PERFORMANCE MEASURES AND SUBJECTIVE EVALUATIONS FOR TWO COLOR DISPLAYS

by

Cristina Christensen

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 Engineering and Operations Research

APPROVED:

H.L. Snyder, Chairman

J.S. Greenstein

B.H. Williges

October, 1985
Blacksburg, Virginia

ABSTRACT

PERFORMANCE MEASURES AND SUBJECTIVE EVALUATIONS

FOR TWO COLOR DISPLAYS

by

Cristina Christensen

The current study investigated the task performance and subjective preference for two color displays with differing image generation technologies, the standard cathode ray tube shadow mask (CRT) display and the newer liquid crystal/cathode ray tube (LC/CRT) display. Six subjects performed three different information processing tasks using each of the two color display technologies and expressed their display preference via evaluation questionnaires. Ambient illumination measurements were obtained to determine preferred conditions for each display.

A four-way factorial design was used to collect task performance data and ambient illumination preferences; performance data were collected as errors per unit task quantity for each of the task types. Subjective evaluations consisted of 20 five-interval bipolar adjective scales and a

forced choice rating on eight display parameters. An analysis of variance procedure and post-hoc Newman-Keuls analyses were employed in the analyses of the performance and subjective bipolar adjective scale data; the forced choice rating scales were evaluated using the Sign Test.

The task performance results indicate that neither display produced better task performance. The subjective data revealed mixed results; while the bipolar adjective scales indicate no differences between the two display technologies, the forced choice rating shows a preference for the LC/CRT display on some display parameters.

A significant difference between the two displays was demonstrated for ambient illumination preferences; the LC/CRT was viewed in greater ambient illumination than the CRT display.

ACKNOWLEDGEMENTS

The author thanks Dr. Harry L. Snyder for his guidance and patience throughout this project. Thanks also are due Dr. Joel Greenstein and for their support.

Appreciation is also given to Tektronix, Inc. for sponsoring this research. Also special appreciation is given to Dr. Gerald Murch and for their research suggestions and for serving as project monitors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The video display unit (VDU) is a common component in many office environments and is used for various information processing tasks (National Research Council, 1983). It is used to present information visually that is stored and processed with computers. Most VDU operators perform their job tasks on monochromatic displays; however, the use of color displays is increasing for both specific and general applications.

Various technologies can be used to generate a colored display image. Two of the technologies are the standard cathode-ray tube (CRT) shadow mask display and the newer field-sequential liquid crystal/cathode-ray tube (LC/CRT) display. The appearances of the visual images which are displayed on the two display screens differ as a direct result of the inherent differences in these image generation technologies.

Although both display technologies employ a cathode ray tube, the LC/CRT uses a rapidly switching liquid crystal shutter over a monochrome CRT. As alternate fields of the

CRT are sequentially addressed, they are presented in either a green or red state which the viewer perceives as green, red, or a visual integration of green and red, amber. Inherent in this technology is the phenomenon of sequential color availability. One of the concerns which arises from sequential color shuttering is the possibility of image "break-up", which is due to the visual system's ability to perceive separately the images of the two color filters as they rapidly alternate during color image production. The lack of perceptual integration may create the appearance of two separate color images when only one is desired.

A second inherent characteristic of the LC/CRT is its high contrast image. This high contrast results as the ambient light reflected by the monochrome CRT component of the liquid crystal display travels through the color and linear polarizers (shutter components) of the LC/CRT twice, while the light emanating from the CRT component passes through only once (Vatne, Johnson, and Bos, 1983). Ambient illumination, therefore, causes less interference with the displayed image on the LC/CRT than on the shadow mask CRT because it is "filtered out" by the color shutter device. This display feature provides a dark background which gives the appearance of almost infinite contrast. Although differences between background and object are necessary for perception of the displayed image, too great a contrast can result in the empty field myopia phenomenon described by Murch (1984). The resulting image can appear to float, which could prove to be a distraction sufficient to cause performance decrements or negatively influence subjective preference.

The different display characteristics and their interactions can result in varied compatibility with the human visual system and may contribute to changes in ocular and visual performance of the operator (Laubli, Hunting, and Grandjean, 1980; Mourant, Lakshmanan, and Chantadisai, 1981; Smith, Tanaka, and Halperin, 1984). The visual system incompatibilities may be expressed, for example, as operator complaints of irritated eyes, blurring of vision, and perceived displayed image instability or illegibility (see, for example, Crane; Dainoff, 1979; Haider, Kundi, and Weissenbock, 1981; National Institute for Occupational Safety and Health, 1981). The potential human visual system incompatibility with the LC/CRT display characteristics was the basis for the present study.

VISUAL DISPLAY PARAMETERS

"Visual Fatigue"

The interaction between the operator and the VDU has produced a broad range of complaints. Smith et al. (1984) listed three major areas of concern. These are (1) the VDU as a source of radiation, (2) the wide variety of somatic complaints of VDU operators, and (3) the fear of office automation. This thesis focused on somatic complaints as an indicator of possible operator visual system incompatibility with the displayed image. It has been asserted (National Research Council, 1983; Ostberg, 1975; Smith et al., 1984) that somatic complaints by VDU operators are neither qualitatively nor quantitatively different from the complaints expressed by other workers performing similar types of work without the use of cathode-ray tube displays. Hart (1976) cautioned that while VDU operators' complaints certainly should not be disregarded, they do need to be considered in perspective because any task requiring prolonged periods of focused attention in the same position combined with an inability to rest will produce feelings of fatigue regardless of whether the task object is a typewriter, manuscript, or video display screen.

Smith et al. (1984) pointed out that there are many factors and multiple interactions of these factors that are responsible for VDU operator complaints. These include the postural relationship of the VDU operator and the machine, VDU display characteristics which affect legibility of the presented information, workplace ambient illumination, type of information processing task (i.e., interactive or data entry), the operators' visual functioning, and uncertainty regarding job security and satisfaction that can accompany office computerization. The present study was concerned with those factors which potentially differ or interact differentially with the two display technologies. In particular, these are the VDU display characteristics, ambient illumination, and type of information processing task. Smith et al. (1984) cited display characteristics and workplace lighting as those factors responsible for VDU operator complaints.

Although the National Research Council (1983) concurred that the cause of VDU operator complaints is a multifactorial and possibly interactive process, it was less certain about correlating specific factors with visual complaints. However, there is agreement that the probable factors responsible for operator complaints are environmental (involving display characteristics and nature of the visual task), ocular (the operator's visual status), and constitutional (both the operator's physical health and emotional state). Stewart (1980), also cited visual factors as being of primary interest, as they are "...the most widely recognized cause of fatigue."

Ostberg (1975) categorized operator complaints into four symptom types. The first category is ocular symptoms, which include feelings that the eyes are uncomfortable, heavy, dry, burning, tender to the touch, and aching. The symptoms of ocular discomfort are possibly due to muscular fatigue, but the mechanism is unknown (National Research Council, 1983).

The second classification is visual symptoms, which are defined by Ostberg (1975) as forms of disturbed normal vision. The symptoms are uncomfortable but do not include physical perception of pain. Some examples of visual symptoms are difficulty in fixating or focussing on an object, seeing double, or seeing objects with color fringes. The first two symptom categories, ocular and visual, have been occasionally experienced by 50% of all VDU operators (National Research Council, 1983).

Systemic symptoms form the third set. This complaint category usually is the most common and includes headaches of all degrees of severity and locations as well as generalized body aches, particularly in the neck, back, and arms.

The final complaint category is behavioral symptoms, which are actions (voluntary or involuntary) taken to compensate for or correct problems in the work environment. An example of these behaviors would be arranging one's posture to make the visual task easier or adjustment of the lighting in the workplace by use of blinds, shields, hoods, or dark glasses.

It is speculated that ocular and visual symptoms have a direct relationship to the display characteristics (see, for example, Laubli et al., 1980; Smith et al., 1984). How these characteristics interact with the human visual system determines the degree of compatability and, therefore, directly determines the contribution the display characteristics make in producing operator complaint.

Snyder (1980) stated that the evaluation of visual displays on any parameter must be made within the context of the visual system of the human operator. Murch (1984) elaborated on the structure of the human operator-machine link by outlining three areas of emphasis as applied to information displays. These are anthropometric, sensory, and cognitive. Of particular interest is the sensory area, as this is most directly related to display characteristics. Murch (1984) defined the sensory area of the human-machine interface as involving "the attributes of products that output or input information to the human sensory systems." The emphasis here is on the visual interface.

Examples of display characteristics which are seen to affect the human visual system are generally agreed upon in the literature (see, for example, Dainoff, 1979; Murch, 1984; National Research Council, 1983; NIOSH, 1981; Ostberg, 1974; Stewart, 1980). The display characteristics most frequently discussed in reference to the quality of the image generated on the display screen are resolution, character legibility, luminance, contrast, and image stability. The research on each of these display characteristics is extensive. For our purposes each will be discussed briefly to include only those features pertinent to the two display technologies under consideration. The attribute of color will also be discussed because the present study is concerned with the evaluation of two color displays.

Luminance

Luminance is the weighted intensity of radiant energy visible to the human eye. The human eye, however, is not equally sensitive to the entire visible light spectrum. It is more sensitive to the middle or green portion of the visible spectrum than to the red and blue (long and short wavelengths, respectively) extremes (Snyder, 1980). The radiant energy emitted is, therefore, "weighted" by the changing sensitivity of the eye to the different wavelengths in the visible spectrum. This "weighting" process converts radiant energy into luminous energy which provides a measurement of luminance or the amount of light emitted from a display. The standard international unit for the measurement of luminance is candelas per meter square (cd/m^2).

As Stewart (1980) and Murch (1984) explained, the relationship between character luminance and background luminance may be a more important consideration than solely the luminance of the characters on the screen. Contrast between background and object is necessary for information to be perceived. An additional consideration is the amount of ambient light which affects the level of light adaptation of the eye and therefore the visibility of the display and the information presented. Contrast is dependent on non-display or environmental lighting levels which can interfere with the information presented on the display by reflections, glare, or reduction of contrast by the addition of ambient light to the characters and background.

Contrast

Murch (1984) defined contrast in terms of the difference in the amount of light emitted from the informational and non-informational areas of the display as well as the ambient light, which has an additive effect to the display luminance. This contrast difference is frequently expressed as a contrast ratio which is calculated using the luminances of the areas with the greatest and least intensity. Murch (1984) stated that expert opinion provides the guidelines for contrast ratio as a minimum of 3:1 and a maximum of 10:1. The draft ANSI standard on VDUs supports 3:1 as a minimum ratio. One can see the need for a minimum contrast ratio as detectability of information involves perception of the difference between character and background. As Timmers, van Nes, and Blommaert (1980) stated, the legibility of displayed text is affected by contrast; in particular, lowered contrast decreases the percentage of correctly recognized words and increases response time. The upper limit for contrast ratio was established to avoid a phenomenon known as empty field myopia (Murch, 1984). This condition occurs when the display background provides a contrast which is perceived as infinite. When the eye has no surface on which to focus or "be grounded," the eye assumes a resting focus which is usually not appropriate for the viewing task and images in the display appear to float or be suspended in space.

In summary, contrast is the difference between the most and least luminance on the display and can be greatly affected by ambient lighting of the surrounding workplace. It is possible to have too much contrast as well as too little for effective visual performance.

Character Legibility

The previous discussion of contrast was primarily concerned with what Murch (1984) calls large area contrast. He also stated, however, that small area contrast or the relationship of one character to its background is often of greater importance. The readability of the display is affected by the character size, shape, and spacing between characters (Bouma, 1980).

The concerns of what contributes to the legibility of written material were not introduced with the video display unit. Demilia (1968) attributed degraded reading performance in part to the typographical characteristics of the written document. The characters displayed on the video screen must fulfill the same requirements for text legibility as does a hard copy document (Starr, 1984). VDUs have the additional problems introduced by the transitory nature of text generation on a display screen (National Research Council, 1983). Therefore, the method by which the character is electronically generated on the screen determines its size, shape, intercharacter spacing, luminance, contrast, stability, resolution, and, therefore, its legibility.

Resolution

The literature does not achieve a consensus in defining resolution, but both Murch (1984) and the National Research Council (1983) specify resolution as the ability of the display to produce small details relative to the eye's capability to perceive and resolve those details. Simply, resolution may be thought of as the crispness of the image (IBM, 1979). The National Research Council (1983) did state that there is clear support for an improvement in performance with an increase in resolution. IBM (1979) attributes these findings to the electrical activity of the brain which, when stimulated by the visual system, is greater for images of high resolution than low resolution. The degree of image resolution is a display characteristic which can affect the readability of information presented on the VDU.

Image Stability

During the discussion of character legibility, some of the legibility constraints for hard copy documents were noted. An additional consideration in information display is image stability. The transitory nature of display generation is an inherent feature of the CRT. Production of a display image is achieved by electrically exciting phosphors which have a specified rate of decay for the trace. Therefore, the phosphors must be continually re-excited or refreshed in order to maintain a constant and stable image. When there is less than an optimum relationship between the refresh rate and the decay of the phosphor trace, the operator may perceive a flickering display (Murch, 1984). Stewart (1980) suggested that flicker can be reduced by either increasing the phosphor decay time or increasing the refresh rate. Both solutions have technical costs; respectively, these constraints are smear for moving images and bandwidth.

The perception of flicker is not solely a function of phosphor decay and refresh rate but is also affected by luminance, contrast, ambient illuminance, eye position during VDU viewing, and the individual variables of age, fatigue, and health. Perhaps of greater interest than the contributing factors to the perception of display flicker is the inability to establish any direct relationship between flicker and visual task performance (National Research Council, 1983). The negative subjective responses of annoyance and headaches to flicker can, however, indirectly reduce visual task performance. Therefore, stability of the image remains a design consideration to maintain operator comfort and thereby achieve increased performance.

Color

The availability of color for visual displays should not be the sole reason for its use; rather, specific justification is recommended. Color displays are used for three reasons: (1) to present information in a "realistic" manner, (2) the technology of the display only can produce color (e.g., light emitting diode), and (3) color displays allow the technique of color coding of information (Snyder, 1980). The addition of color to displays complicates the interaction between the human visual system and the display.

For example, the chromaticity of an image affects the brightness of that image (with luminance level being held constant), this being especially true of the red and blue portions of the visible spectrum (National Research Council, 1983). Red and blue are perceived as having greater brightness, for a given luminance, than green.

A second factor may affect the interaction between eye and display. Different colors are the result of different wavelengths of radiant energy and the non-color-corrected lens of the human eye must change in order to focus clearly these different wavelengths (IBM, 1979; Murch, 1984; National Research Council, 1983). This constant refocusing could lead to ocular discomfort or fatigue and is a confounding effect of color displays.

Stewart (1980) pointed out a third phenomena that may be due to the combination of the first two concerns of (1) apparent luminance differences and (2) the varying refraction of the lens for different colors. The combination of these two effects may cause the operator to perceive different colors as being at different distances in the display, thus producing a pseudo three-dimensional effect which is perceptually distracting (Stewart, 1980) and may, therefore, affect performance. This long-known phenomenon is generally called chromostereopsis.

As Snyder (1980) stated, there is no inclusive model that can be used to delineate all of the color display parameters. However, guidelines are available (Murch, 1984): (1) avoid the simultaneous display of highly saturated, spectrally extreme colors (e.g., blue and red); (2) avoid pure blue for thin lines and small shapes; (3) avoid small changes in reds, purples, and greens since sensitivity to color change varies across the visible color spectrum and can, therefore, be difficult to detect; (4) avoid red and green in the periphery of the display due to the peripheral insensitivity to red and green; and (5) complementary colors make good combinations for color displays (e.g., red and green, yellow and blue).

Illumination

Illumination in the workplace comes from a variety of sources. Lighting (fluorescent, incandescent, and natural) affects the image on the VDU screen. Typical illumination concerns are reflections, glare, and transient adaptation (sensitivity changes due to successive viewing of areas with different luminances) (National Research Council, 1983). Frequently workplace illumination is designed for the traditional desk top and hard copy documentation, and is not necessarily appropriate for the visual needs of the VDU operator. To compensate for the differing visual needs of the VDU operator, many devices (hoods, shields, filters) and operator-initiated work area rearrangements have been used. Often to reduce glare, reflections, and display "wash-out" (decreased contrast), VDU operators prefer a decrease in ambient illumination. Since an office environment can require both VDU and non-VDU tasks to be performed with the same ambient lighting, problems can arise. Illumination levels that provide an adequate viewing environment for paperwork tasks are often greater than optimal illumination levels for the VDU operator.

Task Types

The way the visual system interacts with the display is a function of the tasks the operator performs (National Research Council, 1983; Smith et al., 1984). The diversity of tasks involved in VDU work has been classified by the National Research Council (1983) into six categories, each with its corresponding visual emphasis (Table 1). The nature of the work done with the VDU influences the amount of time the operator spends viewing the display screen and thus the ability of the display characteristics to affect ocular comfort and task performance. Also, if the task requires constant visual shifting from display to a hard copy document, differences in all legibility parameters (contrast, resolution, and character legibility) can require excessive visual adjustment, thereby affecting the operator and performance.

Display Types

There are numerous types of displays used to present visual information (Snyder, 1980). The two display types used in the present study were the cathode-ray tube (CRT) color display and the liquid crystal/cathode-ray tube (LC/CRT) field sequential color display.

The CRT color display uses the standard three gun (red, green, blue) shadow mask technology to produce a color image. This technology is the same as that frequently employed for color image generation in commerical color television. Phosphors capable of producing red, green, or blue colors when electrically excited are placed in hundreds of color triads on the screen of the CRT. Each phosphor color is then excited by an electron beam from the appropriate "color gun", as dictated by the video signal.

Color image production is achieved with a different technique in the LC/CRT field sequential color display. The screen of the CRT used in this display contains a phosphor which, when electrically excited, emits only white light. A

TABLE 1

Some Video Display Terminal Task Categories

Task Category	Input Rate (Strokes/Min)	Visual Emphasis	Decision Making
Data entry	High	Source document (screen/copy/screen checks	Little
Data acquisition	Medium	Screen only	Some
Interactive communication	Medium/ intermittent	Screen only (some keyboard)	Some
Word processing	High/ intermittent	Screen/copy	Varies
Programming	Low/ intermittent	Copy/screen	Great
CAD/CAM	Low/ intermittent	Screen/copy	Great

multicolor display is produced by rapidly switching the liquid crystal off and on while synchronously presenting information on the CRT. In the "on" state of the liquid crystal the light emitted by the CRT passes through the liquid crystal cells unaltered and then travels through a linear polarizer oriented such that only green light may pass. When the liquid crystal is in an "off" state the orientation of the linear polarizer is changed by 90 degrees, thereby allowing only red light to pass onto the viewer. An amber hue is perceived by the eye when red and green emissions are perceptually integrated by the visual system. The rapid switching of red and green colors is made possible by the color shutter device.

Inherent in the LC/CRT is an improved contrast ratio between the display background and characters, especially in high ambient light environments. The physical configuration of the display causes ambient light reflected by the CRT screen to travel through the filter twice (going and coming) while light emitted from the CRT passes through the filter only once. This process reduces the interference from ambient light. A possible consequence of this display feature is improved viewability in high ambient illumination.

Resolution with the LC/CRT is also improved. A color triad shadow mask, as in the CRT, is not as finely spaced a configuration as can be achieved with a monochrome CRT. The resolution obtainable in a color CRT is dependent on the shadow mask component of this technology. Conversely, the LC/CRT resolution is determined only by the spot size of the electron beam. Therefore, greater image resolution is possible with the LC/CRT than with a conventional shadow-mask CRT.

Purpose

The purpose of this research was to investigate the VDU operators' task performance and subjective evaluations for the two color displays of differing technologies and therefore differing display characteristics. Predominately screen intensive tasks were used to increase operator sensitivity to these issues. Additionally, ambient illumination measures were obtained to determine subjectively optimal illumination for each display. These data are used to compare the two displays and to determine any significant differences between them which may result in an improvement or decrement of the typical office VDU task.

RESEARCH OBJECTIVES

This research had the following specific objectives:

(1) To determine if subjective differences exist between the standard CRT shadow-mask display and the liquid crystal shutter CRT display while performing a representative VDU task of relatively long-term duration.

(2) To determine if task performance differences exist between the two display technologies.

(3) To measure ambient illumination level preferences during the use of the two display technologies.

METHOD

Six subjects performed representative text processing tasks designed to investigate their responses to the two color displays. Three types of response variables were evaluated: task performance, subjective evaluation, and ambient illumination preference. Data were collected on task performance by measuring errors relative to the quantity of task material processed for each task type. The second response category, subjective evaluation of preference between the two displays, included ratings with bipolar scales and forced-choice ranking of display image features. Finally, subject-selected ambient illumination levels were obtained periodically for each display.

Experimental Design

The experimental design was a factorial 6 x 3 x 3 x 2 x 2 within-subject design. The independent variables in the study were three tasks, three 2-hr sessions per day (one before lunch, two after), two days per display, and two display technologies.

The dependent measures were divided into performance, subjective, and ambient illumination categories. The performance dependent measures were measured in terms of errors per unit task quantity for each of the specific task types. Scoring for the numbers of errors was task specific and dependent on quantity of task material processed. Task time was fixed at 35 min per task type per session for all task types.

After each 2-hr task session, subjective data were obtained on 20 five-interval, bipolar adjective scales (Crane; Dunn-Rankin, 1983; Epps, 1984). The areas of ocular comfort, visual comfort, and displayed image quality were evaluated by these bipolar scales. The 20 bipolar adjective word sets were selected as representative of sensations or complaints reported by VDU operators. Subjective data also included a final forced choice rating between the two displays.

Subjects

Six female subjects drawn from the university community served as volunteer subjects. Their mean age was 22.8 years. All subjects were experienced touch typists and had previously used a VDU for word processing tasks; in addition, subjects had minimal experience with color displays.

All subjects were required to read and sign an informed consent form prior to participation in the study (Appendix A). Subject-initiated questions were answered, provided the answer would not potentially bias the experimental results. Subjects were paid six dollars per hour for their participation.

<u>Task</u>

Each experimental session consisted of three tasks which varied in screen content and visual emphasis. Task input information was presented either on the screen or the hard copy document.

The first task consisted of editing a variety of documents (Appendix B). This task was selected because it is visually intensive, occurs in many real-world situations, and allows for a readily quantifiable measure of performance (Gould and Grischkowsky, 1984). The original text copy was presented on the display screen. The corrections to the original document were presented on a hard copy of the original text. The subject viewed the edited hard copy of the original document and made the indicated corrections to the copy on the display screen. The corrections involved punctuation and insertion/deletion of characters to correct spelling within the text.

The second task (Appendix B) required the subject to visually search the display screen. Blocks of 80 numbered 6-letter nonsense words were presented on the display screen (Mourant et al., 1981). Subjects were asked to visually search the display screen and locate a specific letter. The information presentation area of the display screen was split horizontally, thereby producing two separate information areas or windows on the display screen. The top portion of the display contained the visual search task and the bottom window was used to record the subject's number selections corresponding to the 6-letter nonsense word which contained the specified letter. There were six letters that subjects were instructed to search for: C, B, S, R, H, D. The presentation of these six letters in this fixed sequence was identical for both display types.

The third task involved data entry of assorted text material (Appendix B). The text entry passages varied in length and composition complexity. For this task all material was presented to the subjects on the video screen in a split display or windowing format. The screen was split horizontally with the top half displaying the text to be typed and the bottom portion of the screen used to accept the subject's typed input. The subject, therefore, read from the top of the display screen and then typed that material onto the bottom of the screen.

The second and third tasks were screen intensive involving either text entry with the copy source being the display screen or a search task consisting of the location of a specific alphanumeric imbedded in random lists of letters presented on the screen. The first task was less screen intensive as it consisted of text editing from a hard copy document. The visual emphasis was varied to determine if this would interact with the two display technologies.

Sufficient task material was generated for each task type based on pilot studies and a hypothetical subject able to touch type at 100 wpm. To assure that subjects would not deplete the supply of task material prepared during the experimental trials, 50 percent more material of each task type was prepared than was expected to be accomplished by the hypothetical subject. This system proved to be adequate in preventing insufficient task material for presentation or repetition of task material.

Procedure

Figure 1 outlines the experimental sequence.

Pre-experimental session. Subjects were tested for compatibility with the SRI Dual Purkinje Image Eye Tracker as this study was conducted in conjunction with another experimental effort. This device necessitated that subjects

Preliminary Evaluation and Instructions

30 min	Pre-test screening	Eye Dominance Visual Acuity Color Normalcy Eye Tracking System Compatibility
10 min	Consent Form	
90 min	Text Processor Instructions/ Sample Tasks (Monochromatic Displays)	

Experimental Procedure (Daily)

8:30 Snellen Acuity; Near, Far
8:35 Contrast Sensitivity
8:38 Dark Focus of Accommodation
8:45 Accommodation Tracking
8:55 Adjustment of Workplace (Chair, Lighting, Work Surface, Contrast)

9:00 Begin VDU Tasks

Session I

- 0 Instructions
- 35' Text Editing Copy-Screen
- 5' Break at Workstation
- 35' Search Task Screen-Screen
- 5' Break at Workstation
- 35' Text Entry Screen-Screen
- 5' Subjective Evaluation
- 15' Acuity
  Contrast Sensitivity
  Dark Focus of Accommodation
  Accomodation Tracking

11:30 Lunch

12:30 Acuity

Contrast Sensitivity

Dark Focus of Accommodation

12:45 Session II

2:15 Session III

4:45 Questionnaire

(Asthenopia and Other Complaints)

Figure 1. Experimental procedure

used in this study not wear corrective lenses and that they have sufficiently large pupils during dark adaptation and an adequate accommodation response to a moving target that they could be tracked by the optometer system. Each subject was then screened for normal color vision using Dvorine Plates and for visual acuity (20/22, minimum, far and near, uncorrected) using the Bausch and Lomb Orthorater.

Subjects were required to demonstrate that they were touch typists. This indicated that the subject's visual attention was directed primarily at the hard copy document or the VDU screen and not the keyboard of the VDU. Upon successful completion of the screening process, a second pre-experimental session was devoted to instructions for the use of the SAMNA III text processor. During this session the basic text processing commands were introduced to the subjects. These included cursor movements, character insertion and deletion, paragraph indentation, and underlining. The subjects were allowed to practice the necessary text manipulation commands to perform the sample tasks provided. The sample tasks required the same activities as did the experimental sessions (see Task subsection). Subjects were allowed to practice until they were able to correctly complete the sample tasks without assistance from the experimenter. The subjects were then scheduled for the experimental sessions.

Experimental sessions. A brief review of the text processing commands was provided at the beginning of the experimental sessions. At the beginning of each 2-hr experimental session adjustment of the workstation was performed (e.g., positioning of chair, document holder, and keyboard). Also at the beginning of each 2-hr session the ambient lighting conditions of the work area were adjusted according to subject preference and recorded (see Ambient Illumination subsection). Each subject was then asked to perform three tasks, each of 35-min duration with 5-min breaks between tasks, in a fixed presentation sequence. The 5-min breaks between tasks allowed for experimental setup of the next task type. This sequence of events comprised one 2-hr session. Subjects performed three sessions each day (one before lunch and two after) for two days on each display for a total of four days. Three of the subjects used the CRT during the first two days, and the other three subjects used the LC/CRT system first. Following each 2-hr session a subjective evaluation consisting of 20 bipolar adjective scales was offered for subject rating. When both displays had been tested, a final evaluation which ranked the two displays by forced choice was obtained.

At the end of the four days an explanation of the research was provided and any further questions were answered. Subjects were financially compensated for their participation at the conclusion of their experimental sessions.

Subjective Evaluation

The questionnaire contained two evaluative formats: bipolar adjective scale rankings and forced choice ratings. The bipolar adjective scales were based on the method developed by Osgood, Tannenbaum, and Suci (1957), which measures reactions to stimulus words and concepts in terms of ratings on bipolar scales defined by contrasting adjectives at each end of the scales (Heise, 1970). These scales measure both directionality and intensity of the subject's reaction (Heise, 1970). The ratings can be combined to analyze the person's feelings on particular issues. This highly generalizable technique of measurement has no standard scales; rather, it depends on the purposes of the research (Osgood et al., 1957).

The selection of scales to be used is based on their relevance. The scales should relate meaningfully to the concepts being presented and make possible the distinctions under consideration. The literature review indicated that the concepts of interest were visual comfort, ocular comfort, and image stability. The literature also indicated the 20 bipolar adjective word sets that were representative of the sensations or complaints reported by the VDU operators.

The positions on the scales were specified with adverbial quantifiers (e.g., VERY CLOSELY, MODERATELY CLOSE) to improve the process of differentiation. Each scale interval position was coded numerically. To aid subjects in the attitudinal discriminations, five equidistant scale intervals were formulated. Instructions for the bipolar adjective scales followed traditional guidelines (Osgood et al., 1957).

The three concepts presented -- ocular comfort, visual comfort, and image stability -- provided a reference or context for the subject's response (Appendix C). The bipolar scales used to assess ocular comfort asked the subjects to indicate how their eyes felt on the following scales: GOOD/BAD, TIRED/RESTED, COMFORTABLE/UNCOMFORTABLE, PAINFUL/PAINLESS, IRRITATED/NON-IRRITATED, and RELAXED/TENSE. Subjects were asked to indicate how their visual related to the displays by responding to the following scales: CLEAR/BLURRY, EASY TO FOCUS/HARD TO FOCUS, ABNORMAL/NORMAL, STABLE/VARIABLE, COMFORTABLE/UNCOMFORTABLE, and TYPICAL/UNUSUAL. Evaluation of image stability used the scales of EASY TO SEE/HARD TO SEE, BLURRY/CLEAR, INTACT/FRAGMENTED, VARIABLE/STABLE, and ANCHORED/FLOATING.

These scales required that subjects make absolute judgments in these three areas. The subjects were familiarized with the scales at the beginning of the experiment. The adjective word sets were reviewed to clarify their meaning if necessary. This portion of the evaluation questionnaire was presented to subjects for their response after each 2-hr experimental session.

The second portion of the evaluation questionnaire was a forced choice rating between the two display technologies. The subjects were asked to make a relative judgment by selecting which display was preferred in eight display feature categories: CLEARNESS, COLOR QUALITY, EASE OF VIEWING, CHARACTER LEGIBILITY, VISUAL COMFORT, IMAGE STABILITY, CONTRAST, and OVERALL. This portion of the evaluation was presented to subjects at the conclusion of the study.

Ambient Illumination

Workplace ambient illumination was provided by both fluorescent and incandescent light sources. The fluorescent light source was overhead and direct with two pairs of twin tubes capable of producing 473 lux measured at the center of the work area. The incandescent light source was indirect and consisted of two 150-watt flood lamps on a rheostat control. The variable rheostat control allowed for a maximum of 237 lux, also measured at the center of the work area. The two light sources combined afforded an ambient illumination of 710 lux. A separate incandescent lighting source was provided for hard copy document illumination. This light source had adjustability for both direction and intensity to provide optimal illumination for the hard copy documentation and to reduce glare and reflection on the VDU screen.

Subjects were allowed to select any ambient illumination type they desired, either fluorescent, incandescent, or both. The incandescent light source also allowed subjects to select the intensity level. Subjects were also directed to position the document light for their preference and to select its intensity level. The adjustments were performed at the beginning of each 2-hr experimental session.

Measurement of the illumination levels was made immediately following subjects' lighting preferences and was obtained at two workplace sites: in the center of the work site (ambient illumination), and directly at the display (display illumination). Orientation for measurement of the ambient illumination was determined by the physical dimensions of the work area. Display illumination was measured at the center of the display. Measurements were obtained with a Sekonic illuminometer in footcandles and then converted to lux.

Apparatus

Display systems. The CRT was a Sony 8-in. diagonal cathode-ray tube. It was a red, green, blue, in-line shadow-mask monitor.

The LC/CRT was a Tektronix, Inc. prototype 8-in. diagonal field sequential liquid crystal cathode-ray tube display.

The two displays were housed in custom-built, identical cabinets to avoid any bias or preference based on cosmetic differences.

Both monitors were driven by an IBM Personal Computer, Model AT, with the following memory capabilities: 256 Kb of RAM, a 20-Mb Winchester disk drive unit, and two 320/360 Kb dual density floppy disk units. The word processing software package in the experimental tasks was SAMNA III, produced by SAMNA Corporation.

The SAMNA software package allowed two colors of text (text and underlined text) to be displayed simultaneously. Underlined text, although always visually displayed, was randomly placed thoughout the text of each task type. Three color combinations of two colors each, one color for text and another color for underlined text, were selected. The color combinations were: 1) green text and red underlined text, 2) green text and yellow underlined text, 3) yellow text and red underlined text. All three color combinations were used in each experimental session.

Data Acquisition System

The IBM PC stored all information and responses to task commands for each subject. The performance, illumination, and subjective data were reduced to computer files and were analyzed using standard packaged statistical programs (SAS Institute, 1982).

Each task type and subset of the task were stored in separate computer files. For example, Task 2, visual search, had a separate computer file for each of the six specified letters. The separation into task subset files reduced experimental session setup time and would have reduced file regeneration time in the event of a hardware or software failure.

Data Reduction

Task 1, text editing, was scored by comparing a master edited hard copy document to the subject's edited computer file. An initial scan of the subject's file was performed using the spell checking software capabilities of the SAMNA III word processing package. This was followed by a visual inspection of the subject's displayed text editing file to match margins, indentations, and line length to the master text editing hard copy document. Discrepancies by word, existing between the master document and the subject's computer file, were scored as errors. Quantity of material accomplished was determined by the number of lines of task material edited for each 35-min session.

Task 2, visual search, was scored by comparing a master listing for each specified letter of the numbered nonsense words containing that specified letter. Discrepancies between the master list of numbers indicating which nonsense words contained the specified letter and the subject's number listings were scored as errors. This process was accomplished by visual comparison between the master listing hard copy document and a hard copy printout of the subjects' responses from their computer file. Quantity of material was determined by the number of possible specified letters that could have been located within the task material accomplished during the 35 min experimental session.

Task 3, text entry, was scored in a manner very similar to Task 1. Quantity of material accomplished was also determined by an identical procedure: counting the number of lines of typed material entered during the 35-min experimental session.

Data Analysis

Statistical analyses were conducted to determine if any of the data measures were significantly affected by the independent variables or their interactions. Task performance data were scored differently for each task type due to the inherent task differences found in editing, searching, and typing. Therefore, the analysis of the task performance data was partitioned by task. An analysis of variance was performed for each task type. All significant effects and interactions were further tested for significance, where appropriate, using post hoc Newman-Keuls tests.

The data were entered into the computer files in three categories: the number of errors, the quantity of task material completed, and the combination of number of errors and quantity of material completed. Error rates were then computed by dividing the number of errors by the quantity of task material completed per standard time unit.

The separation of the data into three categories for analysis was done so that number of errors and quantity of task material completed could be evaluated separately. Analysis of the error rate provides no measure of the components of error rate, although it produces a useful composite score. Also, by separating the error rate components one can determine how each of the components is affected by the independent variables.

RESULTS

Task Performance

Task 1: Text editing. Analyses of variance (ANOVAs) were performed to determine the effects of the independent variables on errors, quantity of task material accomplished, and the error rate (Tables 2, 3, and 4). The interaction of Display x Session was statistically significant for error and error rate ($F = 5.98$, $p = 0.0196$ and $F = 6.60$, $p = 0.0149$, respectively). Post-hoc Newman-Keuls analyses performed for each display comparing Sessions 1, 2, and 3 (Table 5) indicated that the LC/CRT showed no significant differences in number of errors across sessions while the CRT had a significant increase in errors (see Table 5) between Sessions 1 and 3.

The quantity of task material accomplished during this task was not significantly affected by any of the experimental variables (Table 2). Thus, the significance of the Display x Session interaction for error rate (Table 6) can be attributed to the effect of number of errors rather than differences in quantity of material.

TABLE 2

ANOVA Summary Table for Task 1 Quantity of Material

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	1508.58		
Within-subject				
Display (D)	1	36.13	0.06	0.8123
D x S	5	576.36		
Session (Se)	2	80.10	1.07	0.3787
Se x S	10	74.75		
Day (Da)	1	506.68	5.10	0.0735
Da x S	5	99.31		
D x Se	2	143.38	1.11	0.3667
D x Se x S	10	129.06		
D x Da	1	25.68	0.12	0.7456
D x Da x S	5	218.31		
Se x Da	2	113.01	3.77	0.0603
Se x Da x S	10	30.00		
D x Se x Da	2	11.01	0.15	0.8589
D x Se x Da x S	10	71.30		
Total	(71)			

TABLE 3

ANOVA Summary Table for Task 1 Number of Errors

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	19.65		
Within-subject				
Display (D)	1	1.13	0.15	0.7143
D x S	5	7.50		
Session (Se)	2	7.10	2.77	0.1105
Se x S	10	2.56		
Day (Da)	1	5.01	1.16	0.2603
Da x S	5	3.14		
D x Se	2	27.55	5.98	0.0196
D x Se x S	10	4.61		
D x Da	1	10.13	1.21	0.3212
D x Da x S	5	8.36		
Se x Da	2	.85	0.09	0.9123
Se x Da x S	10	9.15		
D x Se x Da	2	.29	0.05	0.9506
D x Se x Da x S	10	5.73		
Total	(71)			

TABLE 4

ANOVA Summary Table for Task 1 Error Rates

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	0.005		
Within-subject				
Display (D)	1	0.00009	0.09	0.7710
D x S	5	0.0009		
Session (Se)	2	0.003	1.84	0.2091
Se x S	10	0.0009		
Day (Da)	1	0.0004	0.42	0.5470
Da x S	5	0.0009		
D x Se	2	0.008	6.60	0.0149
D x Se x S	10	0.001		
D x Da	1	0.003	1.86	0.2305
D x Da x S	5	0.002		
Se x Da	2	0.0009	0.04	0.9594
Se x Da x S	10	0.002		
D x Se x Da	2	0.0002	0.14	0.8723
D x Se x Da x S	10	0.002		
Total	(71)			

TABLE 5

Results of Newman-Keuls Analysis for Task 1 Number of Errors

Display	Session	Number of Errors	Newman-Keuls
LC/CRT	1	3.5	A
	2	4.1	A
	3	2.5	A
CRT	1	1.6	A
	2	3.1	A B
	3	4.4	B

Newman-Keuls: Alpha level = 0.05, df = 10, MS = 4.6, n = 12.

Means with the same letter are not significantly different.

TABLE 6

Results of Newman-Keuls Analysis for Task 1 Error Rate

Display	Session	Error Rate	Newman-Keuls
LC/CRT	1	.060	A
	2	.063	A
	3	.034	A
CRT	1	.025	A
	2	.055	A
	3	.071	A

Newman-Keuls: Alpha level = .05, df = 10, MS = .001, n = 12.

Means with the same letter are not significantly different.

<u>Task 2</u> <u>Text search</u>. Tables 7, 8, and 9 summarize the ANOVA results for this task. The main effect of Session is significant for quantity of task material completed (*F* = 75.59, *p* = .0001). Following the initial session, during which subjects became proficient with the search task, the quantity of material processed significantly increased for Sessions 2 and 3 (Table 10).

The Day x Session interaction was also significant for quantity of task material processed (*F* = 16.75, *p* = 0.0006). As the subjects became more familiar with the task in the first session, their work rate increased in Sessions 2 and 3 on Day 1. On Day 2, work output was significantly different among the three Sessions (Table 11). Work output levels significantly increased from Session 1 to Session 2, and then decreased for Session 3 in contrast to Day 1.

The ANOVAs also revealed that error and error rate were affected by the interaction of Day x Session (*F* = 6.53, *p* = 0.0154 and *F* = 3.96, *p* = 0.0541, respectively). Post hoc Newman-Keuls did not reveal any significant differences among the Day x Session interaction means.

<u>Task 3</u>: <u>Text entry</u>. ANOVAs were performed to determine how the performance measures were affected by the independent variables. These are summarized in Tables 12, 13, and 14.

TABLE 7

ANOVA Summary Table for Task 2 Quantity of Material

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	74960.64		
Within-subject				
Display (D)	1	1860.50	0.42	0.5451
D x S	5	4420.96		
Session (Se)	2	119974.88	75.59	0.0001
Se x S	10	1587.26		
Day (Da)	1	5033.39	3.39	0.1251
Da x S	5	1485.78		
D x Se	2	892.54	1.98	0.1881
D x Se x S	10	449.96		
D x Da	1	1476.06	1.67	0.2532
D x Da x S	5	885.72		
Se x Da	2	17544.35	16.75	0.0006
Se x Da x S	10	1047.69		
D x Se x Da	2	499.02	1.12	0.3649
D x Se x Da x S	10	446.73		
Total	(71)			

TABLE 8

ANOVA Summary Table for Task 2 Number of Errors

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	385.714		
Within-subject				
Display (D)	1	21.13	0.46	0.5270
D x S	5	45.76		
Session (Se)	2	15.93	0.43	0.6634
Se x S	10	37.25		
Day (Da)	1	55.13	0.85	0.3979
Da x S	5	64.56		
D x Se	2	1.04	0.03	0.9666
D x Se x S	10	30.53		
D x Da	1	66.13	0.87	0.3936
D x Da x S	5	75.96		
Se x Da	2	457.04	6.53	0.0154
Se x Da x S	10	70.03		
D x Se x Da	2	5.54	0.37	0.7006
D x Se x Da x S	10	15.03		
Total	(71)			

TABLE 9

ANOVA Summary Table for Task 2 Error Rates

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	0.005		
Within-subject				
Display (D)	1	0.00005	0.03	0.8716
D x S	5	0.0016		
Session (Se)	2	0.0032	3.41	0.0744
Se x S	10	0.0009		
Day (Da)	1	0.0021	1.24	0.3155
Da x S	5	0.0016		
D x Se	2	0.0003	0.35	0.7144
D x Se x S	10	0.0009		
D x Da	1	0.0013	0.89	0.3878
D x Da x S	5	0.0014		
Se x Da	2	0.0056	3.96	0.0541
Se x Da x S	10	0.0014		
D x Se x Da	2	0.00001	0.13	0.8783
D x Se x Da x S	10	0.0001		
Total	(71)			

TABLE 10

Results of Newman-Keuls for Task 2 Session Effect

Session	Quantity of Material	Newman-Keuls
1	182	A
2	311	B
3	295	B

Quantity units = Number of specified letters searched

Newman-Keuls: Alpha Level = .05, df = 10, MS = 1587, n = 24.

Means with the same letter are not significantly different.

TABLE 11

Results of Newman-Keuls Analysis for Task 2 Day x Session

Day	Session	Quantity of Material	Newman-Keuls
1	1	147	A
1	2	301	B
1	3	315	B
2	1	216	A
2	2	321	B
2	3	276	C

Quantity units = Number of specified letters searched.

Newman-Keuls: Alpha level = .05, df = 10, MS = 1048, n = 12.

Means with the same letter are not significantly different.

TABLE 12

ANOVA Summary Table for Task 3 Quantity of Material

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	9098.49		
Within-subject				
Display (D)	1	0.013	0.00	0.9948
D x S	5	296.78		
Session (Se)	2	11.29	0.35	0.7140
Se x S	10	32.40		
Day (Da)	1	406.13	1.94	0.2224
Da x S	5	209.29		
D x Se	2	17.60	1.05	0.3846
D x Se x S	10	16.71		
D x Da	1	74.01	7.04	0.0453
D x Da x S	5	10.51		
Se x Da	2	34.63	1.41	0.2886
Se x Da x S	10	24.54		
D x Se x Da	2	84.43	2.11	0.1724
D x Se x Da x S	10	40.08		
Total	(71)			

TABLE 13

ANOVA Summary Table for Task 3 Number of Errors

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	383.05		
Within-subject				
Display (D)	1	21.13	0.29	0.6123
D x S	5	72.43		
Session (Se)	2	24.01	1.55	0.2596
Se x S	10	15.51		
Day (Da)	1	110.01	8.43	0.0336
Da x S	5	13.05		
D x Se	2	7.86	0.48	0.6318
D x Se x S	10	16.38		
D x Da	1	2.35	0.39	0.5606
D x Da x S	5	6.05		
Se x Da	2	5.56	0.53	0.6062
Se x Da x S	10	10.63		
D x Se x Da	2	4.18	0.31	0.7402
D x Se x Da x S	10	13.48		
Total	(71)			

TABLE 14

ANOVA Summary Table for Task 3 Error Rates

Source	*df*	*MS*	*F*	*p*
Between-subjects				
Subjects (S)	5	0.022		
Within-subject				
Display (D)	1	0.0000001	0.00	0.9982
D x S	5	0.012		
Session (Se)	2	0.0042	2.37	0.1433
Se x S	10	0.0017		
Day (Da)	1	0.021	13.85	0.0137
Da x S	5	0.0015		
D x Se	2	0.0011	0.59	0.5722
D x Se x S	10	0.0018		
D x Da	1	0.0016	0.75	0.4260
D x Da x S	5	0.0022		
Se x Da	2	0.0008	0.77	0.4872
Se x Da x S	10	0.001		
D x Se x Da	2	0.0002	0.11	0.8930
D x Se x Da x S	10	0.0016		
Total	(71)			

The main effect of Day was significant for both number of errors and error rate. Day 2 number of errors and error rates were significantly less for this data entry task than were the means for Day 1 (see Table 15). These results indicate, as would be expected, that practice improved performance. Since the quantity of material did not increase from Day 1 to Day 2, the difference in error rate is due to the reduction in the number of errors for a relatively constant quantity of material entered.

An ANOVA also revealed a significant interaction for Display x Day (F = 7.04, p = 0.0453) for quantity of material entered. While there is no overall difference between the two displays in the quantity of material entered, this interaction clearly indicates a similar level of performance on the two days for the LC/CRT, but a significant and larger increase from Day 1 to Day 2 for the CRT (see Table 15). These results suggest that the LC/CRT provides more consistent work output levels.

Subjective Evaluation

The subjective data from questionnaires were treated as an interval scale. The bipolar adjective scales were constructed such that each segment of the scale corresponded to a numerical quantity (-2 to +2). This overlaying of the

TABLE 15

Results of Newman-Keuls Analyses for Task 3 Display x Day Interaction

Display	Day	Quantity Material	Newman-Keuls
CRT	1	86.9	A
LC/CRT	1	88.9	A C
LC/CRT	2	91.7	B C
CRT	2	93.7	B

Newman-Keuls: Alpha level = .05, df = 5, MS = 10.52, n = 18.

Means with the same letter are not significantly different.

bipolar scale on the numerical scale was used as the basis for parametric analysis of these subjective data. Table 16, 17, 18, and 19 summarize these analyses.

The forced choice evaluation portion of the evaluation questionnaire was analyzed using a nonparametric test. The small subject pool (N = 6) and the method of measurement (ranking) dictated that a nonparametric analysis be performed. Table 20 summarizes this analysis.

Bipolar scale results. An ANOVA was performed to assess the effects of display, day, session, scale, and their interactions on the dependent variables. Statistical significance was obtained for the main effects of Session ($F = 6.37$, $p = 0.0165$), and Scale ($F = 2.60$, $p = 0.0013$), and the interaction of Session x Scale ($F = 2.16$, $p = 0.0004$) as shown in Table 16.

These results generally indicate that time on a task significantly affects the subjective scaling for both displays. Subjects rated both displays as being significantly less desirable as the day progressed from Session 1 to Sessions 2 and 3 (Table 17).

The significant main effect of Scale indicated that subjects scored the 20 bipolar scales differently. A post hoc Newman-Keuls analysis revealed Scale 2 (TIRED/RESTED) from the ocular comfort area was the only scale that

TABLE 16

ANOVA Summary Table for Bipolar Scales

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	55.97		
Within-subject				
Display (D)	1	1.00	0.22	0.6622
D x S	5	4.65		
Session (Se)	2	54.82	6.37	0.0165
Se x S	10	8.61		
Day (Da)	1	0.07	0.03	0.8629
Da x S	5	5.26		
Scale (Sc)	19	6.61	2.60	0.0013
Sc x S	95	2.54		
D x Se	2	0.77	0.47	0.6370
D x Se x S	10	1.61		
D x Da	1	0.23	0.10	0.7683
D x Da x S	5	2.33		
D x Sc	19	0.81	1.60	0.0722
D x Sc x S	95	0.51		
Se x Da	2	3.04	1.70	0.2312
Se x Da x S	10	1.79		
Da x Sc	19	0.42	1.09	0.3769
Da x Sc x S	95	0.39		
Se x Sc	38	1.12	2.16	0.0004
Se x Sc x S	190	0.52		

TABLE 16

(continued)

Source	df	MS	F	p
D x Se x Da	2	2.15	1.79	0.2171
D x Se x Da x S	10	1.20		
D x Da x Sc	19	0.72	1.59	0.0758
D x Da x Sc x S	95	0.45		
Da x Se x Sc	38	0.17	0.65	0.9388
Da x Se x Sc x S	190	0.26		
D x Se x Sc	38	0.22	1.01	0.4633
D x Se x Sc x S	190	0.22		
D x Da x Se x Sc	38	0.17	0.65	0.9450
D x Da x Se x Sc x S	190	0.26		
Total	(1439)			

TABLE 17

Results of Newman-Keuls Analysis for Effect of Session on Bipolar Scale Results

Session	Mean Scale Score	Newman-Keuls
1	0.62	A
2	0.13	B
3	-0.03	B

Newman-Keuls: Alpha level = .05, df = 10, MS = 8.6, n = 480. Means with the same letter are not significantly different.

TABLE 18

Results of Newman-Keuls Analysis for Effect of Scale on BipolarScale Results

Scale		Mean Scale Score	Newman-Keuls
20	BRIGHT/DARK	0.72	A
4	PAINLESS/PAINFUL	0.69	A
18	ANCHOR/FLOAT	0.67	A
19	FUSE/FLICKER	0.49	A
17	STABLE/VARIABLE	0.46	A
11	STABLE/VARIABLE	0.40	A B
13	TYPICAL/UNUSUAL	0.40	A B
10	NORMAL/ABNORMAL	0.35	A B
12	COMFORT/UNCOMFORT	0.31	A B
8	CLEAR/BLURRY	0.29	A B
16	INTACT/FRAGMENT	0.22	A B
7	NORMAL/BURNING	0.22	A B
1	GOOD/BAD	0.13	A B
14	EASY TO SEE/ HARD ...	0.07	A B
3	COMFORT/UNCOMFORT	0.06	A B
15	CLEAR/BLURRY	0.01	A B
9	EASY TO FOCUS/HARD ...	0.00	A B
5	NONIRRITATE/IRRITATE	-0.06	A B
6	RELAXED/TENSE	-0.14	A B
2	RESTED/TIRED	-0.49	B

Newman-Keuls: Alpha level = 0.05, df = 95, MS = 2.5, n = 72.

Means with the same letter are not significantly different.

TABLE 19

Results of Newman-Keuls for Scale-by-Session Interaction: Differences in Session by Scale

Scale 1

Session	Mean	
1	0.63	A
2	0.08	B
3	-0.33	C

Scale 2

Session	Mean	
1	0.25	A
2	-0.63	B
3	-1.08	C

Scale 3

Session	Mean	
1	0.58	A
2	-0.13	B
3	-0.29	B

Scale 4

Session	Mean	
1	1.13	A
2	0.63	B
3	0.33	B

Scale 5

Session	Mean	
1	0.63	A
2	-0.21	B
3	-0.58	B

Scale 6

Session	Mean	
1	0.29	A
2	-0.29	B
3	-0.42	B

Scale 7

Session	Mean	
1	0.58	A
2	0.13	B
3	-0.04	B

Scale 8

Session	Mean	
1	0.75	A
2	0.17	B
3	-0.04	B

Scale 9

Session	Mean	
1	0.54	A
2	-0.13	B
3	-0.42	B

Scale 10

Session	Mean	
1	0.83	A
2	0.21	B
3	0.00	B

Scale 11

Session	Mean	
1	1.00	A
2	-0.04	B
3	0.17	B

Scale 12

Session	Mean	
1	0.92	A
2	0.13	B
3	-0.13	B

Scale 13

Session	Mean	
1	0.92	A
2	0.25	B
3	0.04	B

Newman-Keuls: Alpha level = .05, df = 190, MS = .52, n = 24.

Means with the same letter are not significantly different.

TABLE 20

Results of Sign Test for Forced Choice Subjective Data

Display Characteristics	p
Clearness	0.218
Color Quality	0.032
Ease of Viewing	0.218
Character Legibility	0.218
Visual Comfort	0.218
Image Stability	0.218
Contrast	0.032
Overall	0.218

significantly differed from five other scales, four of which were from the image stability portion of the bipolar scales (Scale 17, VARIABLE/ STABLE; Scale 18, ANCHORED/FLOATING; Scale 19, FLICKERING/FUSED; and Scale 20, BRIGHT/DARK). The fifth significantly differing scale (Scale 4, PAINLESS/PAINFUL) occured in the same assessment area as that of Scale 2, ocular comfort (Table 18). These results indicate that subjects perceived the displayed images of the two displays to be stable as indicated by the relatively high rankings that Scales 17, 18, 19, 20 received. The significant difference between Scales 2 and 4, both ocular comfort scales, can be viewed as an indication of the wide variation in adjective sets used to anchor these two scales. It would appear that the adjective set of Scale 2 (TIRED/RESTED) is more indicative of the subjects' perceptions while the adjective set of Scale 4 (PAINLESS/ PAINFUL) was perhaps semantically too harsh, which resulted in higher rankings. The higher ranking indicates that subjects experienced a relatively painless experimental trial. Subjects, therefore, did not experience pain, but they did perceive their eyes as being tired.

Only scales 1 and 2 (ocular comfort) were significantly different for all three sessions (Table 19). Scales 3 - 13 (remaining five scales for ocular comfort, Scales 3 - 7, and all six scales of visual comfort, 8 - 13) showed a significant difference between Session 1 and the other two Sessions. Scales 14 - 20, dealing with image stability, showed nondifferent scale scores for all three Sessions. Subjective scoring for the ocular and visual comfort scales indicates decreasing comfort over time. Specifically, Scales 1 and 2 showed significant differences for all three sessions while Scales 3 - 13 demonstrated a significant difference between Session 1 and Sessions 2, 3. It may be of interest to note that the Session means declined significantly for all 13 of these scales, although the differences between Sessions 2 and 3 were statistically significant for only Scales 1 and 2.

The sensitivity of the adjective word sets anchoring Scales 1 and 2 could explain the differences in significance between the three Sessions as compared to Scales 3 - 13. Subjective scoring on the image stability scales shows that no image stability change was perceived as the sessions progressed.

Forced-choice preference. A nonparametric Sign Test conducted separately on each of the eight categories indicated a preference for the LC/CRT display for the features of COLOR QUALITY and CONTRAST (Table 20). For all other features (CLEARNESS, EASE OF VIEWING, CHARACTER

LEGIBILITY, VISUAL COMFORT, IMAGE STABILITY, OVERALL) five of the six subjects preferred the LC/CRT display; this proportion, however, is not statistically significant (p = 0.218) for the small sample size used in this study. Nonetheless, the consistent five-of-six or six-of-six subject preference for the LC/CRT across all eight categories, when tested with a Randomization test for matched pairs (Siegel, 1956), demonstrates strong statistical significance (p = 0.0078) and indicates an overall preference for the LC/CRT.

Ambient Illumination

Analyses of variance were performed for both ambient and display illumination measures. The ambient illumination results are summarized in Tables 21, and 22.

Subjects chose to view the LC/CRT display with higher ambient illumination (97.9 lux) than the CRT display (71.0 lux). While the ambient lighting selected for the LC/CRT remained consistently greater relative to the CRT ambient lighting preference, the preferred ambient lighting significantly decreased from Day 1 to Day 2 for the CRT (Table 22). The increased ambient illumination selected for the LC/CRT is more compatible with typical office illumination.

TABLE 21

ANOVA Summary Table for Ambient Illumination

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	185.26		
Within-subject				
Display (D)	1	107.60	6.86	0.0472
D x S	5	15.68		
Session (Se)	2	2.10	0.15	0.8660
Se x S	10	14.36		
Day (Da)	1	43.56	4.50	0.0875
Da x S	5	9.68		
D x Se	2	18.02	2.71	0.1147
D x Se x S	10	6.65		
D x Da	1	76.06	7.67	0.0394
D x Da x S	5	9.92		
Se x Da	2	4.27	0.43	0.6615
Se x Da x S	10	9.90		
D x Se x Da	2	4.52	1.75	0.2231
D x Se x Da x S	10	2.58		
Total	(71)			

TABLE 22

Ambient Illumination Means

Display	Lux
LC/CRT	97.9
CRT	71.0

Display	Day	Lux	Newman-Keuls
LC/CRT	2	100.1	A
	1	94.7	A
CRT	1	90.4	A
	2	51.6	B

Newman-Keuls: Alpha level = 0.05, df = 5, MS = 9.9, n = 18.

Means with the same letter are not significantly different.

The analysis of variance performed for display illumination did not reveal any significant differences (Table 23).

TABLE 23

ANOVA Summary Table for Display Illumination

Source	df	MS	F	p
Between-subjects				
Subjects (S)	5	1611.75		
Within-subject				
Display (D)	1	833.68	2.86	0.1515
D x S	5	291.41		
Session (Se)	2	33.85	1.84	0.2080
Se x S	10	18.35		
Day (Da)	1	62.35	2.25	0.1937
Da x S	5	27.68		
D x Se	2	142.35	2.00	0.1860
D x Se x S	10	78.18		
D x Da	1	360.01	1.54	0.2698
D x Da x S	5	233.88		
Se x Da	2	64.60	1.23	0.3331
Se x Da x S	10	52.53		
D x Se x Da	2	102.26	2.54	0.1286
D x Se x Da x S	10	40.33		
Total	(71)			

DISCUSSION

Task Variation

In general, the results do not offer a clear indication that either display produced better task performance. The mixed statistically significant results for each task type suggest that the varying features and visual emphasis found in each task type interact differently with each display over time. Further, these differences do not offer any clear pattern of how performance was affected by the varying display characteristics inherent in the two display technologies.

The three task types selected provide a representative sampling of actual VDU applications. Each of the three task types influenced the experimental variables in a somewhat different manner, as evidenced by the variability of significant effects and interactions. Thus, the three task types allowed for a broad range of possible human visual system and display characteristic incompatibilities to be expressed.

For Task 1, text editing, the significant Display x Session interaction does not consistently show improved performance for either display. An interpretation of these results indicates a non-different level of performance between the two days for both display technologies; however, the performance trends for the two displays appear to be diverging. It is not clear that either display's performance measures have reached an asymptote. Therefore, it would be of interest for future research to extend the duration of the experiment to observe if asymptotic performance measures would produce significant differences between the two displays.

The same interpretation can be made for Task 3, text entry, where the significant Display x Day interaction does not clearly demonstrate improved performance for either display. Once again, the performance measures appear to diverge as indicated by not reaching asymptotic performance. Extending the duration of the experiment may reveal significant differences between the two displays.

For Task 2, text search, the significant Day x Session interactions do not afford any consistent indication of how the task affected the performance measures. The most general interpretation of these results would be that time on task was a significant factor in performance. A possible explanation for the decline in quantity of material accomplished for Session 3 on Day 2 is fatigue or waning subject motivation. These results would indicate that a maximum was obtained for quantity of material completed and the 2-day length of the experimental sessions allowed the subjects sufficient time to maximize their work output in terms of quantity of material processed.

Time on task also provides a possible interpretation for number of errors incurred for this task type. Subjects exercised greater care and consequently processed material at a slower rate and with fewer errors initially. On Day 2, Session 1, after an approximately 16-hr overnight interruption, subjects were less proficient than during Session 3 of Day 1. By Session 2 of Day 2 subjects had regained their performance proficiency. The increase in errors and error rate in Session 3 could again be indicative of subject fatigue. The expectation that the differing display characteristics found in the two display technologies would significantly affect performance of a visually intensive task was not demonstrated by these results.

Subjective Variation

The subjective data appears to corroborate the possible effects of fatigue as represented by the increase in error rate for Task 1 and the decline in quantity of material for Task 2. The bipolar adjective scales assessed physiological and perceptual sensations, ocular comfort and visual comfort scales. In both ocular comfort and visual comfort a significant decline over Sessions in ranking was demonstrated. This decline can be interpreted as physiological and perceptual fatigue which would agree with the significant performance decrements exhibited as the sessions progressed.

Subjectively, a significant preference for the liquid crystal/cathode-ray tube display was obtained with the forced choice preference evaluation method for the categories of COLOR QUALITY and CONTRAST; in addition, an overall preference for the LC/CRT was demonsrated. The bipolar adjective scale evaluations, however, did not indicate a significant preference difference. Due to the lack of consensus between the two subjective evaluation methodologies, no clear statement of subjective preference can be made, although the general trend suggests preference for the LC/CRT.

Ambient Illumination

The interaction between inherent display characteristics and the human visual system was clearly demonstrated by the preference for greater ambient illumination for the liquid crystal/cathode-ray tube display. The ability of the LC/CRT to reduce the interference of ambient light with the displayed image via the filters employed in color generation allows increased ambient light environments to be used. By allowing greater ambient illumination, both VDU and non-VDU tasks can be performed within the same lighting environment.

These results must be viewed within the limitations that a sample size of six subjects provides. The length of the experiment makes this study the most "long-term" observation of the VDU operator's task performance yet conducted. The constraints imposed by the large time requirements for each subject necessitated that a relatively small sample be used. Of course, a small sample size reduces the power of the statistical test conducted, as was evident in the nonparametric analyses, which required a unanimous decision before statistical significance could be obtained.

The concerns generated at the initiation of the study, image break-up and chromosteropsis, were addressed with each subject. The bipolar adjective scales, 14 - 20, evaluated perception of the displayed image. There were no significant differences demonstrated between the two displays for any of the image stability scales. A final verbal interview also addressed these issues and corroborated the results of the image stability portion of the subjective evaluations. Even with the two displays presented simultaneously and visual attention directed at the issues of image break-up and chromosteropsis, subjects were unable to perceive these effects. These concerns, clearly, did not affect the subjects' perception of the LC/CRT display.

In order to address the limitations brought forth in the discussion, suggestions are offered. First, as previously indicated, increasing the duration of the experiment could clarify trends indicated in the results, but which have not reached an asymptote during the four days of the present experiment. A second concern of small sample size would be ameliorated by increasing the number of subjects. Further, the type of subjects should be expanded to include at least older individuals with corrected vision. Such individuals are clearly present in the VDU operator population, but due to experimental constraints were not represented in this research effort. Finally, the bipolar adjective scales offered were extensive. A more parsimonious selection of scales, particularly ones with higher correlations to VDU operator sensations and complaints, might reveal differences not demonstrated in this study.

Another issue of concern is the small F ratios (less than one) which exist throughout the statistical results. One underlying assumption for parametric analyses of data is normalcy. The data were evaluated using Kolomogorov D and Student's t statistics, and skewness and kurtosis tests for normalcy (SAS Institute, 1982), with the resulting conclusion that the data are not normally distributed. The small F ratios are the result of nonadherence to the assumption of a normal distribution. However, the robustness of the analysis of variance procedure results in a conservative alpha level. The potential for a Type II error is greatly increased, giving rise to the possibility that significance may exist, but remains undetected. With these statistical considerations in mind, the analyses of variance were accepted for the data.

CONCLUSIONS

Three task types, each with its corresponding visual emphasis were selected for this study. This was done to provide a broad range of visual intensities to interact with the two displays. The variable visual involvement with the two displays did not reveal any clearcut performance differences.

The subjective data revealed mixed results. The bipolar adjective scales indicated no differences between the two display technologies. A subjective preference for the LC/CRT display was demonstrated on some display parameters with the forced choice rating scales. However, the specificity of the subjective preference areas and the forced choice evaluation methodology as well as the insignificant results obtained with the bipolar adjective scale evaluation technique does not indicate a consistent subjective preference for either display technology.

A significant difference between the two displays was demonstrated with ambient illumination level preferences. These results indicate that greater ambient illumination was preferred for the LC/CRT display. The display characteristics of the LC/CRT allowed for its use in ambient illumination environments which more closely approaches the typical office lighting situation. The need for costly devices and adaptations to enhance the displayed image in the high ambient environment could, therefore, be avoided. The preference for the LC/CRT display in greater ambient illumination provides increased versatility in office environments.

The potential concerns of image break-up and chromosteropsis due to the inherent technology used to generate the displayed image of the LC/CRT did not result in performance or subjective preference differences. Subjects were unable to perceive these phenomena even after closer inspection directed by the experimenter.

On balance, results which were statistically significant favor the LC/CRT. No statistically significant results favor the CRT over the LC/CRT. The existence of a number of nonsignificant differences suggests that any preference for the LC/CRT be viewed cautiously.

REFERENCES

Beal, D. (1984, September). Martinique. Gourmet, 48 - 52.

Bouma, H. (1980). Visual processes and the quality of text displays. In E. Grandjean and E. Vigliani (Eds), Ergonomic aspects of visual display terminals. London: Taylor and Francis.

Crane, P. M. (undated). Effects of work at video display computer terminals on vision, mood, and fatigue symptoms. Unpublished report for OCLC, Columbus, Ohio.

Dainoff, M. J. Occupational stress factors in secretarial/clerical workers: Annotated research bibliography and analytic review. (1979). U.S. Department of Health, Education, and Welfare.

Demilia, L. A. (1968). Visual fatigue and reading. Journal of Education, 151, (2), 4-34.

Dunn-Rankin, P. (1983). Scaling methods. Hillsdale, New Jersey: Lawrence Erlbaum.

Epps, B. (1984). The influence of insertion-instruction strategies on the noise attenuation capabilities of industrial earplugs. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Gould, J. D. and Grischkowsky, N. (1984). Doing the same work with hard copy and cathode ray tube (CRT) computer terminals. Human Factors, 26(3), 323-337.

Haider, M., Kundi, M., and Weissenbock, R. (1980). Worker strain related to VDUs with differently coloured characters. In E. Grandjean (Ed.) Ergonomic aspects of visual display terminals. London: Taylor and Francis, 53-64.

Hart, D. J. The human aspects of working with visual display terminal. (1976, February). INCA-FIEJ Research Assn. (IFRA),report 7602-30.

Heise, D. R. (1970). The semantic differential and attitude research. In G.P. Summers (Ed.), Attitude measurement. Chicago: Rand McNally.

IBM. Human factors of workstations with display terminals. (1979). G320-6102-1 (2nd ed.). San Jose: Human Factors Center, IBM.

Laubli, T., Hunting, W., and Grandjean, E. (1980). Visual impairments in VDU operators related to environmental conditions. In E. Grandjean and E. Vigliani (Ed.), Ergonomic aspects of visual display terminals. London: Taylor and Francis.

Mourant, R.R., Lakshmanan, R., and Chantadisai, R. (1981). Visual fatique and cathode ray tube display terminals. Human Factors, 23(4), 529-540.

Murch, G. M. (1984, June). Human factors of displays. Paper presented at Annual SID Symposium, San Francisco.

National Research Council, National Academy of Science. (1983). Video displays, work and vision. Washington D.C.: National Academy Press.

National Institute for Occupational Safety and Health. (1981). Potential health hazards of video display terminals: Oakland Tribune, Blue Shield of California, San Francisco Newspaper Agency, Chronicle, and Examiner. U.S. Department of Health and Human Services.

Osgood, C. E., Tannenbaum, P. H., and Suci, G. J. (1957). The measurement of meaning. Urbana: University of Illinois Press.

Ostberg, O. (1974). Fatigue in clerical work with crt display terminals, Goteborg Psychological Reports, 19, (4).

Ostberg, O. (1975). CRTs pose health problems for operators. International Journal of Occupational Health and Safety, November-December, 44 (6).

SAS Institute. (1982). SAS user's guide. Cary, NC: Author.

Schurick, J. M. (1984). Answers to preliminary examination questions leading to PhD. Unpublished examination, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Siegel, S. (1956). Nonparametric Statistics: For the Behavioral Sciences. New York: McGraw-Hill.

Smith, A., Tanaka, S., and Halperin, W. (1984). Correlates of ocular and somatic symptoms among video display terminal users. Human Factors, 26(2), 143-156.

Snyder, H. L. (1980, July). Human visual performance and flat panel display image quality. Virginia Polytechnic Institute and State University, Technical Report HFL-80-1.

Starr, S. J. (1984). Effects of video display terminals in business office. Human Factors, 26(3), 347-356.

Stewart, T. (1980). Problems caused by continous use of visual display units. Lighting Research and Technology, 12, (1), 26-36.

Timmers, H., van Nes, F. L., and Blommaert, F. J. J. (1980). Visual word recognition as a function of contrast. In E. Grandjean and E. Vigliani (Ed.), Ergonomic aspects of visual display terminals. London: Taylor and Francis.

Vatne, R., Johnson, P. A., and Bos, P. J. (1983). A lc/crt field- sequential color display. _1983 SID Digest of Technical Papers_, 28-29.

APPENDICES

APPENDIX A: Informed Consent

The purpose of this research is to evaluate two color displays. The comparison will involve the performance of representative word processing tasks (e.g. proofreading text) on a CRT. The tasks could possibly become monotonous. If at anytime during the experiment you wish to discontinue your participation, you may. Your earnings will be based on the portion of time you have participated. At any time during the research sessions you may ask questions, as you have the right to be fully informed about the purpose and results of this study. The answer to your questions may be delayed until the end of your participation in order not to influence the experimental results. A copy of the results, at your request, will be sent at a later date. You may withdraw your data from the experiment within 48 hours of completing the study. After that time individual data is no longer identifiable in order to preserve participant anonymity.

The experimental process will involve approximately 35 hours over a four day period. You will be paid 6.00 dollars per hour for your participation. Discussion of this experiment with others who could be potential subjects should be avoided until April 15, 1985, as any prior knowledge of the experiment influence the outcome of the experiment.

This research has been reviewed and approved by the Institutional Review Board as acceptable for human participation. Any questions or concerns regarding your participation in this research should be directed to the experimenters, Dr. Harry L. Snyder (961-5358) and Cristina Christensen (961-5495) or Charles D. Waring, Chairman of the Institutional Review Board, (961-5284).

As a participant in this experiment, you have certain rights. The purpose of this sheet is to describe these rights to you and to obtain your written consent to participate.

1) You have the right to discontinue participating in the study at any time for any reason. If you decide to terminate the experiment, inform a member of the research team and he will pay you for the portion of time you have spent.

2) You have the right to inspect your data and to withdraw it from the experiment, if you feel that you should. In general , data are procesed and analyzed after all subjects have completed the experiment. In this experiment, the investigators can supply you with some qualitative information immediately following the experiment. Subsequently, all the data are treated with anonymity. Therefore, if you wish to withdraw your data, you must do so within 48 hours after your participation is completed.

3) You have the right to be informed as to the overall results of the experiment. If you wish to receive a synopsis of the results, include your address (three months hence) with your signature below. If after receiving the synopsis, you would then like further information, please contact the Human Factors Laboratory and a full report will be made available to you.

The faculty and graduate student members of the research team sincerely apreciate your participation. They hope that you will find the experiment a pleasant and interesting experience. If you have any questions about the experiment or your rights as a participant, please do not hesitate to ask. The investigators will try to answer them, subject only to the constraint that the results will not be pre-biased by a detailed answer.

Your signature indicates that you have read and understood your above stated rights as a participant, and that you consent to participate.

____

Signature

____

____

____

Print Name and Address for a Synopsis of the Results

Human Factors Laboratory
Department of IEOR
Room 140 Whittemore Hall
Virginia Tech
Blacksburg, Va 24061
(703) 961-5495

APPENDIX B: Task Examples

Task 1: Text Editing

One of the goals in desgining human-computer dailogues is to minmize the menal load no the user. This maybe particularly critic al ni dialouges incorpaorting speech be cause infornation presen tation is transeint. Discuss varouis approaches to reducig mental workloa ni human-computer dailogues. Cite appropriate existig research as well as araes for futur rese arch. INTRODUCTION The design of "user friendly" huan-computer dialogues hsa recently become a mahor research issue. Onr of the main gaols of this research si to reduce th cognivite of men tal workload associated with intercaitgn with a computer, esp ecially for novices and casual users. Desig i ng user freindly dia logues is nit an easy task however. It requires the desgiren to consider limit- ations of the user's langage abilitie problem solving memory, and attention, nost of whch are not fully unerdstood. The design off inter active dailogues should ideally be ba e d on thoeries of men tal processing. The re search to supp ort such design is presently evlobving slowly, and it si likely that res earch on p timal systems will cont inue for a long time yet to come. This paper is concrened with the apppoaches . . . (Shurick, 1984).

Task 2: Text Search

LETTER TO LOCATE C

1 YHAYKM	21 PGKHEK	41 FKYKSC	61 WBQDRD
2 QDKWUQ	22 ATRDFX	42 VNTHGV	62 DIFKQM
3 WHUKGB	23 IFPUDP	43 NKTMPD	63 DSQNHB
4 NJMJRB	24 VEVBFN	44 KQLFQS	64 DRPYVK
5 SWJYKB	25 AQMLAP	45 GXPAGO	65 EMYLED
6 PJGDBW	26 AUXTUF	46 MHUURA	66 BHMWUG
7 KSWPFP	27 HEFTQL	47 VUBRNU	67 FLPEGU
8 YTWDGY	28 COZPSO	48 WKSEIL	68 BLPKHD
9 RLEEMR	29 JSAAAP	49 NVRBLA	69 QTNYVW
10 YPKKPJ	30 AIMQYV	50 TFSQZF	70 QWNKHK
11 ZPVDFH	31 ZTNYNJ	51 GJVCNP	71 ETLEGC
12 JDAFRV	32 OZHEYK	52 LEELTT	72 SZETXN
13 YOUEPF	33 NGJKCD	53 HTTCWI	73 SYJOXT
14 PKXENB	34 SBXNMU	54 NVEPJR	74 ASFIKH
15 GUNLMG	35 ENPPYI	55 HOXMQX	75 IXIOMV
16 GJBOYC	36 GVGYQH	56 ONTOPM	76 GCJKQV
17 AEZDOK	37 VBZJZG	57 FKVBXH	77 XUZYBL
18 IMFFRT	38 EYXPUK	58 QFUWBM	78 VAZMNA
19 NSHIYQ	39 TRJSZY	59 ZPXMAK	79 HYTUQV
20 IGIFPG	40 GSOTHR	60 ALBAJY	80 DEZHQR

(Mourant et. al.,1981)

Task 3: Text Entry

COFFEE (Coffea arabica)

Water is the number one beverage in the world today. In second place is coffee, the background of which is as intriguing as that if the far- flung spices and herbs. According to legend, the flavorful wonders of coffee were discovered in the third century A.D. Monks fleeing from persecution found refuge in the highlands of Abyssinia across the Red Sea. One night a father tending the flock ran to the monastery and cried, "The animals are bewitched. They gambol and play as if it were spring morn." The prior tried to reassure him. He would go to the fole. He found that the monk was not seeing things. Night after night the animals continued to frisk about instead of sleeping as they should. The prior, after much study of the plants the animals browsed upon, was convinced that the sleeplessness was caused by the leaves and fruit of an unfamiliar shrub that grew in profusion there. He picked some of the ripe cherrylike fruit, chewed the seeds, felt exhilarated, and was very wakeful that night.

Whether or not legend is correct, Abyssinia and Arabia were the original homes of the coffee shrub. The Turkish . . . (Beal, 1984).

APPENDIX C: Questionnaire and Evaluations

Pre-experimental Questionnaire

1) How long have you directly worked with a computer word/text processing system?

___________Months

2) How many hours per day are you usually engaged in working with a word/text processor?

___________Hours Per Day

3) Approximately what percentage of the time you are working with a word/text processor is spent looking at the display screen (as opposed to looking at paper documents.)?

4) Have you ever worked with a color display?
If so, for how many hours per day?

___________Hours Per Day

Evaluation Instructions

After each two hour session you will be asked to judge the display on several descriptive scales. Your judgment should be based on how the display affected you. Instructions are provided for use of the scales.

If you feel that the display's effect on you is very closely related to one end of the scale, then you should mark the scale as follows:

	X					
Good	2	1	0	-1	-2	Bad

OR

					X	
Good	2	1	0	-1	-2	Bad

If you feel that the display's effect on you is moderately closely related to one end of the scale, then you should mark the scale as follows:

		X				
Good	2	1	0	-1	-2	Bad

OR

				X		
Good	2	1	0	-1	-2	Bad

Which end of the descriptive scale you choose will, of course, depend on which word best reflects your feelings about the display.

If you feel neutral about how the display affected you, then you should mark the scale as follows:

```
                       X
Good _____:_____:_____:_____:_____ Bad
       2     1     0    -1    -2
```

Please place your marks in the space provided, not on the boundaries:

```
                 this         not this

                  X              X
Good _____:_____:_____:_____:_____ Bad
       2     1     0    -1    -2
```

Please complete all of the items; do not omit any of the items. Please make only one mark for each item.

Immediately following your performance of tasks on the display, please rate it on the following descriptive scales.

Make your judgments based on how your <u>eyes</u> feel.

Good	_____ 2	_____ 1	_____ 0	_____ -1	_____ -2	Bad
Tired	_____ -2	_____ -1	_____ 0	_____ 1	_____ 2	Rested
Uncomfortable	_____ -2	_____ -1	_____ 0	_____ 1	_____ 2	Comfortable
Painless	_____ 2	_____ 1	_____ 0	_____ -1	_____ -2	Painful
Irritated	_____ -2	_____ -1	_____ 0	_____ 1	_____ 2	Non-irritated
Relaxed	_____ 2	_____ 1	_____ 0	_____ -1	_____ -2	Tense
Burning	_____ -2	_____ -1	_____ 0	_____ 1	_____ 2	Normal

Immediately following your performance of tasks on the display, please rate it on the following descriptive scales.

Make your judgments based on how your <u>vision</u> feels.

Clear	2	1	0	-1	-2	Blurry
Easy to focus	2	1	0	-1	-2	Hard to focus
Abnormal	-2	-1	0	1	2	Normal
Stable	2	1	0	-1	-2	Variable
Uncomfortable	-2	-1	0	1	2	Comfortable
Typical	2	1	0	-1	-2	Unusual

Immediately following your performance of tasks on the display, please rate it on the following descriptive scales.

Make your judgments based on your evaluation of <u>the images on the display</u>.

Easy to see	2	1	0	-1	-2	Hard to see
Blurry	-2	-1	0	1	2	Clear
Intact	2	1	0	-1	-2	Fragmented
Variable	-2	-1	0	1	2	Stable
Anchored	2	1	0	-1	-2	Floating
Flickering	-2	-1	0	1	2	Fused
Bright	2	1	0	-1	-2	Dark

Final Evaluation Instructions

After performing tasks on both displays, you are asked to judge the two displays on descriptive scales and to indicate a preference for one of the displays over the other.

Your instructions for the descriptive scales remain the same as the previous evaluations you have completed.

In the second part of the final evaluation you will be given several display features and be asked to choose which display you prefer for each of the features listed. Place an "X" to the right of the display you prefer for each display feature listed.

Please do not omit any of the items in the evaluation. Please make only one mark for each item.

Final Evaluation

Please place your mark based on your overall reactions while performing on the two displays.

The difference between the two displays is:

Very small _____:_____:_____:_____:_____ Very large
-2 -1 0 1 2

In comparing the two displays I have a

Strong preference _____:_____:_____:_____:_____ No preference
2 1 0 -1 -2

Please decide which of the two displays you would prefer, based on the features listed below:

Clearness	Display B ___	Display A ___
Color Quality	Display A ___	Display B ___
Ease of Viewing	Display A ___	Display B ___
Character Legibility	Display B ___	Display A ___
Visual Comfort	Display A ___	Display B ___
Image Stability	Display B ___	Display A ___
Contrast (Brightness)	Display A ___	Display B ___
Overall	Display B ___	Display A ___

<u>Final Evaluation Interview</u>

(Two Displays Visible)

1. In using Display ___ (to indicate LC/CRT display) did you notice:

a. a band splitting the display screen horizontally?

b. the image on the screen "breaking up" while working in the display? (After the subject answers the question, ask the subject to rapidly scan the screen diagonally; move their hand rapidly in front of the screen; and quickly bring their teeth together to aid in the possible perception of "image break-up".)

c. the colors in the display changing as you changed your viewing angle? (After subject answers the question, demonstrate changes in chromaticity due to viewing angle.)

d. the letter characters "swimming" on the display, in particular with the green text and the red highlighting?

2. Which color combination (text and highlighting colors) did you like the best? On which display?

3. Which of the two displays would you prefer to keep on using?

a. Do you think this display should cost more than the other display?

b. If yes, how much more should the display cost, 10%, 20%, 30% ...?

The vita has been removed from the scanned document