: a serif font and a sans serif font. The two fonts were HALO104 and HALO107 (for the serif and the sans serif, respectively) of the Halo-88 graphics package, and they were selected because of their size, style, and spacing comparability.

TABLE 4: Experimental Misconvergence Levels

Misconvergence Type		Misconvergence Amount (arcmin)				
Number	Code	0.0	1.38	2.75	4.13	5.50
1	G-B					
2	R-G					
3	R-B					
4	RG-B					
5	BG-R					
6	RB-G					

RESULTS

Analysis

The three dependent variables - error, time and subjective rating - were analyzed using separate analysis of variance (ANOVA) procedures. The Greenhouse-Geisser correction (as described in Winer, 1971) was applied to all significant main effects and interactions. The results of these corrections are presented with the analyses of variance. Significant interaction effects were investigated using simple-effect F-tests and post-hoc comparisons.

The correlations between the dependent variables were calculated to determine the nature of their inter-relationships. Pearson correlations coefficients for the (1) error rate and response time, (2) response time and subjective ratings, and (3) error rate and subjective ratings were -0.141 ($p = 0.0005$), -0.133 ($p = 0.0011$), and 0.144 ($p = 0.0004$), respectively. These low correlations indicate three weak but significant trends. Subjects made fewer mistakes when they took more time to read a passage. Subjects preferred those passages which required less time to read. The more preferred passages resulted in greater errors.

To determine if significant dependent measures (time, errors, or subjective ratings) were correlated with the luminance or chrominance differences between different components of a misconverged character, stepwise regression analyses were performed. Four luminance predictors and six chrominance predictors were tested. Based on the results the stepwise regression analysis, the predictors (luminance/chrominance values) of the significant dependent measures were determined. The Pearson correlation coefficients between the predictors and the dependent measures were determined.

Error Rate

The ANOVA summary table for spoiler-selection error is shown in Table 5. The main effect of misconvergence type ($p < 0.05$) was significant. Figure 7 shows the average error rates for each of the misconvergence types, and Table 6 lists the results of the Newman-Keuls post-hoc analysis. Figure 7 shows RG-B misconvergence to result in the most errors, while G-B resulted in the fewest errors. Table 6 shows this difference to be significant. No other significant differences among misconvergence types were found.

Reading Time

Table 7 shows the ANOVA summary table for the reading time required for selection of the spoiler word. None of the main effects or interactions were significant.

TABLE 5: ANOVA Summary Table for Error Rate

Source	df	MS	F	p
Subjects (Subj)	9	1.945		
Font Type (FT)	1	0.000	0.00	1.0000
Misconvergence Type (MT)	5	0.077	2.96	0.0214
Misconvergence Amount (MAMT)	4	0.022	1.19	0.3320
FTxSubj	9	0.015		
MTxSubj	45	0.026		
MAMTxSubj	36	0.018		
FTxMT	5	0.019	0.54	0.7428
FTxMAMT	4	0.005	0.36	0.8354
MTxMAMT	20	0.012	0.59	0.9171
FTxMTxSubj	45	0.034		
FTxMAMTxSubj	36	0.014		
MTxMAMTxSubj	180	0.020		
FTxMTxMAMT	20	0.022	1.22	0.2396
FTxMTxMAMTxSubj	180	0.018		
Total	599			

Using the Greenhouse-Geisser correction (epsilon = 0.6265), $p < 0.05$ for MT.

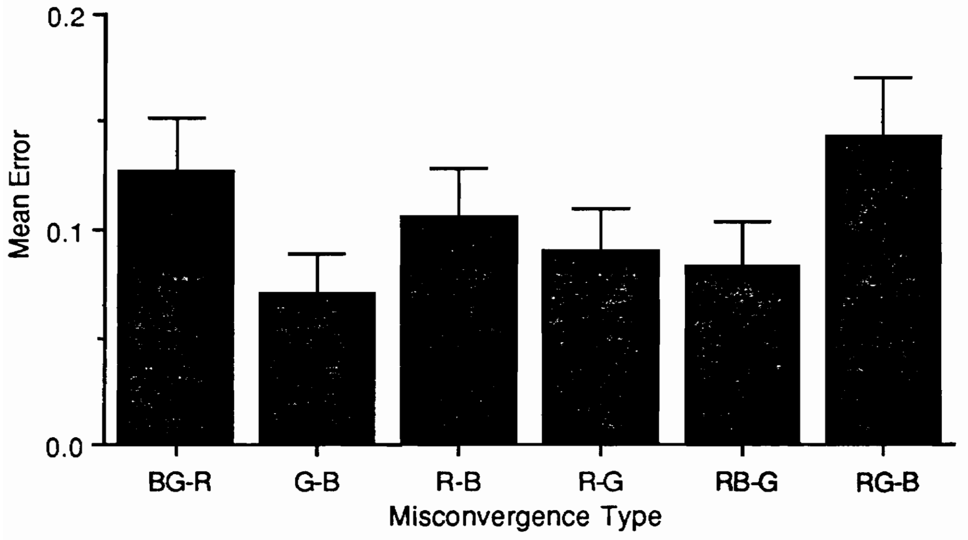


Figure 7: Main effect of misconvergence type on error rate. Error bars indicate + 1 standard error of the mean. (N = 100)

TABLE 6: Newman-Keuls Results for Misconvergence Type and Error Rate

Misconvergence Type	Mean Error	Group	
RG-B	0.14333	A	
BG-R	0.12667	A	B
R-B	0.10667	A	B
R-G	0.09000	A	B
RB-G	0.83333		B
G-B	0.07000		B

Groups with the same letter are not significantly different.

TABLE 7: ANOVA Summary Table for Reading Time

Source	df	MS	F	p
Subjects (Subj)	9	471.868		
Font Type (FT)	1	0.191	0.12	0.7420
Misconvergence Type (MT)	5	4.468	1.57	0.1892
Misconvergençe Amount (MAMT)	4	4.052	1.04	0.4006
FTxSubj	9	1.654		
MTxSubj	45	2.854		
MAMTxSubj	36	3.900		
FTxMT	5	8.187	1.75	0.1417
FTxMAMT	4	4.795	1.46	0.2337
MTxMAMT	20	2.853	1.03	0.4254
FTxMTxSubj	45	4.666		
FTxMAMTxSubj	36	3.277		
MTxMAMTxSubj	180	2.760		
FTxMTxMAMT	20	2.599	0.93	0.5561
FTxMTxMAMTxSubj	180	2.810		
Total	599			

Subjective Ratings

The ANOVA summary table for the subjective ratings is presented in Table 8. The main effects of font type ($p < 0.0001$), misconvergence type ($p < 0.001$), and misconvergence amount ($p < 0.001$) are significant, as well as the interaction between misconvergence type and misconvergence amount ($p < 0.001$).

Font Type. Figure 8 shows the average subjective ratings for the two font types. This figure shows that the sans serif font was rated higher than was the serif font, 4.94 versus 4.22. Since the interactions between font type and misconvergence type and font type and misconvergence amount are not significant, the sans serif font type preference is maintained for all misconvergence types and amounts.

Misconvergence Type. The misconvergence type main effect is shown in Figure 9, and Table 9 lists the Newman-Keuls post-hoc analysis results. As shown in Figure 9, the BG-R misconvergence was rated the highest, and the R-B misconvergence was rated the most objectionable. As Table 9 shows, the BG-R misconvergence condition results in significantly higher ratings than any of the other conditions. The G-B, RG-B, R-G, and RB-G conditions do not differ from each other and result in intermediate ratings.

TABLE 8: ANOVA Summary Table for Subjective Rating

Source	df	MS	F	p
Subjects (Subj)	9	19.314		
Font Type (FT)	1	78.482	105.22	0.0001
Misconvergence Type (MT)	5	22.606	21.09	0.0001
Misconvergence Amount (MAMT)	4	201.366	70.56	0.0001
FTxSubj	9	0.746		
MTxSubj	45	1.072		
MAMTxSubj	36	2.854		
FTxMT	5	0.090	0.37	0.8681
FTxMAMT	4	0.353	0.85	0.5003
MTxMAMT	20	4.577	10.50	0.0001
FTxMTxSubj	45	0.245		
FTxMAMTxSubj	36	0.413		
MTxMAMTxSubj	180	0.436		
FTxMTxMAMT	20	0.476	1.46	0.0990
FTxMTxMAMTxSubj	180	0.325		
Total	599			

Using the Greenhouse-Geisser correction, the following results were obtained:

MT	epsilon = 0.5298	$p < 0.001$
MAMT	epsilon = 0.8127	$p < 0.001$
MTxMAMT	epsilon = 0.0914	$p < 0.001$

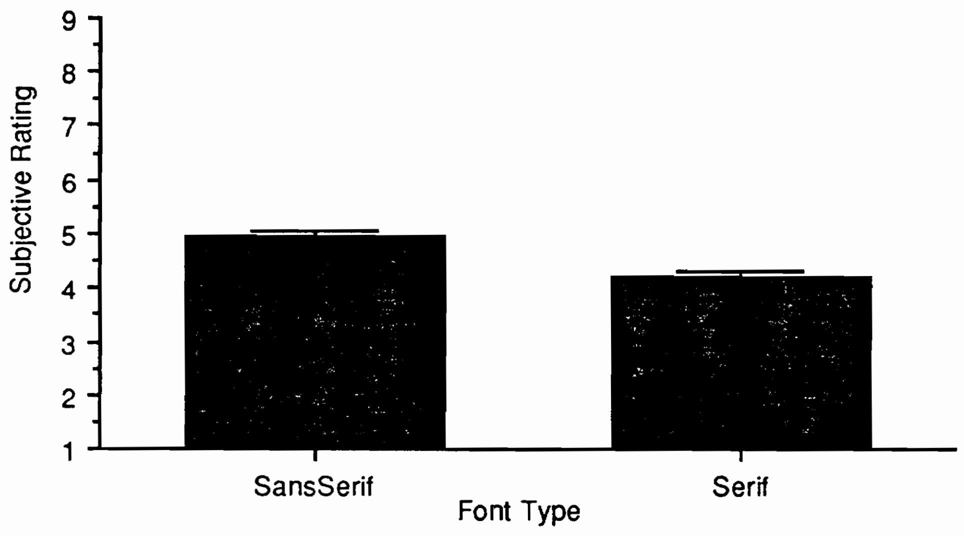


Figure 8: Main effect of font type on subjective rating judgements. Error bars indicate + 1 standard error of the mean. (N = 300)

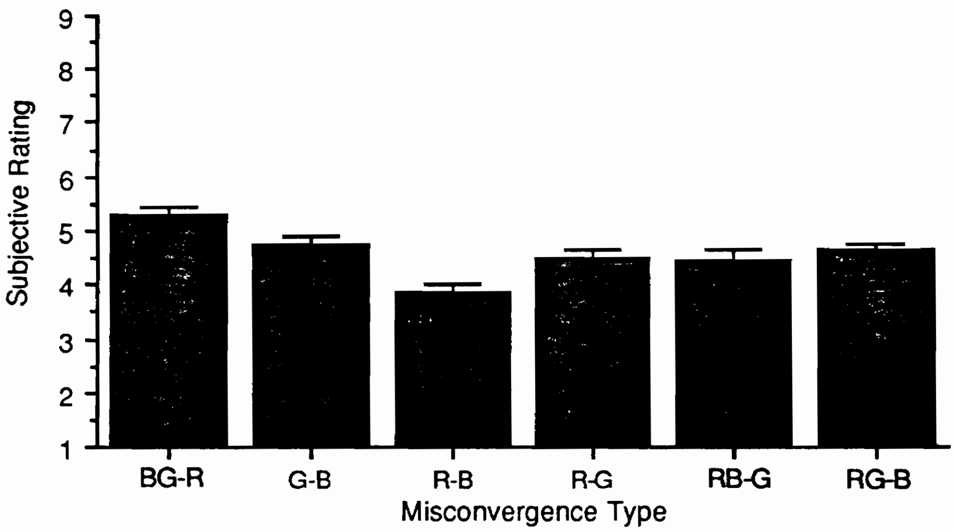


Figure 9: Main effect of misconvergence type on subjective rating judgements. Error bars indicate + 1 standard error of the mean. (N = 100)

TABLE 9: Newman-Keuls Results for Misconvergence Type and Subjective Rating

Misconvergence Type	Mean Rating	Group		
BG-R	5.31667	A		
G-B	4.74000		B	
RG-B	4.64667		B	
R-G	4.49667		B	
RB-G	4.44667		B	
R-B	3.85000			C

Groups with the same letter are not significantly different.

The R-B misconvergence type results in the lowest ratings, which differ significantly from all the other misconvergence conditions.

Misconvergence Amount. The significant main effect of misconvergence amount is plotted in Figure 10, and Table 10 shows the results of the post-hoc Newman-Keuls analysis. This figure and table show that as misconvergence amount increases, the subjective ratings decrease. The difference between the 0 and the 1.38 arcmin misconvergence conditions was not significant, but the differences among all remaining misconvergence amounts were significant.

Misconvergence Type and Amount. The final significant effect found with the subjective rating dependent variable was the interaction between misconvergence type and amount. Figure 11 shows the misconvergence amounts within each misconvergence type. In this figure, misconvergence types are plotted along the horizontal axis, subjective ratings are plotted along the vertical axis, and different types of bars represent the five different misconvergence amounts within each misconvergnce type.

Figure 11 shows that the subjective preference for the various misconvergence types are affected differently for increasing amounts of misconvergence. Some misconvergence

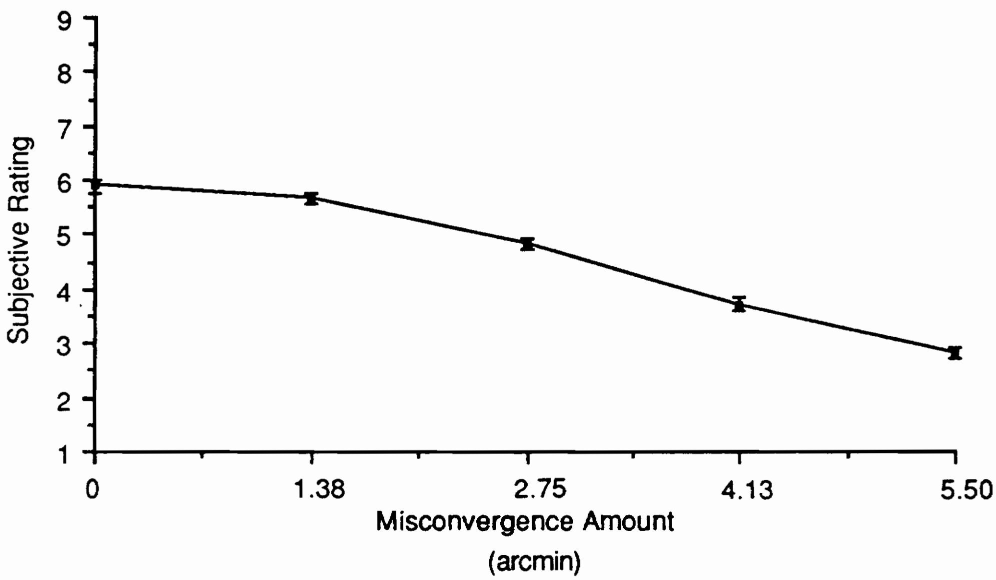


Figure 10: Main effect of misconvergence amount on subjective rating judgements. Error bars indicate ± 1 standard error of the mean. (N = 120)

TABLE 10: Newman-Keuls Results for Misconvergence Amount and Subjective Rating

Misconvergence Amount (arcmin)	Mean Rating	Group
0.00	5.88611	A
1.38	5.64167	A
2.75	4.84167	B
4.13	3.71389	C
5.50	2.83056	D

Groups with the same letter are not significantly different, $p > 0.05$.

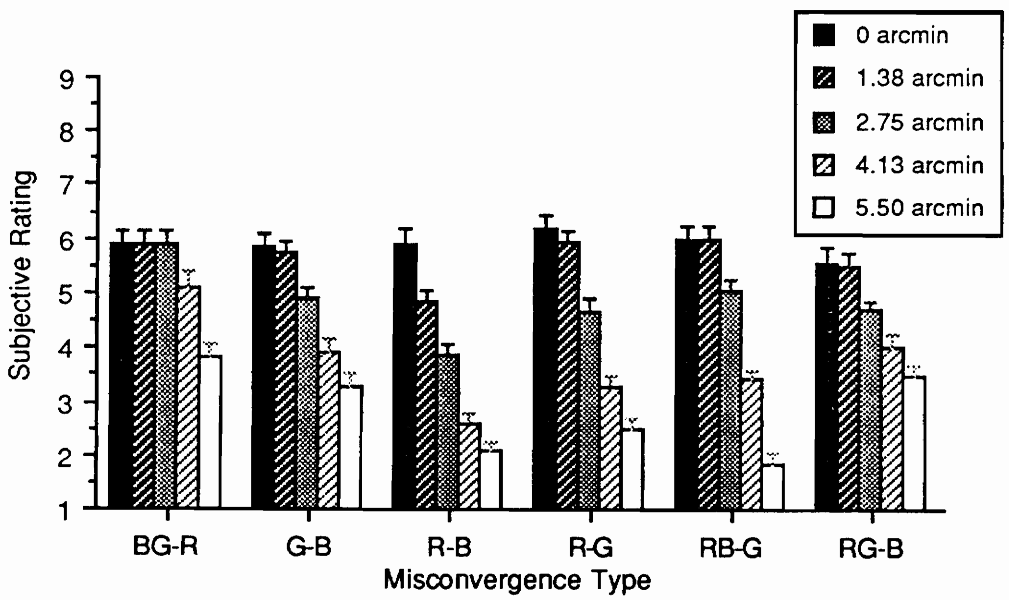


Figure 11: Interaction between misconvergence type and misconvergence amount on subjective rating judgements. Error bars indicate + 1 standard error of the mean. (N = 20)

types appear to "resist" the degrading effects, while other types degrade rapidly.

The results of the simple effect F-Tests, presented in Table 11, show that significant differences of misconvergence amount exist within each of the misconvergence types ($p < 0.001$). A Newman-Keuls analysis was conducted to determine the locations of these significant effects, and the results of this analysis are shown in Table 12. For all misconvergence types except the R-B, no significant difference exists between the 0.0 arcmin and the 1.38 arcmin misconvergence conditions. For the BG-R condition, this "no significant difference" group also includes the 2.75 arcmin misconvergence condition. For all misconvergence types, significant differences exist between the 2.75 arcmin, the 4.13 arcmin, and the 5.50 arcmin misconvergence conditions. Also, for all misconvergence types except the BG-R, significant differences exist between the 1.38 and the 2.75 arcmin conditions.

To determine the "acceptable" limits of misconvergence amount for each of the types, the range of subjective ratings for each type was found by subtracting the lowest average rating from the highest average rating for each type. This range was divided by 10, and the misconvergence amount required to result in this 10% decrease in ratings

TABLE 11: Simple-Effect F-Tests for MAMT at each MT for Subjective Rating

Source	df	MS	F	p
MT*MAMT*SUBJ	16	0.43590		
BG-R	1	46.24300	106.09	<0.001
G-B	1	50.58775	116.05	<0.001
R-B	1	70.85425	162.55	<0.001
R-G	1	74.69425	171.36	<0.001
RB-G	1	86.73450	198.98	<0.001
RG-B	1	41.21225	94.55	<0.001

df is corrected using epsilon = 0.0914

TABLE 12: Newman-Keuls Results of MAMT for each MT for Subjective Rating

BG-R Misconvergence

Amount	Mean	Group
0.00	5.88	A
1.38	5.88	A
2.75	5.90	A
4.13	5.10	B
5.50	3.83	C

G-B Misconvergence

Amount	Mean	Group
0.00	5.83	A
1.38	5.75	A
2.75	4.90	B
4.13	3.93	C
5.50	3.28	D

R-B

Amount	Mean	Group
0.00	5.90	A
1.38	4.83	B
2.75	3.87	C
4.13	2.57	D
5.50	2.08	E

R-G

Amount	Mean	Group
0.00	6.17	A
1.38	5.92	A
2.75	4.67	B
4.13	3.27	C
5.50	2.47	D

RB-G

Amount	Mean	Group
0.00	5.97	A
1.38	5.98	A
2.75	5.03	B
4.13	3.40	C
5.50	1.85	D

RG-B

Amount	Mean	Group
0.00	5.57	A
1.38	5.50	A
2.75	4.68	B
4.13	4.02	C
5.50	3.47	D

Within each misconvergence type grouping, misconvergence amount groups with the same letter are not statistically different, $p > 0.05$.

was calculated for each misconvergence type. Table 13 shows the results of this analysis. As would be expected, the more favorably a misconvergence type was rated over the misconvergence amounts, the greater the tolerable amount of misconvergence. Also, this table shows a total range of 0.5 to 3.2 arcmin, with the average acceptable limit being 1.7 arcmin.

Luminance and Chrominance Contrast

To determine if luminance and chrominance contrasts affected error rate or subjective preference, regression analyses were performed. Specifically, to test luminance as a predictor of error rate, the average errors for each misconvergence type were tested in a stepwise regression analysis with each of the following four predictors shown in Figure 12: (A) the stationary misconvergence fringe luminance, (B) the character luminance, (C) the misconverged primary luminance, and (D) the luminance difference between the two misconverged colors.

To test luminance as a predictor of subjective ratings, the average subjective ratings for each misconvergence type were tested in a stepwise regression with the four luminance predictors. The average ratings were determined using only the degraded misconvergence amount conditions (1.38, 2.75, 4.13, and 5.50 arcmin) for each misconvergence type.

TABLE 13: Maximum Acceptable Misconvergence Amounts for Each Misconvergence Type

Misc. Type	Misc. Amount (mm)	(arcmin)
BG-R	0.55	3.2
RB-G	0.35	2.0
G-B	0.29	1.7
RG-B	0.28	1.6
R-G	0.26	1.5
R-B	0.09	0.5

All values based on a 10 % decrease in the total subjective rating range for each misconvergence type.

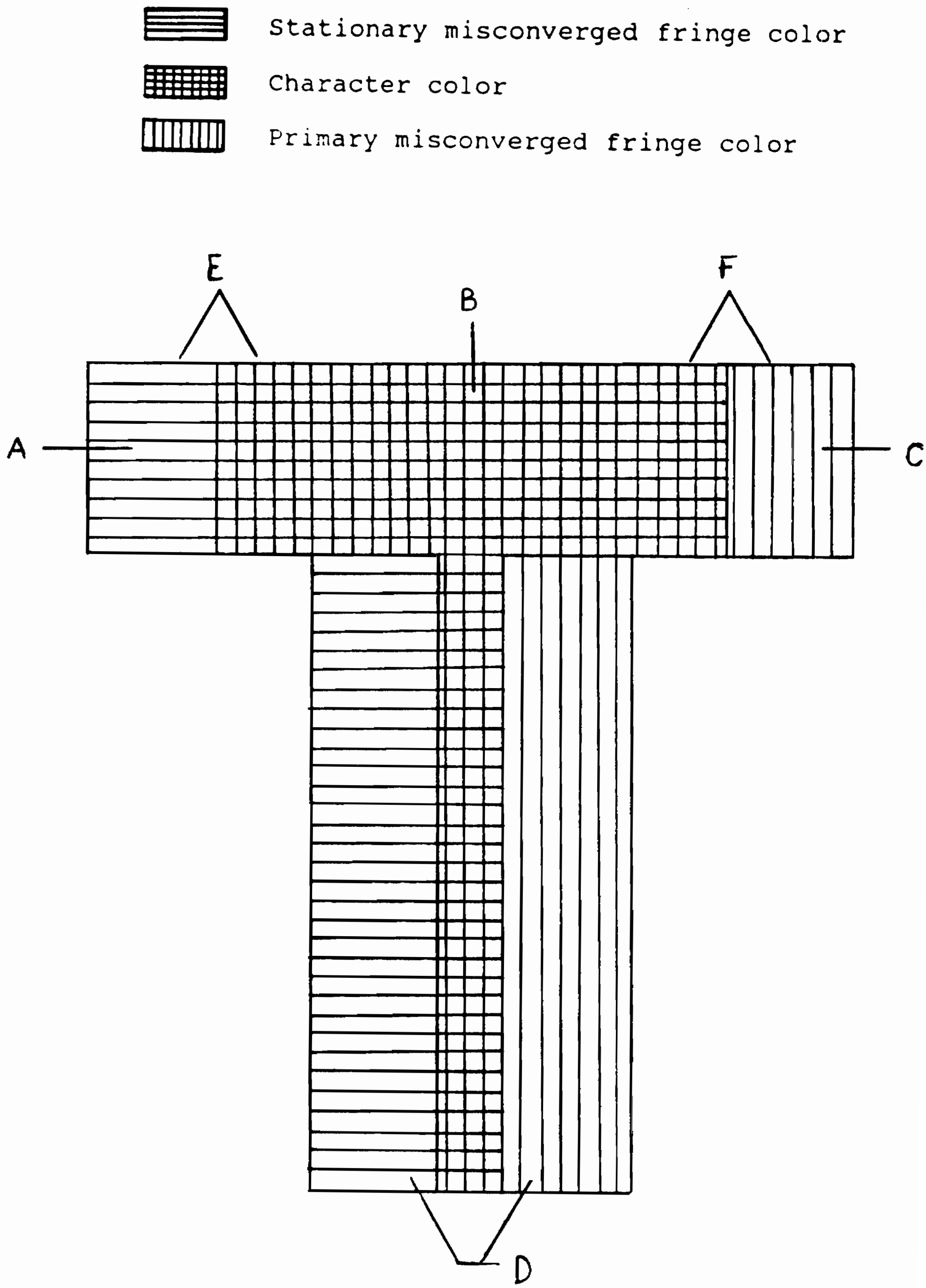


Figure 12: Misconverged character components analyzed in luminance and chrominance contrast analyses

The chrominance predictors of error rate and subjective rating were the following color differences: (A) between the stationary misconvergence fringe and the background, (B) between the character and the background, (C) between the primary misconvergence fringe and the background, (D) between the two misconvergence fringes, (E) between the stationary misconvergence fringe and the character, and (F) between the character and the primary misconvergence fringe. In addition, four color difference metrics (L*a*b*, Yu'v', L*u*v* and weighted L*u*v*) were used to determine the color difference predictors.

Based on the stepwise regression analyses, predictors for error rate and subjective rating were found, and are presented in Table 14. Any predictor not shown in Table 14 was not significant at the $p = 0.15$ level in the stepwise regression analysis. Error rate was found to be predicted by the luminance and chrominance (Yu'v' scale only) difference between the misconvergence fringes. Subjective ratings were found to be predicted by the luminance and chrominance (Yu'v' and weighted L*u*v* scales only) difference between the stationary misconvergence fringe and the background.

Once the predictors of rating were identified, the Pearson correlation coefficients between these predictors

TABLE 14: Luminance and Chrominance Predictors, Pearson Correlation Coefficients, and p Values for Error Rate and Subjective Rating

	Predictor		
	Luminance	Chrominance	
Dependent Variable	.	Yu'v'	weighted L*u*v*
Error Rate	D r = 0.718 p = 0.1083	D r = 0.687 p = 0.1319	
Subjective Rating	A r = 0.746 p = 0.0886	A r = 0.861 p = 0.0276	A r = 0.773 p = 0.0715

None of the L*a*b* or L*u*v* color differences were found to correlate with error rate or subjective rating at the p = 0.15 level.

D = the difference between the misconvergence fringes

A = the difference between the stationary misconvergence fringe and the background and the dependent variables were determined. These correlations and the p-values associated with them are also presented in Table 14. As this table shows, only the correlation between the stationary misconvergence fringe/background chrominance contrast (Yu'v' only) and the subjective rating was significant at an alpha level of 0.05. This finding shows that the higher the chrominance contrast between the stationary misconvergence fringe and the backgound, the higher the subjective rating.

These analyses also show that two color metrics, the L*a*b* and the unweighted L*u*v*, had no factors that correlated with subjective ratings. This color metric finding agrees with previous research (Ansley, 1991; Lippert, 1985; Sayer et al., 1990) in asserting the superiority of the weighted L*u*v* and the Yu'v' color metrics for predicting the subjective quality of CRT-presented images.

DISCUSSION

Error Rate

This research found that misconvergence amount and font type do not significantly increase errors made in the selection of spoiler words for the Tinker Speed of Reading task. The ANOVA results show misconvergence type to significantly affect error rate.

The Newman-Keuls analysis shows G-B misconvergence resulted in the lowest error rates and RG-B misconvergence resulted in the highest error rates. The difference between these conditions was significant. All remaining misconvergence types were not significantly different from each other, or from either of these two extreme conditions.

Herb (1990) found that the main effect of misconvergence type did not significantly affect error rate; however, the interaction of misconvergence type with target size was significant. Post-hoc analyses of this significant interaction revealed that only the smallest target size condtion (0.5 cm height at 60 cm viewing distance) contained significant differences among the misconvergence types. These misconvergence types were divided into three groups: the RB-G, the G-B, and the BG-R conditions resulted in the lowest error rates and the RG-B condition resulted in the highest error rate. The difference between these groups was significant. The third group, the R-G and the R-B misconvergence conditions, resulted in intermediate error rates, and did not differ significantly from either of the other two misconvergence type groups.

To compare Herb's (1990) results to the present study, the interaction effect of misconvergence type and target size must be understood. The smallest target size used by Herb was 0.5 cm (28.7 arcmin), the same size as the characters used in the present study. The other target sizes were 1.0 cm and 1.5 cm (57.3 and 86.1 arcmin, respectively). The misconvergence amounts (0 to 5.50 arcmin in the horizontal direction) were constant for all three target sizes in the Herb study. Thus, the larger the character, the smaller the percent or effective degree of misconvergence. While misconvergence amount did not affect error rate in the Herb study, perhaps the greater overall symbol distortion which occurs in the smallest target size does result in higher error rate. The finding that 0.5 cm characters were affected by misconvergence type in terms of error rate was found in this present study.

In comparing the effects of misconvergence type on error rate, the results of this study agree with those of the Herb (1990). The misconvergence type yielding the lowest error rate was the G-B in both studies. The R-G and the R-B conditions resulted in intermediate error rates in both experiments. Also, the RG-B condition resulted in the highest error rates in both studies.

The effects of misconvergence type upon error is similar between the present study and the Herb (1990) study. Additionally, the magnitude of error rates were similar. In this study, error rates ranged from 7 to 14 percent, while in the Herb study error rates ranged from 9 to 15 percent.

Reading Time

The results of this research agree with previous misconvergence research. In general, misconvergence does not result in a degradation of performance as measured by reading time. This finding is consistent with the research conducted by DeVilbiss (1990), Herb (1990), and Ansley (1991). The task, misconvergence types, and the misconvergence amounts tested in the DeVilbiss research were identical to the task, misconvergence types, and misconvergence amounts used in this study. DeVilbiss found that neither misconvergence type nor amount significantly affected the time required to complete the reading task. Ansley (1991) also found that neither misconvergence type or amount significantly affected the time required to complete the reading task. Herb (1990) used a visual search task but also found no misconvergence type or amount effects on task-completion time.

One misconvergence study (Snadowski et al., 1966) found that misconvergence amount significantly affected the time required for task completion. Unlike the other misconvergence research, this study required subjects to identify characters based on their color, not their shape. This finding does not conflict with the other research, since misconvergence distorts symbol color more than it distorts shape.

Subjective Ratings

In this study, the three main effects -- font type, misconvergence type, and misconvergence amount -- were all found to have a significant effect on subjective ratings. Further, the interaction between misconvergence type and amount was found to affect subjective ratings.

Font type. The sans serif font was rated consistently higher than the serif font. This rating was constant across all misconvergence types and all misconvergence amounts. This trend indicates a general preference for the sans serif font. If the rating difference were more pronounced at increasing amounts or particular types of misconvergence, then some tentative conclusions could be reached regarding the superiority of the sans serif font over the serif for maintaining image quality in misconverged conditions. However, the constant level of preference for the sans serif font indicates that this preference is unrelated to misconvergence.

Misconvergence type. The subjective ratings differed across misconvergence types. The highest ratings were associated with BG-R misconvergence, which was significantly different from all other misconvergence types. The lowest ratings were associated with R-B misconvergence, which also differed from all other misconvergence types. Intermediate ratings were associated with G-B, RG-B, R-G, and RB-G, which were not significantly different. The range of the subjective ratings was from 5.32 to 3.85.

These results differed somewhat from the results of previous studies. While previous studies (Ansley, 1991; DeVilbiss, 1990; Herb, 1990; Merrifield, Haakenstad, Ruggiero, and Lee, 1979; Robertson and Jones, 1984; Silverstein and Lepkowski, 1986) also found subjective ratings to vary across misconvergence type, differences in the ordering of these ratings exist among the studies. The factors underlying these observed differences must be examined for a better understanding of the effects of misconvergence type upon subjective preference judgements.

Differences in subjective ratings may be expected to occur if different criteria are established for the rating. Thus, the basis for the rating must be investigated.

Ansley (1991) tested the same reading task, misconvergence types, and misconvergence amounts that were tested in this present study. Ansley also varied screen luminance and tested additional misconvergence types. Of the misconvergence types that were identical to those used in this study, two significantly different groups were found based on subjective ratings. The higher-rated group consisted of the BG-R, the G-B, and the RG-B misconvergence conditions, and the lower-rated group consisted of the R-G, the RB-G, and the R-B misconvergence. While the Ansley (1991) study found fewer distinctions between the groups than did this present study, possibly due to the greater number of conditions in the Ansley study, the order of the ratings for the different misconvergence types was identical for the two studies.

DeVilbiss (1990) found the same general pattern of results as Ansley (1991) and the present study, with one conflicting result. DeVilbiss found BG-R misconvergence to be ranked intermediately, and Herb (1990), Ansley (1991), and the present study found it to be among the most preferred types. This unusual finding may be explained by the electron beams luminance proportions. While Herb, Ansley, and the present study tested the same electron beam luminance proportions (green, 50%; red, 30%; and blue, 20%), DeVilbiss tested different proportions (green, 69%; red, 23%; blue, 8%). The DeVilbiss findings that the most-preferred misconvergence types were G-B and RG-B, and the least-preferred types were R-G and R-B agreed with Ansley (1991) and the present study results.

Herb employed a visual search task and found different results from DeVilbiss (1990), Ansley (1991), and the present study. Herb found G-B and RG-B misconvergence types to be rated as highly objectionable, but these types were preferred in the other studies. Further, the R-G and the RB-G types were preferred in the Herb study, but were rated as highly objectionable in the other studies. The two other types of misconvergence, BG-R and R-B, were rated similarly in the Herb research and the other three studies.

The similarity of the findings among the present study, the Ansley study, and the DeVilbiss study suggest that task demands affect the preference for misconvergence type. Even though the luminance levels were different in this study and the DeVilbiss study, the tasks were similar and the results were similar. The Herb study and this present study used the same luminance proportions, but found different results, perhaps because of the differing tasks. For reading, people may prefer colors with a high luminance contrast relative to the black background such as white or cyan (especially noticable in the Ansley and DeVilbiss studies), and for the visual search task (of the Herb, 1990 study), they may dismiss this preference.

Merrifield, et al. (1979) found that yellow misconvergence was more objectionable than magenta. They did not test other misconvergence types. While these results may appear straightforward, this study employed a "contaminated" manner of obtaining the subjective ratings. First, subjects were asked to identify misconvergence, then they were asked to rate the misconvergence. Upon detecting misconvergence, subjects were likely to begin to give lower subjective ratings for the image. That is, subjects were asked to find a flaw in the image, then to rate the image. Naturally once they found a flaw, their ratings would decrease. The subjects were "set up" to rate the most easily detected misconvergence as most objectionable. While one may argue that the most easily detected misconvergence also is the most objectionable, DeVilbiss (1990) found that cyan misconvergence was the most-preferred type, yet it was not the least easily detected.

Robertson and Jones (1984) found no difference between symbols or text presented on a CRT in terms of misconvergence type or amount preferences. The only misconvergence type effect they did find was a preference for cyan misconvergence. This finding agrees well with the results of DeVilbiss, Ansley, and the present study, but the other findings of Robertson and Jones do not agree with other misconvergence research. The remaining six character colors tested by Robertson and Jones were found to be rated as equally objectionable. All other misconvergence research has shown distinctions among preferences for the various misconvergence types.

This unusual result may be due to the similarity of the colors tested in the study. Combining the primary colors of the electron beams in unique combinations results in only four distinct secondary colors (yellow, magenta, cyan, and white). Since seven character colors were tested, the electron beam combinations were not unique: the proportions of the primary colors were changed to create different secondary colors. For example, magenta is created by combining red and blue. Increasing the red beam component results in a color that is more pink than purple. The secondary colors used in this study may have been similar enough to one another that distinguishing among them was difficult.

Silverstein and Lepkowski (1986) found that misconvergence involving a green-varying beam was detected readily. Varying the red beam gave intermediate detection thresholds, and varying the blue beam resulted in the most difficult detection. This study did not test actual misconvergence (where the electron beams are improperly aligned) but increased the spot width of a primary color component of a secondary-color line. The subjects were not detecting misconvergence, but were looking for a primary color fringe around the secondary-color line. Further, no subjective ratings were measured. One might expect more easily detected misconvergence types to be less preferable, but the findings of DeVilbiss (1990) indicate that this expectation is not always fulfilled.

Silverstein and Lepkowski found that the green primary was the most easily detected, the red was the next most easily detected, and the blue was the least easily detected. Assuming the beam luminance proportions were similar to those used in other studies (DeVilbiss, Herb, Ansley, and the present study), these findings are easily explained by the luminance (and chrominance) contrast between the primary color and the background. Since the green beam had the highest luminance proportion, a change in the width of the green spot would be more readily detected than a change in the lower-luminance red and blue spots. Also, the red beam was of a higher luminance than the blue beam, so a change in the red spot size would be more easily detected than a change in the blue spot size.

Other explanations for the different findings between the misconvergence studies also should be considered. Herb (1990) suggested that the disparate findings may be due to luminance proportions of the electron beams used in forming the character colors. The more similar the character color luminance to the misconverged fringe luminance, the more annoying (and thus the lower the subjective rating) the misconvergence type. The results of Herb (1990), DeVilbiss (1990), Silverstein and Lepkowski (1986) and Merrifield et al. (1974) suggest this hypothesis is reasonable. However, this present study and the Ansley study used luminance contrasts (green, 50%; red, 30%; and blue, 20% of peak white luminance) that were identical to those used by Herb, and the results of the misconvergence type preferences between the studies were somewhat different. This difference may be due to the nature of the tasks in experiments.

Related to the idea that luminance contrast affects misconvergence preference is the idea that chrominance contrast affects preference. That is, color difference between the character color and the misconverged color fringe may affect preference judgements.

Misconvergence amount. This study found that increasing amounts of misconvergence reduced subjective ratings. Other studies suggest that approximately the same amount of misconvergence is unacceptable. Even the pattern for subjective ratings under conditions of varying amounts of misconvergence amount is fairly well-maintained across studies.

Ansley, Herb, and DeVilbiss found that misconvergence amounts of 1 to 2 arcmin (0.18 mm to 0.36 mm) were the maximum acceptable amounts: increasing misconvergence beyond those limits decreased subjective ratings. This study found the same pattern of results. For all misconvergence types (except the R-B), the difference between 0.0 and 1.38 arcmin misconvergence was not significant, but beyond 1.38 arcmin (except for the BG-R misconvergence), all other amounts were significantly decreased in their ratings. Further, the average amount of "tolerable" misconvergence, the amount required for a 10% decrease in subjective ratings, was 1.7 arcmin, with a range of 0.5 to 3.2 arcmin depending on misconvergence type.

Merrifield et al. (1979) found that misconvergence amounts were correlated with subjective ratings. Although subjects detected 0.13 mm (0.55 arcmin) misconvergence, only when the level reached 0.33 mm (1.4 arcmin) did 88% of the participants express subjective objections.

Robertson and Jones (1984) found that misconvergence amounts between 0.2 and 0.3 mm were acceptable to 90% of the participants. This study did not control the viewing distance, so the visual angle subtended by these guidelines is not known. If the subjects were approximately 60 cm from the screen, these guidelines correspond quite well with the previously mentioned studies.

Silverstein and Lepkowski (1986) found much higher levels (2.5 - 5.9 arcmin required for detection, suggesting that even higher levels would be required for a decrease in image quality) than other researchers. However, this study did not examine actual misconvergence, but measured spot width increases. The unusual findings of this study are explained by the nature of the task. Greater changes in spot size relative to the changes in color separation (that occur under true misconvergence conditions) must occur before the spot-size misconvergence can be detected.

Misconvergence type and amount interaction. This interaction showed that as misconvergence amount increased, image quality decreased. The rate at which the quality drops, however, varies depending on the type of misconvergence and the ratings associated with that type. For the low-rated R-B misconvergence, the ratings drop sharply with increasing misconvergence. For the high-rated BG-R misconvergence, the ratings drop sharply only beyond the 2.75 arcmin misconvergence condition.

In this study the difference between the 0.0 and the 1.38 arcmin conditions were not significant for all misconvergence types except R-B. This "no significant difference" group was extended to include the 2.75 arcmin condition for the BG-R misconvergence type. All misconvergence types had significant differences among the 2.75, 4.13, and 5.50 arcmin conditions. Except for the BG-R condition, all types had significant differences between the 1.38 and the 2.75 arcmin conditions.

Herb (1990) also found that the 0 and 1.38 arcmin misconvergence amount ratings were not significantly different from one another for the G-R, B-R, BG-R, and RB-G conditions. For the B-R condition, the 1.38 arcmin condition was found to be not significantly different from the 2.75 arcmin condition. For the G-B and GR-B conditions, all misconvergence amounts were significantly different from one another.

While the groupings of the specific misconvergence types vary between the present study and Herb (1990), the patterns are similar. For most types of misconvergence, 0.0 arcmin to 1.38 arcmin of misconvergence are acceptable, but beyond 1.38 arcmin, the subjective ratings decrease significantly. Rarely can a particular character color be misconverged up to 2.75 arcmin with no apparent loss in image quality ratings, but even for these "misconvergence tolerant" colors, beyond 2.75 arcmin the ratings decrease significantly.

Luminance and Chrominance Contrast

This study found that the higher the Yu'v' chrominance contrast between the stationary misconverged color and the background, the higher the subjective rating. All other luminance and chrominance contrasts tested did not significantly correlate with the dependent measures of subjective rating or error rate.

While this finding may at first appear to be counter-intuitive, one must realize that the stationary misconvergence color appears to the left of the misconverged character. Thus, a reader (since English is read left to right) encounters the border created by the left fringe before the actual character or the right fringe, and the contrast of this left-most character with the background may "establish" in the reader's mind the total contrast. That is, if the left-most fringe appears to be highly contrasted with the background, the reader may "see" the entire character as being of a higher contrast. This reading-sensitive phenomenon may help explain the different subjective preference findings of the misconvergence studies which used non-reading tasks.

CONCLUSIONS

This research, in agreement with previous work, shows that neither misconvergence type nor amount affects the time required to perform a task. The errors made during task performance also are not affected by misconvergence amount. Further, this research suggests that font type (whether sans serif or serif) does not affect reading time or error rate.

Misconvergence type, however, does affect error rate, so the use of white characters should be avoided, particularly when the blue beam is misconverged. If only one beam is misconverged, the character color resulting from the two non-misconverged beams should be presented. In general, the use of cyan characters will reduce error rate.

All misconvergence studies indicate that misconvergence amount and type significantly degrade image quality. This study in particular indicates a significant effect exists for font type, but this effect probably is unique to the fonts chosen. The amount of misconvergence required to significantly decrease subjective ratings of image quality is 1 to 2 arcmin. The types of misconvergence which are least preferable appear to be those in which small Yu'v' chrominance difference exists between the stationary misconverged color and the background. The greater the difference between the stationary misconverged fringe chrominance and the background, the more preferable the misconvergence type.

REFERENCES

Ansley, D.A. (1991). Effects of color CRT misconvergence and character luminance on reading performance and perceived image quality judgements. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

CIE (Commission Internationale de l'Eclairage). (1959). Colors of light signals. Paris: CIE Publication No. 2 (W-1.3.3).

DeVilbiss, C.A. (1990). An analytical model to predict detection threshold and performance data for misconvergence on a shadow-mask CRT. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Feldman, S.C. (1972). Basic reading rate scale. Journal of Reading, 15 (6), 451-452.

Feldt, L.S. (1959). Tinker speed of reading test. In O.K. Buros (Ed.) The fifth mental measurement yearbook (pp. 780-781). Highland Park, NJ: Gryphon Press.

Hale, S.L. and Billmayer, H.J. (1988). Use of color CRT's in aircraft cockpits: A literature review. (Technical Note 3-88). Aberdeen Proving Ground, Maryland: U.S. Army Human Engineering Laboratory.

Herb, I.M. (1990). Effects of color CRT misconvergence, target size, and nontarget density on visual search performance. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Human Factors Society, Inc. (1988). American National Standard for human factors engineering of visual display terminal workstations. (ANSI/HFS 100-1988). Santa Monica, CA: Human Factors Society, Inc.

Lehrer, N.H. (1985). The challenge of the cathode-ray tube. In L.E. Tannas Jr. (Ed.) Flat-panel displays and CRTs (pp. 138-176). New York: Van Nostrand Reinhold.

Lippert, T.M. (1985). Unitary suprathreshold color-difference metrics of legibility for CRT raster imagery. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Merrifield, R.M. (1987). Visual parameters for color CRT's. In H.J. Durrett (Ed.), Color and the Computer (pp. 63-81). Boston: Academic Press.

Merrifield, R.M., Haakenstad, L., Ruggiero, F.T., and Lee, J.N. (1979). Electronic flight instrument system (EFIS) misconvergence testing (Technical Report SYST-B-8764-10-79-085). Seattle, WA: Boeing Commercial Airplane Company.

Post, D.L. (1983). Color contrast metrics for complex images. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Robertson, P.J. and Jones, M.R. (1984). Subjective reactions to misconvergence on a colour display. Displays, 165-169.

Rogers, S.P. and Gutmann, J.C. (1983). CRT symbol subtense requirements. SID 83 Digest of Technical Papers, 166-167. Playa del Rey, CA: Society for Information Display.

Sanders, M.S. and McCormick, E.J. (1987). Human factors in engineering and design (6th ed.). New York: McGraw-Hill Book Company.

Sayer, J.R., Sebok, A.L., and Snyder, H.L. (1990). Color difference metrics: Task performance for multichromatic CRT applications as determined by color legibility. SID 90 Digest of Technical Papers, 265-268. Playa del Rey, CA: Society for Information Display.

Silverstein, L.D. (1987). Human factors for color display systems: Concepts, methods, and research. In H.J. Durrett (Ed.), Color and the Computer (pp. 27-61). Boston, MA: Academic Press.

Silverstein, L.D. and Lepkowski, J.S. (1986). The perception of color primary spatial distribution in color information displays. SID 86 Digest of Technical Papers, 416-419. Playa del Rey, CA: Society for Information Display.

Snadowsky, A.M., Rizy, E.F., and Elias, M.F. (1966). Symbol identification as a function of misregistration in color additive displays. Perceptual and Motor Skills, 22, 951-960.

Snyder, H.L. and Taylor, G.B. (1979). The sensitivity of response measures of alphanumeric legibility to variations in dot matrix display parameters. Human Factors, 21, 457-471.

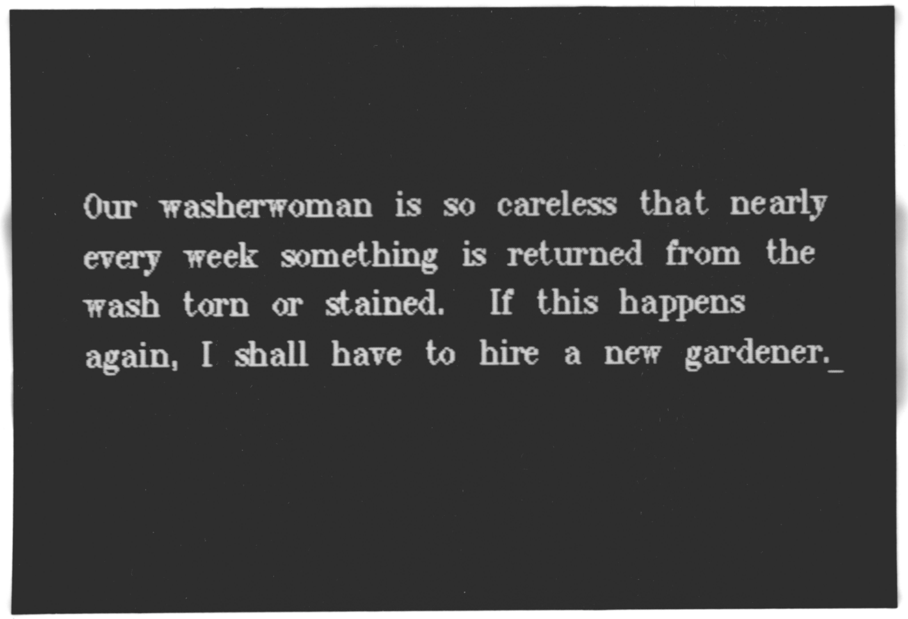
Snyder, H.L. (1988). Image quality. In M. Helander (Ed.), Handbook of human-computer interaction (pp. 437-474). New York: North Holland.

Walraven, J. (1984). Perceptual artifacts that may interfere with colour coding on visual displays. In C.P. Gibson (Ed.), Proceedings of the workshop on colour coded vs. monochrome electronic displays (pp. 13.1-13.11). Farnborough, England: Royal Aircraft Establishment.

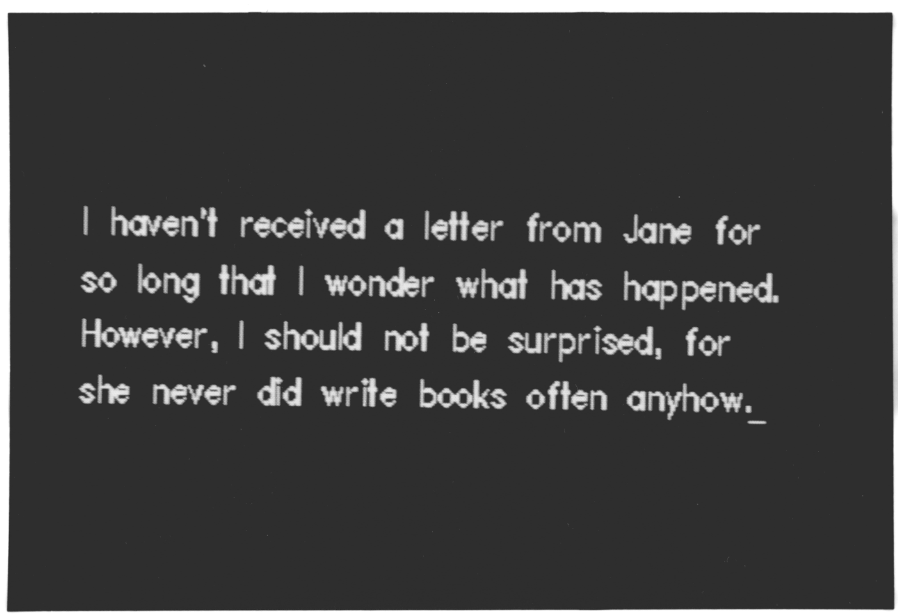
Winer, B.J. (1971). Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill Book Company.

APPENDIX A

SERIF AND SANS SERIF TEXT



Serif Text



Sans Serif Text

APPENDIX B

PARTICIPANTS' INSTRUCTIONS

INSTRUCTIONS

Pre-Screening Explanation

The purpose of this research experiment is to examine the effect of misconvergence on reading performance and subjective image quality. A color CRT screen produces images by combining the beams from three electron guns in different arrangements. Each gun activates a primary color: red, green, and blue. When these beams do not overlap properly, misconvergence results, and image quality degrades.

Your contribution to the study of this problem will be to view a display under various levels of misconvergence. Before participating in this experiment, you will be asked to take a vision test to make sure you will be able to see the experimental display clearly. If you have any questions regarding the purpose of this experiment, please ask them at this time.

Pre-Experiment Instructions

- You will see a screen that says "READY?" When you are ready to begin, press the left mouse button.
- You will then see a paragraph. Read the paragraph and look for the out-of-context word. When you find it, hit the RIGHT mouse button.

- A screen of blocks outlining the words will appear in place of the words. Move the mouse to move the cursor over the box which corresponds to the out-of-context word and hit the left mouse button.
- After you select the box, a list of adjectives will appear. Move the cursor next to the word which best describes the image quality of the paragraph you just saw. Press the left mouse button to pick this word.
- The "READY?" screen will appear again. Follow the same procedure.

Please work as quickly and as accurately as possible. You will be given some practice trials to get used to this experiment before actually starting. If you get tired or just want to take a break, go ahead and do so, but wait until the "READY?" screen appears before stopping. Remember, you always press the left mouse button except when you find the word in the paragraph.

Do you have any questions?

VITA

Angelia L. Sebok was born in Charleston, West Virginia on December 18, 1965. She obtained her B.S. in Engineering Science and Mechanics, Biomedical option and Psychology minor, from Virginia Polytechnic Institute and State University in December, 1988. As an undergraduate, Angelia was a member of the Virginia Tech Rescue Squad. In graduate school she joined Tau Beta Pi, Alpha Pi Mu, Phi Kappa Phi, and the Human Factors Society (student member).

Angelia L. Sebok