Subjective Image Quality of CRT Displays under Ambient Glare: Assessing the ISO 9241-7 Ergonomic Technical Standard

Gary S. Olacsi

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science

in

Industrial and Systems Engineering

Dr. Robert J. Beaton, Chair

Dr. A.M. Prestrude

John Deighan

April 10, 1998

Blacksburg, Virginia

Keywords: Glare, CRT, Glare-Filter, ISO 9241-7 Ergonomic Standard

Subjective Image Quality of CRT Displays under Ambient Glare: Assessing the ISO 9241-7 Ergonomic Technical Standard

by
Gary S. Olacsi
Dr. Robert J. Beaton, Chairman
Industrial and Systems Engineering Department

(ABSTRACT)

This thesis evaluates the ISO 9241 Part 7, *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 7: Requirements for display with reflections*. The thesis involved two phases of effort that evaluated the photometric measurements required in the ISO standard in terms of subjective image quality judgments. In phase one, seven monitors were evaluated photometrically according to the ISO 9241-7 standard to determine whether they were Class I, II, or III. Additionally, glare filters were attached to monitors to see if they change the ISO classification of the monitor.

The results of phase one indicated that positive polarity always yielded either Class I or Class II because the Large Source, *Screen Image Luminance Ratio* ($LR_{BDS/FDS}$) and the *Specular Reflection Luminance Ratio* ($LR_{BDS/BD}$) always passed. Conversely, negative polarity always produced Class II or Class III because the Small Source, *Specular Reflection Luminance Ratio* ($LR_{BDS/BD}$) failed every time. Also, the AF 150 and HF 300 were the best filters for reducing glare on monitors. The BF 10 and AF 100 or AF 200 ,on the other hand, were the worst because they intensified screen reflections.

In phase two, human image quality judgments were collected to determine if people rated Class I, II, or III monitor-filter combinations differently under different lighting and different screen polarity conditions. Specifically, phase two assessed the effects of seven monitor/add-on glare filter combinations, five ambient lighting conditions, and two screen polarities on subjective image quality ratings. Each participant provided subjective image quality judgments by viewing alphanumeric text on the CRT screens. Subjective scale values also were correlated with ISO classifications and two ISO metrics*: screen image luminance ratio* (Diffuse, 200 lux) and *specular reflection image luminance ratio*.

The ANOVA findings indicated that specular glare significantly degrades image quality ratings more than diffuse glare. The author contends that this finding is the result of an experimental context effect. In other words, the specular glare was so influential on subjective ratings of image quality that subjects paid little or no attention to reductions in contrast from the diffuse lighting conditions. The correlation analysis showed that the *specular reflection luminance ratio* and the negative polarity classifications did index subjective quality ratings.

Finally, this thesis establishes a human factors basis to justify the measurement requirements in the ISO 9241-7 standard. Specifically, the findings show that it underemphasizes the contribution of the *specular reflection luminance ratio’* and overemphasizes the contribution of the *screen image luminance ratio* to compliance classification calculations, because the procedure gives equal weight to both ratios.

ACKNOWLEDGMENTS

The success of this research endeavor was the result of cooperative efforts of many individuals. First, I would like to thank my committee chair, Dr. Robert Beaton, for providing the much needed guidance and technical expertise to complete this research. His comments, suggestions, and corrections over the last nine months were always helpful. John Deighan brought me up to speed on the photometric system and provided valuable insight on taking ISO measurements.

My family has been very supportive throughout the research process. My loving thanks goes to my mother, father, sister Lori, sister Lisa, and cousin Ron for being there to support me. They helped me to realize the importance of setting and obtaining goals in life.

PREFACE

This research was undertaken as a team project. The author conducted his research in conjunction with Joy Kempic. Both researchers evaluated the ISO 9241-7 standard using human performance data. The difference was that the author evaluted the ISO standard using subjective measures of human performance, while Joy Kempic evaluted the ISO standard using objective measures. Therefore, much of the background and literature review sections are similar in the two theses since they deal with similar topic areas. Both researchers contributed to writing these sections and have a strong appreciation for the research process. Additionally, the researchers worked together in the Displays and Controls Laboratory at Virginia Tech to take the 500 measurements on the monitors and filters according to the ISO 9241-7 standard. Therefore, the Phase One Methods and Results sections are also similar in the two theses. Note: Phase Two of these theses greatly differ since Joy Kempic performed an objective human performance evaluation, while the author performed a subjective human performance evaluation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

In today's society, computer workstations are common tools in homes and offices. Since the beginning of electro-optical visual display terminal (VDT) use in the late 1970's, people have complained of several problems associated with the use of these devices. The most prevalent class of complaints is visual performance-related (Rancourt and Grenawalt, 1986). In particular, VDT operators complain of blurred vision, redness, burning, dry eyes, and visual fatigue. Investigators have determined that glare is a contributing factor to visual discomfort and eye-strain (Cakir, Hart, Stewart, 1980). One cause of glare in the workplace is that VDTs are prone to reflections of ambient illuminance. In fact, studies by Laubi, Hunting, and Grandjean (1982) and Stammerjohn, Smith, and Cohen (1981) reported that 45 and 85% of VDT operators complain of bothersome screen reflections. Therefore, screen reflections are a serious problem for VDTs users and efforts should be made to reduce them.

U.S. National and international standards have been developed to address reflected glare problems in cathode ray tube (CRT) workstations. Essentially, they measure the ability of a display to sustain satisfactory image quality in workplace luminous environments that produce screen reflections. The *American National Standard for Human Factors Engineering of Visual Display Terminal Workstations* (ANSI/HFES 100-1988) approaches the problem of glare indirectly by providing display hardware requirements for luminance, contrast, and resolution. Unfortunately, this standard does not directly measure the design attributes of the display (and/or glare filter) surfaces that are responsible for disruptive screen reflections. In other words, the tests prescribed in the standard produce results that indicate only the effects that reflected glare has on display luminance, contrast, and resolution output for a specific workplace lighting environment. On the other hand, International Standard Organization 9241 Part 7 of the Ergonomic Requirements for Office Work with Visual Display Terminals (ISO 9241-7) quantitatively characterizes the reflectances of display device surfaces so that their effects on human performance can be determined.

ISO 9241-7 directly measures the display hardware design attributes that produce screen reflections (i.e., the diffuse and specular reflectances of the display device surfaces). Thus, this standard has the potential to provide display and glare filter manufacturers the necessary information to assess product performance in work environments susceptible to reflected glare. However, there are two main problems with this standard that need to be investigated. First, the ISO 9241-7 does not address an important component of specular reflections: the reflected image structure or sharpness. This is unfortunate because researchers have noted that a more sharply defined or structured image is perceived as more annoying and noticeable than a less structured image of the same intensity (Rancourt and Grenewalt, 1986). This omission may be due to the inability of investigators in the field to agree upon a single quantitative measurement method for characterizing display specular reflectance.

A second problem with the ISO standard is that it is loosely based human performance data. It is entirely unclear how the standard developers arrived at the specific luminance ratio requirements. Therefore, the question of whether the three ISO classifications truly represent perceivable display performance differences remains to be addressed.

With the two problems above in mind, the present thesis was designed to assess the validity of the ISO standard. Particularly, whether the ISO standard can effectively model CRT viewability under various glare conditions and whether it can also assess potential improvements of glare filter use.

BACKGROUND

In general, glare is defined as a “brightness within the field of vision that is sufficiently greater than the luminance to which the eyes are adapted so as to cause annoyance, discomfort, or loss in visual performance and visibility” (Sanders and McCormick, 1993). Glare can be categorized by its effects on people into either discomfort or disability. Discomfort glare is caused by intense bright light or non-uniform distributions of light that are in the visual field. A person perceives the light as annoying or disturbing, but it does not interfere with a person’s ability to perform work. On the other hand, disability glare does interfere with a person’s ability to perform work by superimposing stray light caused by the glare onto the retinal image. Stray light reduces the contrast of an item being viewed to the point where it may be impossible to see it. A good example would be seeing the outline of a person standing in front of a window on a sunny day. People often experience discomfort glare at a visual display terminal and not disability glare (Cakir et al., 1980).

Glare is further defined by type: direct or reflected. Direct glare occurs when glare sources are in the direct field of view of a person. Reflected glare occurs when the glare sources are reflected from a surface and into the visual field. This thesis focuses on reflected glare since this is what occurs on computer screens and can be reduced by the use of glare filters.

Reflected glare on a display is caused by two types of screen reflections: diffuse and specular. Diffuse reflections create a light or bright haze over some or all of the display and are not well defined images. In contrast, specular reflections produce sharp mirror like images on the first surface of the display. Visual display terminals can have both diffuse and specular reflections at the same time.

Effects of Screen Reflection Types

Specular and diffuse reflections create several problems for the operators of visual display terminals. To understand these problems, it is necessary to explain how and where the reflections are formed on a CRT. An untreated CRT monitor usually consists of a curved glass placed over the phosphor layer of the screen. The glass surface reflects about 4% of the light that falls onto it (Grandjean, 1987; Cakir et al., 1980). The light that is reflected from the curved glass creates the specular (sharp mirror-like) images. These images appear to the observer to be behind the visual display screen. For some displays, they appear to be as much as 32 cm behind the screen (Sanders and

McCormick, 1993). Not all of the light is reflected from the glass surface. Some light is transmitted through to the phosphor layer of the screen. The phosphor layer of the screen is responsible for producing the display images that an operator must detect. The light that is reflected from the phosphor layer creates most of the diffuse (light haze) images. However, some of the light that is reflected from the first surface scatters and creates diffuse images as well.

Specular reflections are annoying and distracting to the operator for two reasons. First, people tend to shift their eyes toward bright spots in their visual field. This effect is known as *phototropism* (Sanders and McCormick, 1994). The second problem is that the eye has difficulty focusing on the display images, since the characters on the screen and the reflections are in different focal lengths. When the eye is focused on the characters, the reflected image is blurred and out of focus. Because the eye naturally wants to bring the blurred image into focus, it will shift its attention to the specular reflection: thereby blurring the character images. Therefore, the eye's focus is constantly oscillating between screen characters and the reflected image (Cakir et al., 1980).

The degree to which the operator feels annoyed or distracted by specular reflections is dependent upon two characteristics of the specular image. One characteristic is the intensity or brightness of the image (Cakir et al., 1980; Rancourt and Grenawalt, 1986). The brighter the specular image is, the more distracting it becomes. Another characteristic is the sharpness of the image. A more sharply defined image is perceived as more annoying and noticeable than a less sharply defined image of the same intensity (Cakir et al., 1980; Rancourt and Grenawalt., 1986).

One reason that diffuse reflections are annoying to the operator is that they reduce the contrast of the characters on the screen. In order to read text on a display, there must be a certain amount of luminance difference between the characters and the background. Diffuse reflections reduce contrast by illuminating all areas of the screen image, which decreases the luminance difference. The screen is, therefore, more difficult to read in the presence of strong diffuse reflections. This phenomenon can be mathematically represented by the modulation equation:

$$M = \frac{\left(L_{\max} + L_{amb}\right) - \left(L_{\min} + L_{amb}\right)}{\left(L_{\max} + L_{amb}\right) + \left(L_{\min} + L_{amb}\right)} = \frac{\left(L_{\max} - L_{\min}\right)}{\left(L_{\max} + L_{\min} + 2L_{amb}\right)} \qquad \text{(Eq. 1)}$$

where L_{max} represents the maximum luminance of the display

L_{min} represents the lowest luminance of the display

L_{amb} represents the ambient luminance reflected from the display

If the minimum luminance (L_{min}) equals the maximum luminance (L_{max}), then the modulation value is zero and the display is unreadable. Therefore, higher modulation values are better. A display with no ambient light (in a dark room) will yield a higher modulation value then the same display with ambient light. Thus, the display will be harder to read with ambient light. For example, if the display is viewed in a completely dark room ($L_{amb}=0$) and $L_{max}=100$ cd/m^2 and $L_{min}=10$ cd/m^2, then the modulation value is 0.82. Suppose now there is ambient light present in the room and $L_{amb}=10$ cd/m^2, then the modulation value is 0.69; a rather substantial decrease in modulation.

There are U. S. National technical standards that specify proper luminance modulation and contrast ratios (highest luminance to lowest luminance). For example, the *American National Standards for Human Factors Engineering of Visual Display Terminal Workstations* states that character luminance modulation must be greater than or equal to 0.5, which is equivalent to a contrast ratio of 3:1. It also states that a preferred modulation is 0.75 or 7:1. The Illuminating Engineering Society of North America (IES) also recommends contrast ratios for visual display terminals. Its standard states that contrast ratios should range from 3:1 to 15:1, and that the preferred range is 5:1 to 10:1 (IES RP24, 1989). It is important to note that these documents came out in the late 1980's when most displays used light characters on a dark background (negative contrast). Today, people mainly use dark characters on a light background (positive contrast). The IES has, however, made a recommendation for positive contrast displays. It states that the preferred range is also 1:5 to 1:10.

Methods of Reducing Glare

Although it is always best to reduce glare at its source, such as reducing the ambient illumination or covering windows, this approach is not always practical for all VDT workplaces. Such is the case where there are other tasks that need to be performed in conjunction with VDT work, which require sufficient ambient illumination. These tasks include reading documents, looking at the keyboard, filing, and safely walking around the work area. The following is a discussion of methods that reduce glare at the VDT screen.

There are currently two main approaches to reducing the glare from CRTs. The first approach involves methods which treat the first surface (glass face plate) of the CRT screen and are performed by the original equipment manufacturer (OEM). These methods include etching and quarter-wave thin film coatings. The second approach uses various after-market, stand alone filters which are placed in front of the CRT screen to reduce glare. Examples of these glare filters include: neutral density filters, polarization filters, and micromesh filters.

It is important to note that in order for a CRT anti-reflection treatment or glare reduction method to be effective it must do two things. It must first enhance the display contrast by reducing diffuse reflections from the phosphor surface. Additionally, it must reduce the intensity and the sharpness of specular reflections from the glass face-plate of the CRT (Beaton, et al., 1985).

One method of reducing glare by treating the glass face plate of the CRT is to etch or roughen its surface. Etching is a chemical or mechanical treatment applied to the outer layer of the glass surface. An advantage of this method is that it blurs the edges of specular reflections, making them appear diffuse. Etching can reduce the reflectivity of the glass from 4% to 2% (Cakir et al., 1980). There are, however, some disadvantages to etching. Not only does etching blur the reflected image, but also may blur the image emitted from the phosphor screen. Therefore, it can reduce the legibility of the screen. Another disadvantage is that etching does not improve the contrast of the display by attenuating diffuse reflections at the phosphor surface.

Quarter-wave thin film coatings are another anti-reflection treatment that reduces specular reflections. It acquired its name because the film thickness is equal to one quarter the wavelength of light. Quarter-wave thin film coatings reduce specular reflections by setting up an interference pattern for light. Because air to coating interface has a different index of refraction then the air to glass interface, the incoming ambient light wave will interfere with a reflected light wave from the glass and cancel each other. The operator will never perceive this light. An advantage of this method is that it does not blur images produced from the phosphor screen. There also are disadvantages in using this method. One disadvantage is that it does not reduce diffuse reflections at the phosphor surface. Another disadvantage is that quarter wave coatings do not reduce the sharpness of the specular reflections. A final disadvantage is that the coating is highly susceptible to fingerprints, dust, and scratches. In other words, if the screen becomes soiled, then the phosphor image will tend to look smeared.

Another anti-reflection technology is a neutral density filter. A perfect neutral density filter works by absorbing equal amounts of light from each wavelength in the visual spectrum. However, neutral density filters are not perfect and they tend to lose some of their absorption properties at either end of the visual spectrum. This can be a problem if an operator needs true color representation. Neutral density filters also improve the contrast of the display by reducing ambient light twice, once by the filter on its way to the phosphor surface and then again when it is reflected back from the surface. The screen luminance, however, is reduced only once through the filter. Since the ambient light is filtered more than the screen luminance, the contrast is improved. Figure 1 illustrates how a filter reduces light transmission effect. A problem with neutral density filters is that they are made from plastic, glass, or gelatin material and their first surface can be a source of specular reflections.

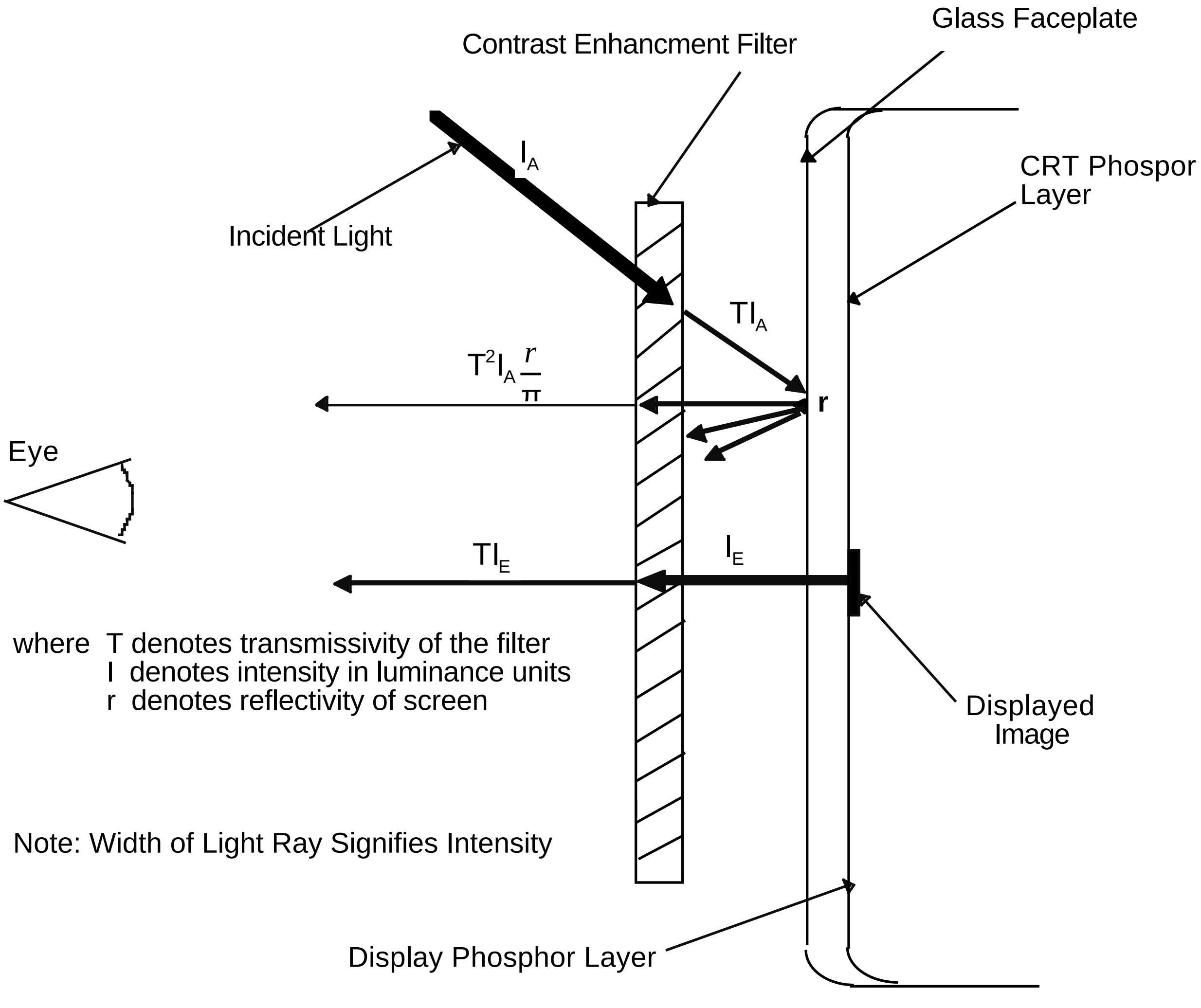


Figure 1. Contrast enhancement through the use of neutral density and polarization filters.

Polarization filters also improve the contrast of the display. To understand how a polarization filter works, a general discussion of light is needed. Light can be considered a wave form that oscillates in any direction such as up/down, left/right, and all directions in between the two. A polarizing filter only allows the waves that are traveling in the polarizing direction of the filter to pass through it. Therefore, a polarization filter improves contrast the same way as a neutral density filter: first by reducing ambient light as it passes through the filter to the phosphor surface and then again, as it passes back through the filter and to the observer. Again, the image luminance is only filtered once so characters appear more distinct on the screen. One problem with polarization filters is that they have highly polished and reflective first surfaces that makes them extremely susceptible to specular reflections. One way to remedy the problem is to apply an anti-reflection coating to the first surface.

A final glare reduction technology is to use micromesh filters. These filters are made of fine fabrics which resemble black nylon stockings stretched over CRT screens. Micromesh filters only allow light that is perpendicular to the mesh to pass through it and it absorbs or scatters the rest of the light. The ambient light that does pass through the filter is either reflected from the glass surface or the phosphor layer. In either case, the reflected light is further reduced by passing through the filter once again while the screen luminance is only passed though the filter once. Thus, a micromesh filter reduces specular reflections and enhances display contrast. However, this filter does have some drawbacks. The overall display luminance is reduced when the screen is viewed from oblique (indirect) angles. Thus, unless the display is viewed head-on, there may not be enough luminance contrast to read the screen. A second drawback is that micromesh filters are difficult to clean and tend to collect dust and dirt, which further reduces luminance from the screen.

Methods of Physically Measuring Glare

There are several methods to quantitatively evaluate glare on a visual display. This background review will focus on two methods. One is the reflectance transfer function and the other is the ISO 9241 part 7 standard.

The *reflectance transfer function* (RTF) measures both the intensity and the sharpness of the first surface specular image (Beaton and Snyder, 1984). However, it does not directly measure the loss in contrast that display images suffer at the phosphor layer. RTF is determined by a Fourier Transform and either a line spread function or an edge spread function. RTF is defined as “the modulation spectrum of a reflected edge-target [or line-target] image divided by the modulation spectrum of the edge [or line] image measured directly” (Beaton 1993). Essentially, RTF describes how efficiently the display produces a specular image. Therefore, higher RTF values indicate sharper and more intense images produced on the first surface, an undesirable quality.

Another method of physically measuring reflections on displays or filters is the recently approved ISO 9241 part 7. The purpose of ISO 9241-7 is to determine acceptable VDT image quality when the VDT is subjected to certain glare conditions. The ISO standard classifies or predicts the image quality of VDT monitors based on their usefulness in particular office environments. The ISO standard states that all monitors must comply with one of three classes: Class I, II, or III. Class I monitors have the best image quality under glare conditions and are considered suitable for general office use. Class II monitors are suitable for most office environments, but have poorer image quality than Class I under the same glare conditions. Class III monitors have the worst image quality under glare conditions and are only suitable for specially controlled luminous environments (ISO 9241-7, 1997).

The ISO standard measures glare on a VDT screen or filter using luminance ratios. Luminance ratios are similar to contrast ratios but they are not the same. The ISO standard defines a luminance ratio as "the ratio between two luminances, either or both of which may be the composite of two or more other luminances" (ISO 9241-7, 1997). The ISO standard states two types of luminance ratios that must be met.

One type is the *screen image luminance ratio* with specular and diffuse reflections. The screen image luminance ratio is the ratio between the higher and lower luminances of an image that has both specular and diffuse reflections on it. In other words, the ratio not only includes the luminances of the image but also the specular and diffuse reflected luminances. The image stated by the ISO standard is either a dark square of pixels surrounded by a bright patch of pixels (positive polarity) or the image is a bright square of pixels surrounded by a dark patch of pixels (negative polarity). Thus, the screen image luminance ratio is dependent on the polarity of the screen. These formulas can be expressed mathematically as:

$$\text{Screen Image Luminance Ratio} = \frac{L_{background} + L_{diffuse} + L_{specular}}{L_{foreground} + L_{diffuse} + L_{specular}} \geq 3 \text{(positive polarity)} \qquad \text{(Eq. 2a)}$$

$$\text{Screen Image Luminance Ratio} = \frac{L_{foreground} + L_{diffuse} + L_{specular}}{L_{background} + L_{diffuse} + L_{specular}} \geq 3 \text{(negative polarity)} \qquad \text{(Eq. 2b)}$$

where $L_{background}$ is the luminance of the background of the image
$L_{foreground}$ is the luminance of the foreground of the image
$L_{diffuse}$ is the diffuse reflected luminance
$L_{specular}$ is the specular reflected luminance

The second type of luminance ratio is the *specular reflection luminance ratio*. It is the ratio between the luminance of the background of an image with specular and diffuse reflections on it to the luminance of the background with only diffuse reflections on it. Therefore, the formula does not change for the different polarities, but the required limit does change. The required limit is different because positive polarity displays do not show specular reflections as well as do negative polarity displays (Sanders and McCormick, 1993). Therefore, if a display has specular reflections, it is better to use positive polarity than negative polarity. The formulas for the specular reflection luminance ratio are as follows:

For positive polarity,

Specular Reflection Luminance Ratio = $$\frac{L_{background} + L_{diiffuse} + L_{specular}}{L_{background} + L_{diifuse}} \leq 1.25 \qquad \text{(Eq. 3a)}$$

For negative polarity,

Specular Reflection Luminance Ratio =

$$\frac{L_{background} + L_{diiffuse} + L_{specular}}{L_{background} + L_{diffuse}} \leq 1.2 + \frac{\left(\dfrac{L_{foreground} + L_{diffuse}}{L_{background} + L_{diffuse}}\right)}{15} \qquad \text{(Eq. 3b)}$$

where $L_{background}$ is the luminance of the background of the image
$L_{foreground}$ is the luminance of the foreground of the image
$L_{diffuse}$ is the diffuse reflected luminance
$L_{specular}$ is the specular reflected luminance

The ISO standard utilizes a combination of luminance ratios, monitor polarities, and glare source aperture sizes to evaluate monitor performance. In particular, the monitor class (Class I, II, or III) is determined only according to the luminance ratios acquired at different glare source sizes. And this class must be determined separately for both positive and negative monitor polarities, so that a display might be rated Class I for positive polarity and Class II for negative polarity. The exact determination of a display class involves applying two different glare source test conditions: large and small glare source. These two test conditions use large (15°) and small (1°) aperture integrating spheres as the glare sources, respectively. The screen image and specular reflection luminance ratios then are calculated under both the large and small glare source test conditions. In summary, eight different luminance ratios need to be calculated for each monitor: four ratios each for positive and negative polarity.

To be Class I, the polarity of the monitor must pass the screen image and the specular reflection luminance ratio for both the large and the small glare sources. A Class II polarity only has to pass the screen image luminance ratio and the specular reflection luminance ratio for either the large or the small sources. In other words, it can pass the screen image luminance ratio and the specular reflection luminance ratio for the large source and fail one or both ratios for the small source. The opposite also can be true. The polarity of the monitor can pass the ratios for the small source and fail one or both for the large source. A Class III polarity of the monitor also only has to pass the screen image and the specular reflection luminance ratio for either the large or the small sources, but the calculations for the luminance ratios use different luminance levels for the sources than Class I or II. Table 1 shows the luminance levels used in the calculations of the luminance ratios for the three classes and the two sources. These luminance levels are different from those used in taking the luminance measurements. While taking the measurements, the luminance of the sources are set equal to or greater than 2000 cd/m^2.

Table 1. Luminance Values Used in the Calculations of ISO Luminance Ratios

Class	**Requirement**	**Luminance of Large Sources (cd/m^2)**	**Luminance of Small Sources (cd/m^2)**
I	must pass luminance ratios for both sources	200	2000
II	pass luminance ratios for either source but not both	200	2000
III	pass luminance ratios for either source but not both	125	200

The discussion of luminance ratios demonstrates that the ISO standard characterizes specular and diffuse reflections in one out of two respects. As mentioned earlier, for a CRT anti-reflection treatment or glare reduction method to be effective it must do two things. It must first enhance the display's contrast by reducing diffuse reflections from the phosphor surface. Additionally, it must reduce the intensity and the sharpness of specular reflections from the glass face-plate of the CRT. First, the ISO standard measures the contrast of the display using the screen image luminance ratio. Second, it measures the intensity of the specular image using the specular reflection luminance ratio. However, the ISO standard fails to analyze the sharpness of the specular image.

LITERATURE REVIEW

Several interesting observations were found during a review of the literature regarding the effects of CRT glare and glare attenuation techniques. Overall, most of the researchers agree that high ambient illumination ultimately is responsible for loss in screen contrast and for formation of specular reflections. However, investigators do not entirely agree upon the most effective methods to achieve contrast enhancement and to reduce specular reflections. A notable downfall of most of the studies' reviewed was that they were only meaningful for specific situations and did not allow for generalization to other luminous environments. In other words, because these studies did nothing to characterize the reflectivity of display surfaces, predictions could not be made as to how CRTs perform in other luminous environments. The Reflectance Transfer Function (RTF), however, was found to quantitatively characterize a monitor’s ability to diffuse (blur) and attenuate first surface reflections.

The studies could be described as belonging to two types of investigations. The studies either approached the glare problem from a hardware (physical measurement), photometric measurement, and analytical perspective or by way of human subjective and objective performance measures. Unfortunately, only a few studies were able to effectively incorporate both approaches. In other words, only a couple studies were able to characterize the reflected light in a quantitative manner, thereby allowing its effect on human performance to be determined. Also, ambient lighting conditions and controls seemed to significantly differ between studies in terms of glare source design and incident angle to either an observer or photometer. These ambient lighting condition inconsistencies made it difficult to compare the results of one study to another because, in general, glare increases with the luminance, size, angle of incidence, and proximity of a glare source to an operator’s line of sight (ANSI/HFES 100-1988).

Physical Measurement Studies

Beaton and Snyder (1984). Beaton and Snyder described a series of experiments in which they measured the capabilities of 10 anti-glare filters. Specifically, they measured the ability of the filters to suppress specular reflections. Additionally, they measured whether the displayed image quality was degraded by the filters. Finally, they measured the display image luminance lost by adding the filters. The glare filters were placed on polished and etched CRTs and measurements were performed for a number of off-axis glare source angles.

Modulation transfer functions (MTF) were obtained for each of the CRTs while the glare filter was fixed in front of it. In general, the MTF concept is an image quality metric that shows the amount of achievable luminous power at each of the spatial frequencies of interest. In basic terms, display systems that are capable of transmitting sufficient luminous power at high spatial frequencies are perceived by the observer as having sharper edges. Those that are incapable of transmitting high frequencies appear blurred (Hunter, 1988). Therefore, the MTF for glare filters indicates whether a filter tends to degrade the displayed images by scattering or diffusing them. The four off-axis vertical angles used for these measurements were 15, 30, 45, and 60 degrees. In addition, the author reported that the maximum display luminance intensity of 86.68 cd/m^2 for the no-filter condition. Quantitatively documenting display maximum luminance intensity is an important control for determining the effects of glare (contrast loss) on a particular CRT's performance.

Beaton and Snyder found the circular polarizer and quarterwave coated neutral density filters produce the least reduction in image quality overall. The micromesh filters, in comparison, produced the greatest reduction. Moreover, because of its diffusing effect, the etched CRT produced lower MTF values than the polished display did overall.

RTFs also were collected for each of the display and filter combinations. The RTF concept is like the MTF in that they are both spatial frequency measures. The difference is that the RTF indicates the surface's ability to diffuse or eliminate specular reflection independent from displayed image quality. Moreover, small RTF values describe surfaces that effectively reduce specular reflection by eliminating high spatial frequency information, unlike the MTF test for which larger MTF values are viewed desirable. The three horizontal angles used were 20, 30, and 45 degrees. The luminance of the glare source was 3000 cd/m^2. Beaton and Snyder found that the etched CRT diffused specular reflections more than the polished did, regardless of the glare filter used.

Lastly, Beaton and Snyder produced a signal-to-noise ratio (SNR) by dividing the measure of the displayed image quality (MTF) by the measure of the "noise" (RTF). These ratios describe the relative improvement in display quality that a particular glare filter produces by attenuating glare sources. Beaton and Snyder found that the SNR of the etched CRT proved to be greater than the SNR for the polished CRT due to its capacity for diffusing specular reflections. In addition, the quarterwave coated neutral density filter was found to produce the largest SNR with the etched CRT, whereas the micromesh, green and high neutral density filters performed well with the polished CRT. Notable is that only the reflected image sharpness component of the CRT's first surface reflectance was quantitatively characterized by the authors (i.e., RTF measurements), while diffuse reflections and display surface diffuse reflectance was not.

Rancourt, Grenewalt, Hunter, and Snyder (1986). Rancourt, Grenawalt, Hunter, and Snyder (1986) studied the effects of several glare reduction techniques including: chemical etching, chemical etching combined with a quarterwave filter, and a polished surface with a quarterwave filter. They compared these three techniques to a polished CRT alone in terms of the ability of each to reduce specular reflections without degrading image quality. Rancourt et al. used the MTF and RTF metrics previously mentioned in the Beaton and Snyder study (1984) and found considerable differences between the treatment conditions. The MTF measurements were conducted with a collimated incandescent light that was varied between 20 and 50 degrees from normal, under overhead fluorescent lighting, and in a dark room conditions.

Rancourt et al. found that the etched surface reduced the MTF by about 30% over the polished surface, with or without a quarterwave filter. In addition, they found that placing the quarterwave filter on either the etched or polished CRT surfaces had no effect on image quality under any of the lighting conditions tested. Moreover, glare angle was found to have no effect. The RTF measurements of specular glare suppression ability suggested that the etched, etched with quarterwave, and polished with quarterwave were all much more effective at reducing reflection than the polished only condition. Also, calculated signal-to-noise ratios (same as in Beaton and Snyder, 1984) indicated that the etched conditions caused the highest overall image quality and that adding an anti-reflection coating to either an etched or polished CRT surface increases both of their SNRs. It is important to note that only the reflected image sharpness component of the CRT's first surface reflectance was quantitatively characterized by the authors (i.e., RTF measurements), while diffuse reflections and display surface diffuse reflectance was not.

Stammerjohn, Smith, and Cohen (1981). Stammerjohn et al. (1981) investigated VDT workstation designs at five San Francisco area businesses to determine their consistency with recommended design specifications. Measurements were taken and reported on workplace illumination and glare for each of the workplaces. VDT operators were then surveyed to identify factors which contribute to worker dissatisfaction, annoyance, and health complaints. Stammerjohn et al. found illumination levels ranging from 300 to over 1000 lx, 75% of which had illumination levels of 501-700 lx. In 1984, Snyder recommended workplace illumination levels of about 200 lx for workstation with supplementary lighting. Cakir et al. (1979) suggest that illumination levels between 300 and 500 lx are sufficient for VDT workstations without supplementary lighting. The Human Factors Society (ANSI, 1988), along with Sanders and McCormick (1993) has recommend ambient illumination levels of 200-500 lx. An important finding of the Stammerjohn et al. investigation was that a majority of the workstations had levels above the recommended lighting requirements.

The glare measurements taken in this study are representative of early attempts to quantitatively measure specular and diffuse reflections on VDT screens and relate their effects to a subjective image quality evaluation (glare questionnaire). Results from the questionnaire included the finding that 17% of the VDT screens had reflected glare that could make reading characters on parts of the screen difficult. Of particular interest, were the findings that 85% percent of the participants complained of screen glare and 70 % of character brightness (presumably resulting from the loss in contrast from relatively high ambient illumination levels). Glare sources for the workstations included windows and light fixtures, with luminance levels of up to 2100 cd/m^2. Specular and diffuse components of reflections were measured directly together, at one of the sites, with a hand-held photometer. Specifically, the measurements included the maximum reflected luminance levels of reflections on VDT screens with the display turned off. They recorded luminance values between 3 and 50 cd/m^2.

Human Performance Studies

Beaton, Murch, and Knox (1985). Beaton, Murch, and Knox (1985) conducted three studies to determine the effects of various anti-glare filters on perceived image quality under dark, diffuse, and specular lighting conditions. Moreover, a different CRT was used in each of the studies in order to determine if the effects of the filters are consistent for different CRT technologies.

The conditions for the first study consisted of a high resolution monochrome CRT with one of the following placed in front of it: a neutral density spray filter, a rectangular oriented micromesh filter, a mechanically abraded glass filter, a quarterwave coated glass filter, or a quarterwave coated filter with a low neutral density layer. The maximum screen luminance produced by this CRT was 68.5 cd/m^2. A non-filtered, polished CRT display condition served as a baseline for this experiment. This study presented a series of paired comparisons to the participant under both dark and diffuse ambient lighting conditions (900 lx from two glare sources displaced off-axis 0.86m vertically and +/- 1,37m horizontally). Beaton et al. found that the spray and etched filters yielded the lowest perceived image quality ratings when viewed in the dark, whereas the quarterwave with the low neutral density layer was found to produce perceived image quality improvements over the baseline condition. None of the other four filters were found to be significantly different from the baseline condition. For the conditions employing the diffuse glare source, the spray filter was judged to degrade the displayed image. Whereas the quarterwave filter with a neutral density layer was found to enhance a displayed image's quality. The authors believed that the softening or blurring produced by the etch and the luminance nonuniformity associated with the spray filter, were what caused the perceived image degradation.

The second study was conducted using a medium resolution color CRT to investigate the effects of the same six filters used in the first experiment along with a louvered filter, a green density filter, a diagonally oriented mesh filter, a glass circular polarizer, and a quarterwave coated filter with a high neutral density layer. The physical characteristics (maximum display luminance and glare source orientation) of the first study were employed here again in this study.

In this experiment, participants were asked to compare each filter to the no-filter baseline condition using an 11-point rating scale ranging from extreme degradation to extreme improvement. Participants judged the quarterwave with a high density layer to mildly degrade the image, whereas the rectangular mesh, the etch, and the louvered filter moderately degraded display images. As in the first experiment, the spray filter yielded the most degraded images. Also, the spray was again found to reduce image quality in the diffuse glare source condition, whereas the quarterwave coated filter and the diagonal mesh produced moderate improvement. The quarterwave filters with the low and high neutral density layers and the circular polarizer were judged to greatly enhance image quality relative to the no-filter condition. The most important finding of this study was that the differences between the filters could not have been a function of CRT type or bandwidth settings.

A third study was conducted using a high resolution monochrome CRT and the six filters which produced image quality enhancement capability. The purpose of this experiment was to determine the effects of uniform and patterned specular reflection on perceived image quality. These two specular reflection types constitute the two glare levels used. The uniform glare source consisted of a large 48% reflective gray-colored matte board illuminated with 1500 lx of diffuse light which was located 1.52m behind the subject. For the Patterned glare condition, the matte board consisted of alternating white (90% reflective) and black (5% reflective) stripes. This large matte board also was illuminated with 1500 lx of diffuse light and the stripes were 10.16 cm wide, thereby forming a vertical square-wave grating pattern.

Beaton et al. found that the circular polarizer degraded perceived image quality for the uniform glare condition, whereas the quarterwave coating with either low or high neutral density layers were found to enhance the displayed information. According to the authors, one possible reason for this finding was that the quarterwave coating reduced the first surface reflectance and the neutral density layer provided contrast enhancement. In the case of the circular polarizer, contrast enhancement was achieved but specular reflection reduction was not. The patterned specular reflection condition yielded similar results for all of the filters but the diagonal micromesh filter, which received a slightly higher image quality rating than the no-filter baseline condition. The authors believe that this finding suggests that the mesh was able to attenuate the high spatial frequency information of the patterned reflection. Another important conclusion on the part of Beaton et al. is that contrast enhancement and anti-reflection treatments are both needed to improve image quality in the presence of a specular glare source. This fact was demonstrated by the subpar performance of the circular polarizer and quarterwave filter with no density layer.

In summary, there were three findings from this study that had far-reaching implications for future research. First, glare filter investigations must have a dark or low ambient lighting condition for a baseline. Secondly, both diffuse and specular glare sources must be used to test whether the filter can both enhance contrast and cut specular (patterned) reflections effectively. Third, the validity of the human performance measures used was found to be high. In other words, VDT operators can discriminate and detect even small differences between lighting filter conditions regardless of the psychophysical data collection procedure employed.

Garcia and Weirwille (1985). Garcia and Wierwille (1985) investigated the effects of 'no-glare' and 'glare' lighting conditions on reading response time and proportion of correct responses. It is important to note that diffuse reflections were actually present in their 'no-glare' condition while specular reflections were present in their 'glare' condition. The same glare sources were used for both the 'no-glare' (diffuse) and 'glare' (specular) experimental conditions. They included two sets of fluorescent tube lights recessed in the ceiling (located above the VDT) and several fluorescent tubes suspended from the ceiling behind the operator. However, a shield was used for the 'no-glare' (diffuse) condition to redirect the light in order to prevent direct light from striking the CRT surface. Unfortunately, glare source illuminance and luminance levels as well as incident angles were not reported. Only one CRT was employed in this experiment and no anti-reflection or contrast enhancement treatments were implemented. The difficulty of the reading passages also was varied by the authors for the 'no-glare' (diffuse) and 'glare' (specular) lighting treatments.

For the human perform measurement procedure, the participants were asked to read each of the passages very carefully because their reading comprehension was going to be tested. The specular lighting conditions yielded the most interesting results of the study. There was more time required to read 'easy' passages under specular conditions than diffuse. However, the time required to read ‘difficult’ passages surprisingly decreased. This finding is inconsistent with prior research that would expect an intense specular reflection to have either no effect or an increased reading time. Garcia and Wierwille rationalize this finding by saying that the participants may have been so annoyed by the intensity of the glare source that they rushed through the reading task in an effort to end their contact with the experimental condition quickly. If their explanation had been accurate, than more incorrect responses would have been recorded for the more difficult passages. However, there was no increase in incorrect responses for the more difficult passages.

Isensee and Bennett (1983). Isensee and Bennett investigate the effects of diffuse reflected glare on subjective discomfort ratings under two video polarities, three video luminance, and three ambient illumination conditions. The polarities of video were negative (light characters on a dark background) and positive (dark characters on a light background). The levels of video luminance (photometrically measured) were 120.1, 65.2, and 10.3 cd/m2. The levels of ambient illuminance, measure at the keyboard, were 420, 260, and 100 lx. Isensee and Bennett (1983) did, however, note that ambient illumination levels measured at the center of and parallel to the front surface of the CRT produced readings that were around 20% smaller than those at the keyboard. Also, the illumination was emitted by a single four-lamp fluorescent ceiling source located directly above the CRT. Finally, the authors ensured the absence of specular reflections from light colored clothing by having the participants wear a black cape. This eliminates a confounding variable.

Isensee and Bennett found that reflected glare was judged to be less uncomfortable at lower levels of ambient illuminance and higher levels of video luminance, overall. Also, they found a significant interaction between video luminance and ambient illumination. Specifically, as video luminance increased, the average discomfort rating for the levels of ambient illuminance were more similar. In other words, the average discomfort ratings for the illuminance levels was less variable as video luminance amounts were increased. At that time, Isensee and Bennett believed this finding was an artifact caused by the ratings approaching the low end of the discomfort scale (floor effect). However, this finding can be explained by the fact that higher video luminance usually means higher contrast capability. Further, higher video luminance output (modulation) can help lessen the effects of diffuse reflections that act to reduce the contrast of displayed images. Another important finding was that positive polarity video was preferred to negative polarity in terms of comfort for all the combinations of video luminance and ambient illumination.

Habinek, Jacobson, Miller, and Suther (1982). Habinek et al. (1982) studied the effects of three anti-reflection treatments, two display polarities (positive and negative), and two lighting conditions (‘no-glare’ and ‘glare’) on correct reading rates (CRR) and subjective preference ratings. Just as in the Garcia and Weirwille study, diffuse reflections were actually present in their ‘no-glare’ condition while specular reflections were present in their ‘glare’ condition. Three identical Sylvania CRTs were fitted, respectively, with a micromesh filter, a quarterwave coated filter, and a no-filter. The no-filter polished CRT served as the baseline condition. A fourth condition of a Clinton CRT with an etched front surface was also employed. The ‘no-glare’ (diffuse) condition employed fluorescent lighting located directly over the VDT, while the ‘glare’ (specular) condition used a bank of bare fluorescent tubes suspended from the ceiling behind the participant. Both sources provided 1450 lx of ambient illumination. In addition, glare source incident angle, operator line of sight, and maximum screen luminances were not recorded.

For the ‘no-glare’ (diffuse) lighting condition, no CRR differences were found among the three anti-reflection treatments and the baseline condition. However, significant differences were realized for the preference data. Specifically, the polished baseline treatment was judged to have “sharper” characters. For the ‘glare’ (specular) lighting condition, the micromesh filter produced the highest CRR, while the polished baseline produced the lowest CRR. Subjective preference results indicated that participants found the micromesh filter display to be the easiest to read under ‘glare’ (specular) lighting, whereas the polished CRT was significantly more difficult to read than the others.

In summary, the authors could not recommend one anti-reflection treatment over another treatment, but they were all rated as superior to an unpolished CRT when specular glare is present. However, there is one main problem with the methodology the authors employed. They may have confounded the CRT anti-reflection treatments with the recorded contrast ratios because the subjects were allowed to adjust the brightness controls of the CRTs for the each of the conditions.

Morse (1985). Morse (1985) studied the effects of six anti-glare filters on subjective evaluations of the brightness, sharpness, contrast, color, glare and preference for each filter. In addition, photometric measurements of luminance, MTF, contrast, and glare were make in order to identify the optical characteristics of the filters that are responsible for particular subjective preferences. A quarterwave filter with a neutral density layer, a micromesh filter , a circular polarizing filter, a quarterwave filter with a blue density layer, an etched filter, and a no-filter polished CRT surface were included in the investigation. The specular glare source used consisted of fluorescent light fixture with two 48 inch tubes reflected from a piece of white fome-core (48 X 18 inches) suspended above and behind the displays. Overhead fluorescent lights provided additional ambient lighting. Unfortunately, no low ambient or dark condition was administered. For each pair of filter-CRT combinations, participants were instructed to make a comparison as to the brightness, sharpness, contrast, color, glare, and overall preference.

Morse did not directly measure the MTF of the display the way that Beaton and Snyder (1984) and Hunter et al. (1987) did. Instead, he calculated the MTF using a prediction equation that takes viewing distance, spot luminance, and average ambient illuminance into account. Therefore, the calculated MTF values could not account for the reflected glare differences or changes in resolution and luminance that resulted from the participants adjusting the brightness controls.

Morse found that, for the preference results, the quarterwave filter with a neutral density layer, the quarterwave filter with a blue density layer, the micromesh filter, and the circular polarizing filter were all rated moderately high for all six criteria. However, the etched filter and no-filter conditions were rated significantly lower for most of the criteria. Calculated estimates indicated that the quarterwave filters produced the highest MTF values, whereas the etched surface produced the lowest MTF. Also, the quarterwave filters, circular polarizer, and micromesh filters were all much more effective in reducing reflected luminance levels than were the etched and polished conditions. Calculated correlations between the subjective and objective measures yielded moderately high r values.

Physical (Photometric) and Human Performance Measurement Studies

Hunter (1988). Hunter's dissertations was composed of two experimental phases. One objective of the first phase was to investigate the effects 16 glare filters with different surface treatments and transmissivity ratings. This group of filters included three polished with 31-92%, six etched with 31-92%, one polished AR coated with 62%, one etched AR coated with 62%, and two mesh with 50 & 37% transmissivity. The polished glass filter with a 92% transmissivity rating served as a baseline for all the measurements in this phase. Three ambient lighting conditions and two monochrome CRT resolution conditions were also investigated for the filters. The second objective of Hunter's first phase was to quantitatively model the relationship between the measured image quality of the display/filter combinations and the recorded human performance data. His models were based on signal to noise measurements similar to the SNRs used by Beaton and Snyder (1984).

The glare sources employed in the phase one experiments consisted of two, fluorescent light fixtures. One was located directly above the CRT and provided an approximately 650 lx source of diffuse illumination. The other fixture was placed directly behind and above the participant producing a specular image in their line-of-sight. This specular glare source produced a reflection of sufficient luminance to simulate an environment with a poorly located lighting fixture or window.

One of Hunter's general findings was that none of the treated filters significantly improved readability and legibility performance more than the polished glass filter that served as the baseline. However, he did identify various filter treatments which severely degraded performance. Specifically, the etched and low transmission filters degraded human performance over the baseline. Another finding was that the mesh and quarterwave filters improved perceived image quality for the specular glare condition. Hunter's modeling efforts produced mixed results. Although good fit models were produced for perceived image quality, the models Hunter developed for the reading and legibility tasks proved to have little utility.

Hunter's second experimental phase examined the separate and combined effects that filter transmission and diffuse illumination have on readability, legibility, and perceived image quality. Specifically, five contrast enhancement filters with transmissivity rating between 11% to 92% were measured under five ambient illumination experimental conditions. The diffuse glare source used in his second phase consisted of a four-lamp, fluorescent light fixture placed above the CRT. The five discrete levels of illumination used were 0 lx, 320 lx, 1200 lx, 2000 lx, and 2800 lx.

This experimental phase indicated that even during occurrences of extreme contrast loss, the participants were able to maintain good reading and legibility performance. In addition, an inverse correlation was found between illuminance level and perceived image quality. It is important to note that Hunter did not control for specular reflections from light clothing in this diffuse glare experiment, as Isensee and Bennett (1983) did by placing a black cape over their participants. Also, Hunter makes no attempt to quantitatively characterize CRT phosphor surface diffuse reflectance.

Lloyd, Mizukami, and Boyce (1996) . Lloyd et al. propose an ambient lighting-display parameter interaction model. It has the potential to provide display and glare filter manufactures with the necessary tools for predicting and assessing the improvements of their product's performance in work environments susceptible to reflected glare. The authors developed an equation which, in theory, predicts the subjective image quality of displays with reflections.

Predicted conspicuity or disturbance ratings (Rp) are calculated as follows:

$$R_p = b_1 * M_s + b_2 / M_{da} + b_3 * W_b / M_{da} + b_4 * M_s * W_b \qquad \text{(Eq. 4)}$$

where M_s is the modulation of the reflected image
M_{da} is the modulation of displayed image when reflections are present
W_b is the width of blur of the specular image
b_1 is 16.31
b_2 is 0.7575
b_3 is -0.003726
b_4 is -0.06572

The authors identified three display parameters that are needed for the equation. The authors took measurements of six different displays so that they could calculate the three display parameters. They used three CRT displays and three Liquid Crystal Displays (LCDs). One of the display parameters is the modulation (contrast) of the reflected image (M_s). This indicates the intensity of the reflected image. For example, if a display is used in a dark room, the M_s would equal 0. The M_s value is similar to the specular reflection luminance ratio used in the ISO-9241-7. Another display parameter is the modulation (contrast) of displayed image when diffuse and/or specular reflections are present (M_{da}). This parameter is similar to the screen image luminance ratio used in the ISO 9241-7. The only difference is that this study uses a modulation value and ISO standard uses a ratio. The last display parameter is the width of the blur of the specular image (W_b). A small Wb value indicates that the display or filter does not blur specular images very well. As can be seen, there are a lot of parallels between the physical measures described in this study and those in the ISO standard. The largest difference is that this study addresses the component of specular reflections which the ISO standard does not: reflected image sharpness (edge sharpness).

Lloyd et al. also conducted a human performance study. The purpose was to relate the 3 parameters to subjective analysis of specular glare for the same 6 monitors. Participants were asked to rate the conspicuity or the disturbance of the specular glare under two display polarities for each the six monitors. The glare source used was an integrating sphere and the words "Lighting Research Center" appeared on each of the displays. The specular glare source was positioned with respect to the participant and the display so that the center of its reflected image appeared about 10 cm from the center of the display's active area. Also, the mean illuminance on the display was 120 lx, provided primarily by overhead fluorescent luminares fitted with parabolic louvers. Lloyd et al. found that ratings improved (decrease) as the modulation of the reflected image decreases, the modulation of the displayed image in ambient lighting environment increases, and the width of blur function increases. Additionally, the modulation of the reflected images has a stronger affect on the ratings when the specular reflection is sharper.

There are several problems with this study. One problem is that it uses unrealistic specular glare sources (integrating spheres) which are hardly representative of ambient lighting fixtures in office work environments. Also, a section of the results suggest that the displays may have been given anti-reflection treatments even though the authors don't address whether they had been treated or not. Another problem is that they used both CRTs and LCD's to build their model. The reflective properties of CRT's and LCD's vary greatly. Therefore, they should have built separate models for the two types of displays. A final problem with the study is that the authors did not hold certain variables constant on the displays for the human performance study. These variables included display resolution, line spacing, and character size.

Conclusions and Objective

There are several important themes that can be identified from the literature review of photometric measurement, analytic, and human performance research. These themes are directly relevant to the present research and dictate the design of the experimental investigation.

1. Photometric and human performance measures should both be incorporated in order to characterize reflected light in a quantitative manner, thereby allowing its effect on human performance to be determined.

2. Glare filters performance must be measured under a number of different ambient lighting conditions. In other words, both specular and diffuse glare sources can be used to determine a filter's strengths and weaknesses (many environments contain both). Additionally, a dark or low ambient lighting condition that serves as a baseline must be implemented so that the glare filter can be evaluated under an optimal viewing condition.

3. A range of filter technology should be evaluated. For example, the review of the literature to date showed the need for investigations into the effects of various levels of filter transmission. One advantage of this approach is that it allows for increased generalization to other using environments.

4. Another important consideration is that glare filter evaluations should be carried out on a range of CRTs. The luminance, contrast, uniformity, and resolution characteristics are all capable of affecting the results of the study and are found to vary between monitors. Also, many monitors today have anti-reflection coatings applied to the front of their glass face-plates. Investigating a crossection of contemporary CRTs with various anti-reflection treatments applied to them directly, yields increased generalization to other using environments.

5. The reflected image sharpness (edge sharpness) component of specular reflections should be investigated in order to assess a filter or monitor's ability to diffuse and attenuate first surface reflections. Researchers have noted that a more sharply defined reflected image is perceived as more annoying and noticeable than a less sharply defined image of the same intensity.

6. Finally, some standardization is needed in glare investigations for ambient lighting conditions and controls in terms of glare source design and incident angle to either an observer or photometer. Ambient lighting condition inconsistencies make it difficult to compare the results of one study to another because, in general, glare increases with the luminance, size, angle of incidence, and proximity of a glare source to an operator's line of sight.

The objective of this research was to assess the subjective image quality of CRT displays viewed under ambient lighting conditions and then to evaluate the findings with respect to the ISO 9241-7 standard. The objective of this research is therefore divided into two phases. In Phase 1, seven monitors were evaluated according to the ISO 9241-7 standard to determine whether they are Class I, II, or III. Additionally, anti-reflection filters were attached to monitors to see if they change the class of the monitor. In Phase 2, human performance measurements were taken to determine if people perform differently for Class I, II, or III monitors or monitor-filter combinations under different lighting conditions and different screen polarities. Specifically, subjective scale values were evaluated against two ISO metrics and ISO classifications with correlational techniques to see whether ISO measurement results indexed subjective judgments of image quality.

PHASE 1

Methods for Physical Measurements

Seven monitors were classified according to the ISO 9241-7 standard. For each of these seven monitors, six filters were attached and evaluated according to the ISO 9241-7 standard. In other words, a grand total of 49 different monitor or monitor- filter combinations were classified according to the ISO 9241-7 standard. The specifications for the monitors and the filters are shown in Table 2 and Table 3, respectively.

Table 2. Monitor Specifications

Monitor	Coating
SamSung Model # CSN5987	uncoated
Packard Bell Model # 1020	1-layer spin coat
Sony Trinitron 100	1-layer spin coat
Nokia	2-layer spin coat
Sony Trinitron 200	4-layer PET/AR laminate
Mitsubishi Model# TF870	4-layer direct coat
Panasonic Pana Flat PF 70	AR Bonded Panel

Table 3. Filter Specifications

Filter Brand Name	Coating	Transmissivity
3M Anti-Glare Filter Model BF10	1-sided Anti-Reflection coating	45%
3M Anti-Glare Filter Model AF150	2-sided Anti-Reflection coating	45%
3M Anti-Glare Filter Model AF 100 or AF200 *	2-sided Anti-Reflection coating	31%
3M Anti-Glare Filter Model 400	1-sided Anti-Reflection coating	44%
3M Anti-Glare Filter Model 450	2 sided Anti-Reflection coating	53%
3M Circular Polarizing Filter Model HF 300	Circular polarizer, 1 sided Anti-Reflection coating	42%

* Note: AF 100 and AF 200 are the same filter but the names are different for different size filters.

The display-filter combinations were evaluated for compliance with the ISO 9241-7 standard in a university-based photometric measurement facility. The photometric evaluation required numerous photometric measurements to determine glare attenuation performance. The photometric equipment consisted of a two-dimensional CCD detector (Photometrics, Model AF 200), mounted on a large-area XYZ translation stage (Areotech, Model 101SMB2-HM). The translation stage and photometer were coupled to a vibration-isolated optical bench, which contained a jig for accurate placement of the displays under test (see figure 2). All measurements were performed to the conditions in the ISO 9241-7 standard. Additionally, all photometric equipment was calibrated to NIST-traceable standards for luminance.

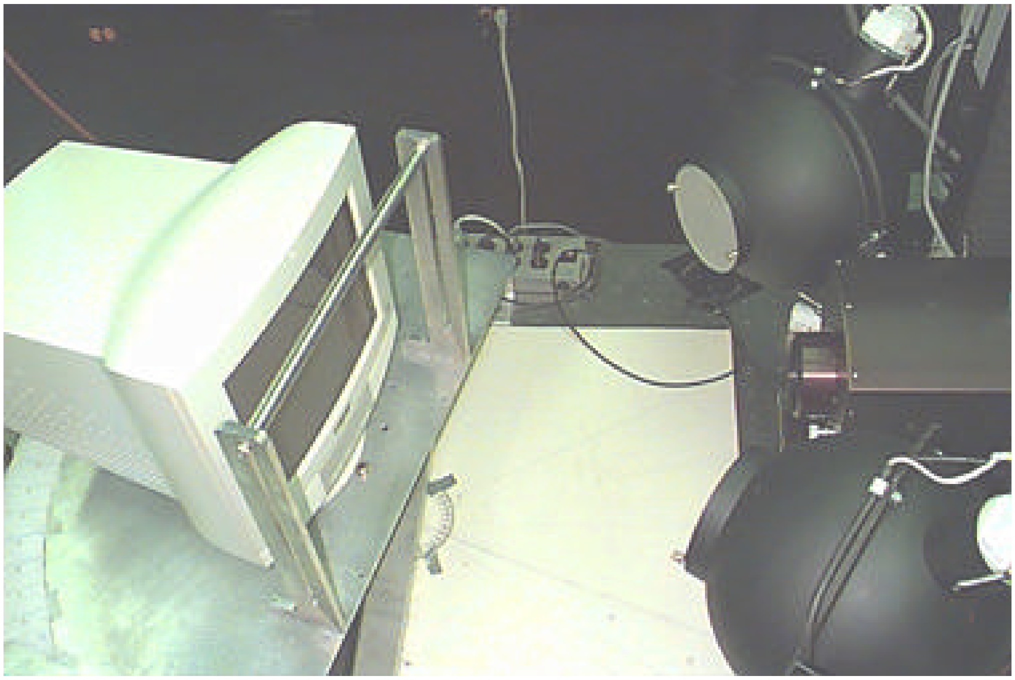


Figure 2. General photometric measurement equipment configuration: including display jig, integrating sheres, and microphotometer.

ISO Test Procedure

According to the ISO 9241-7 standard, 10 measurements are needed from each monitor/filter combination. These 10 measurements were used to calculate four luminance ratios. Table 4 summarizes the major elements of these 10 measurements. The following paragraphs describe the procedures used to take the 10 measurements. Please refer to the standard for more detailed information.

Table 4. Major Measurement Components

Measurement Name	**Light Source**	**Item Measured**	**Item Focused On**
$L_{DS(small,15^{\circ})}$	one small	Display	virtual iimage of source as reflected on the 1st surface of item measured (500 +500=1000 mm)
$L_{D(sTD,\ 15^{\circ})\ small}$	one small	Reflectance Standard	virtual iimage of source as reflected on the 1st surface of item (500 +500=1000 mm)
$L_{DS(extended,\ 15^{\circ})}$	one extended	Display	virtual iimage of source as reflected on the 1st surface of item (500 +500=1000 mm)
$L_{D(STD,\ 15^{\circ})}$ extended	one extended	Reflectance Standard	virtual iimage of source as reflected on the 1st surface of item (500 +500=1000 mm)
$L_{D(0^{\circ})}$	two extended	Display	1st surface of item (500 mm)
$L_{D(sTD,\ 0^{\circ})}$	two extended	Reflectance Standard	1st surface of item (500 mm)
$L_{B(15^{\circ})}$ positive polarity	None	Display	Phosphor dots of CRT (approximately 500 mm)
$L_{F(15^{\circ})}$ positive polarity	None	Display	Phosphor dots of CRT (approximately 500 mm)
$L_{B(15^{\circ})}$ negative polarity	None	Display	Phosphor dots of CRT (approximately 500 mm)
$L_{F(15^{\circ})}$ negative polarity	None	Display	Phosphor dots of CRT (approximately 500 mm)

Small-Source Measurements @ 15 Degrees: Small Specular. This measurement was performed with the monitor off and turned 15 Degrees from normal. The photometer was focused on the monitor's glass surface in the case of monitor measurements and on the glass surface of the filter for monitor-filter combinations. It should be noted that the filters were positioned carefully on the monitor such that two out of the three reflections (filter glass and monitor glass surface reflections) were superimposed just above the center of the screen. A separate luminance measurement was made of the third reflection off the phosphor surface and added to the measurement from the superimposed glass surfaces, as these are additive photometric quantities. The diffuse reflectance standard was then placed in front of the monitor or filter and the $L_{DS(STD, 15)}$ measurement was taken. Figure 2 shows the equipment configuration for these measurements.

Extended-Source Measurements @ 15 Degrees: Large Specular. This measurement was performed with the monitor off and turned 15 Degrees from normal. The photometer was focused on the monitor's glass surface in the case of monitor measurements and on the glass surface of the filter for monitor-filter combinations. It should be noted that the filters were positioned on the monitor such that all three reflections (filter glass, monitor glass and phosphor plane reflections) were superimposed at the center of the screen where all the previous measurements were taken.
The diffuse reflectance standard was then placed in front of the monitor or filter and the $L_{DS(STD, 15)}$ measurement was taken. Figure 3 shows the equipment configuration for these measurements

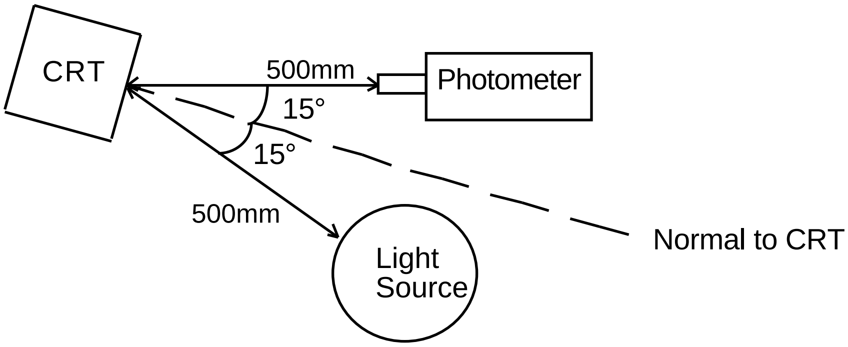


Figure 3. Equipment configuration for specular measurements.

Measurements Perpendicular to Center of Screen. Screen grid coordinates were used to ensure that all photometric measurements were taken at center of the CRT. With both luminance sources turned on and set to 2000 cd/m^2 (15 degree aperture sources), a scan was taken of the diffusely reflected luminance ($L_{D(0)}$). Monitor power was off for this measurement. The diffuse reflectance standard then was placed directly over the center of the CRT screen or filter and a second measurement was made.
Figure 3 shows the equipment configuration for diffuse measurements.

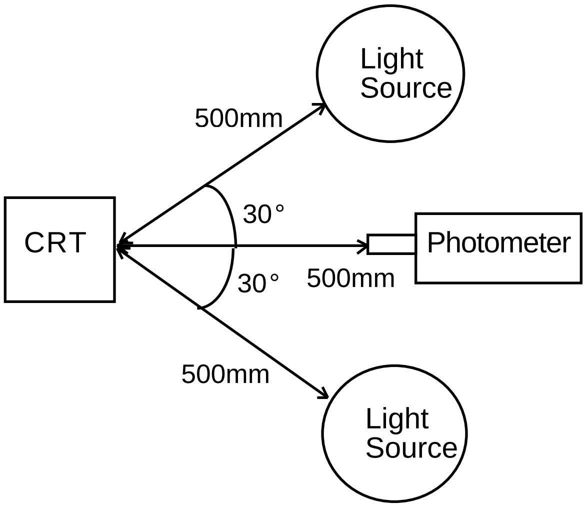


Figure 4. Equipment configuration for diffuse measurements.

Measurements @ 15 Degrees to Center of Screen For Positive Polarity. A five minute waiting period was observed after the monitor was switched to positive polarity. With the same setup mentioned above (except with a black box and a white background), the luminance foreground $L_{F(15)}$ was calculated. The luminance background $L_{B(15)}$ then was measured by selecting the background area and calculating an average area luminance for it. Figure 4 shows the test configuration for these measurements

Measurements @ 15 Degrees to Center of Screen For Negative Polarity. For the next two measurements, the power to the luminance sources remained off and the monitor was turned 15 degrees from normal. After the monitor was turned on and allowed to warm up for 20 minutes, a 50mm white box was displayed on its screen. A single scan was taken of the top portion of the box (white foreground) and an adjacent portion of its black background. The luminance foreground $L_{F(15)}$ was measured by selecting the foreground area and calculating an average area luminance for it. The luminance background $L_{B(15)}$ was then measured by selecting the background area and calculating an average area luminance for it. Figure 5 shows the test configuration for these measurements.

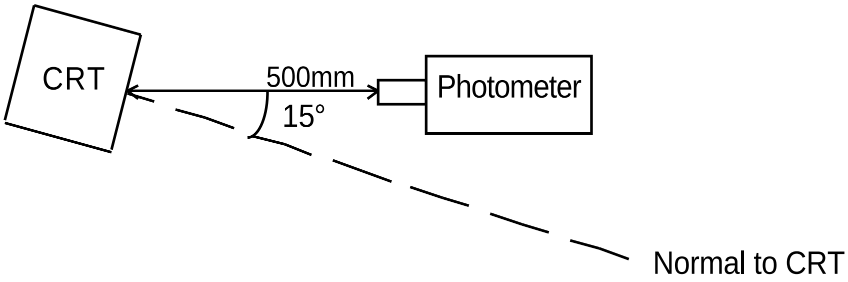


Figure 5. Equipment configuration for polarity measurements.

To ensure the stability of the monitor luminance output, the following additiioinal procedure was implemented after the 20 minute warm-up. First, both the contrast and brightness controls were turned off completely. Then, the luminance was increased until the outer edge of the raster lines were just perceivable in a dark room. Finally, the contrast was turned up until the luminance output for a full white screen reached 75 cd/m^2. This luminance output level was determined by following the above procedure for all the monitors to determine common maximum luminance output level (dictated by the lowest of the seven monitors with the contrast turned all the way up).

RESULTS AND DISCUSSION

After the ten measurements were taken for each monitor or monitor/filter combinations, the luminance ratio equations were calculated for both sources and both polarities. The Classes where then determined for each polarity. Tables 5 shows the results of the ISO standard measurements. Table 6 shows whether adding the a filter improved, degraded, or kept the same Class as the display by itself.

TABLE 5. Monitor and Filter Classifications

	Monitor	AF 150	BF 10	PF 400	PF 450	HF 300	AF 100 or AF 200
SamSung	II	I	II	II	II	II	II
	FAILED	III	FAILED	II	III	II	III
Packard Bell	I	I	II	I	II	II	II
	III	II	III	II	III	II	III
Sony Trinitron 100	I	I	II	II	II	II	II
	III	II	FAILED	II	III	II	III
Nokia	II	I	II	II	II	I	II
	III	III	FAILED	II	II	II	III
Sony Tinitron 200	I	I	II	II	II	I	II
	II	II	III	II	II	II	III
Mitsubishi	I	I	II	II	II	I	II
	II	II	III	II	II	II	III
Panasonic Pana Flat	I	II	II	I	II	I	II
	II	II	III	II	II	II	II

*Positive polarity in white and negative polarity in gray

TABLE 6. Filter Effect on ISO 9241-7 Monitor Classification

	Monitor	AF 150	BF 10	PF 400	PF 450	HF 300	AF 100 or AF 200
SamSung	II	+	0	0	0	0	0
	FAILED	+	0	++	+	++	+
Packard Bell	I	0	-	0	-	-	-
	III	+	0	+	0	+	0
Sony Trinitron 100	I	0	-	-	-	-	-
	III	+	-	+	0	+	0
Nokia	II	+	0	0	0	+	0
	III	0	-	+	+	+	0
Sony Tinitron 200	I	0	-	-	-	0	-
	II	0	-	0	0	0	-
Mitsubishi	I	0	-	-	-	0	-
	II	0	-	0	0	0	-
Panasonic Pana Flat	I	-	-	0	-	0	-
	II	0	-	0	0	0	0

Note: Positive polarity in white and negative polarity in gray
‘+’ means filter improved one class
‘++’ means filter improved two classes
‘0’ means filter kept the same class as the monitor alone
‘-’ means filter degraded one class

Monitor-Filter Combinations

As shown in table 6, certain glare filters work better on some monitors than other monitors. The following paragraphs describe which filters work best and worst on the different CRT displays examined.

SamSung. The AF 150 is the only filter that improved the positive polarity classifications on the SamSung monitor. All of the filters improved the negative polarity class except the BF 10.

Packard Bell. Only the AF 150 and PF 400 kept the Packard Bell monitor a Class I in positive polarity. All other filters degraded the Class. The only filters to improve the negative polarity class were the AF 150 and PF 400.

Sony 100. The AF 150 was the only filter that kept the Sony 100 a Class I in positive polarity. In negative polarity, the AF 150, PF 400, and CP 300 improved the class to Class II.

Nokia. The AF 150 and CP 300 improved the Nokia to a Class I in positive polarity. In contrast, the HF 300, PF 400, and Pf 450 were the only filters to improve the negative polarity class to a Class II.

Sony 200. Only the AF 150 and CP 300 kept the Sony 200 a Class I in positive polarity. All other filters degraded the Class. All filters kept the same negative polarity class as the Sony 200 with the exception of the BF 10 and AF 100, which degraded it to a Class III.

Mitsubishi. Only the AF 150 and CP 300 kept the Mitsubishi a Class I in positive polarity. All other filters degraded the Class. All filters kept the same negative polarity class as the Mitsubishi with the exception of the BF 10 and AF 100, which degraded it to a Class III.

Panasonic. Only the PF 400 and CP 300 kept the Panasonic a Class I in positive polarity. All other filters degraded the Class. All filters kept the same negative polarity class as the Panasonic except the BF 10 which degraded it to a Class III.

Filters

AF 150. In general, AF 150 always improved or remained the same as the monitor itself with one exception: Panasonic positive polarity.

BF 10. In general, this filter degraded the polarity Class of a monitor when placed in front of its screen. It even failed a Class III for the negative polarity when placed on the SamSung, Sony 100, and Nokia.

PF 400. In negative polarity, the PF 400 improved the smaller, lower end monitors and kept the same class for the bigger, higher end monitors. In positive polarity, the PF

400 either kept the same class or it degraded the class. Because the standard explicitly states that measurements should not be performed on displays that have significant luminance fall off when viewed at oblique angles, it is difficult to determine the validity of these findings.

PF 450. Regardless of the monitor, this filter was always a Class II in positive polarity. In negative polarity, the PF 450 seemed to keep the same class as the monitor with the exception that it improved the SamSung and Nokia monitors. Because the standard explicitly states that measurements should not be performed on displays that have significant luminance fall off when viewed at oblique angles, it is difficult to determine the validity of these findings.

HF 300. In general, the HF 300 either kept the polarity Class the same as the monitor alone or it improved the polarity Class. It especially improved the SamSung negative polarity and the Nokia positive polarity classes.

AF 100 or AF 200. Regardless of the monitor, the AF 100 or AF 200 filter was always a Class II in positive polarity and a Class III in negative polarity.

General Findings

One general finding was that positive polarity always yielded either Class I or Class II because the Large Source, Screen Image Luminance Ratio ($LR_{BDS/FDS}$) and the Specular Reflection Luminance Ratio ($LR_{BDS/BD}$) always passed. Conversely, negative polarity always produced Class II or Class III because the Small Source, Specular Reflection Luminance Ratio ($LR_{BDS/BD}$) failed every time (note: it would occasionally pass when the calculations were performed for a Class III monitor). Additionally, the negative polarity, large source, specular reflection luminance ratio ($LR_{FDS/BDS}$) always passed.

Also, the AF 150 and HF 300 were the best filters for reducing glare on monitors. The BF 10 and AF 100 or AF 200 seem to be the worst because they can intensify screen reflections. Table 7 illustrates the effects of certain filters on ISO classification.

Table 7a. Tabulation of Filter Effects on Monitor Class in Positive Polarity

Monitor Class	AF 150	BF 10	PF 400	PF 450	HF 300	AF 100 or AF 200
I	0	-	-[1]	-	0[2]	-
II	+	0	0	0	[3]	0

Table 7b. Tabulation of Filter Effects on Monitor Class in Negative Polarity

Monitor Class	AF 150	BF 10	PF 400	PF 450	HF 300	AF 100 or AF 200
II	0	-	0	0	0	-[4]
III	+[5]	-[6]	+	0[7]	+	0
FAILED	+	0	++	+	++	+

Note: The symbos mean the following:

‘-’ means filter degraded one class
‘0’ means filter kept the same class as the monitor alone
‘+’ means filter improved one class
‘++’ means filter improved two classes

The superscripts mean the following:

[1] = filter degraded the class 3 times but kept the same class twice
[2] = filter kept the same class 3 times but degraded twice
[3] = filter improved the class once and kept the same class once
[4] = filter degraded twice and kept the same class once
[5] = filter improved the class twice and kept same class once
[6] = filter degraded the class twice and kept same class once
[7] = filter kept the same class twice and improved class once

PHASE 2

Methods

Participants

Fourteen individuals (seven male; average age of 24) participated in this experiment and were paid for their time. These participants were from the Virginia Polytechnic Institute and State University population. Participants had 20/20 or 20/20 corrected near vision and were screened using an Ortho-Rator test. Those participants who wore contact lenses and glasses were allowed to participate but had to meet the vision requirements. People who wore tinted lenses were not allowed to participate. Additionally, each participant spoke English as a native language.

Equipment

The experiment was conducted in the Displays and Controls Laboratory at Virginia Politechnic Institute and State University. A desktop computer (Model Apple Macintosh 8100/120) was used with three cathode ray tube (CRT) monitors: SamSung, Mitsubishi, and Sony Trinitron 100. The following filters were attached to some of the monitors: 3M Anti-Glare Filter Model BF10, 3M Circular Polarizing Filter Model HF 300, and 3M Anti-Glare Filter Model AF150. Three CRTs without an add-on filter and four CRT/add-on filter combinations were selected for use in this study:

TABLE 8. Monitors and filters Used in the Human Performance

Monitor	Positive Polarity	Negative Polarity
SamSung	Class II	Failed
SamSung/AF150	Class I	Class III
SamSung/HF300	Class II	Class II
Sony	Class I	Class III
Sony/BF10	Class II	Failed
Mitsubishi	Class I	Class II
Mitsubishi/AF150	Class I	Class II

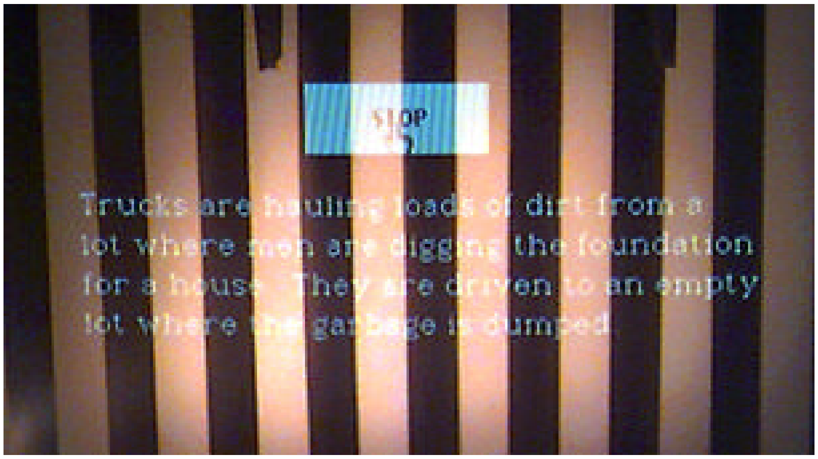
Experimental Conditions

The monitor/filter combination had seven levels. Three monitors and four monitor/filter combinations were selected for the human performance study. The monitors and monitor/filter combinations were chosen based on their ISO classifications. There is at least one display from each of the classes plus filters that improved, degraded, and kept the same Class as the display alone. Table 7 shows the monitor/filter combinations used in the study.

Ambient illumination consisted of five levels: one for a dark condition, three for diffuse light, and one for specular light. For the dark condition, the only light sources were the monitor under test and an incandescent lamp. The lamp had a 15 watt light bulb and had orange filters placed over it so that it did not interfere with dark adaptation. The lamp was necessary so that the participants could use the keyboard to type the out of context word in the dark.

For the diffuse light conditions, fluorescent light sources were positioned so that they did not create any specular (sharp, mirror like) images on the screen. Additionally, the participants sat behind a viewing booth so that their clothing did not create any specular images on the screen. The three levels of diffuse lighting were 200, 500, and 800 lux. The 200 lux condition was chosen since it was the same amount of light used in the diffuse measurements for the ISO standard. Additionally, the ANSI/HFS 100 standard (1988) indicates that office lighting levels should be between 200 and 500 lux. Thus, 500 lux was chosen because it was the upper range suggested by ANSI. Since most office environments are illuminated much brighter than 500 lux (Stammerjohn et al., 1981), 800 lux was chosen as the last level.

For the specular condition (Figure 6), a white and black vertical bar pattern was positioned to reflect in the display. The width of each bar was 19 millimeters. Two incandescent lights shined on the bar pattern so that the luminance emitted from the center white vertical bar was 2000 cd/m^2. This is equivalent to the lighting level used for the specular condition in the ISO standard. The specular bar pattern was positioned so that it overlapped the information that the subject needed to read from the screen.



Trucks are hauling loads of dirt from a lot where men are digging the foundation for a house. They are driven to an empty lot where the garbage is dumped.

Figure 6. White and black vertical bar pattern specular reflection used in the experiment

The last independent variable is the polarity of the screen and it has two levels: positive and negative. The positive polarity screen had dark characters on a bright background and the negative polarity screen had bright characters on a dark background.

The experimental dependent measure was a 9-point subjective rating scale (Beaton et al., 1985; Hunter, 1988).with adjectives (anchors) to assist observers in making their judgment (i.e., 1-worst imaginable, 5-passable, 9-best imaginable).

Several experimental conditions needed to be controlled in the experiment. These conditions included: setting the monitor’s contrast and brightness levels; setting the resolution and font; placing the chin rest; and measuring the lighting levels.

The contrast and brightness levels were set on each monitor by following a specific procedure. First, the brightness and contrast levels were turned completely down. Then, the brightness level was raised until the raster lines were just perceivable in a dark room. Finally, the contrast level was raised until it reached 75 cd/m^2.

The pixel addressability used was 832 by 624. This was the highest addressability that all three monitors could display. The font used for the passages was New York. Characters were set to same heights and same widths on each monitor. Characters subtended 0.67 degrees of visual angle. The chin rest was placed 40.64 cm (16 inches) from the screen.

The illuminometer was not used to measure the lighting levels during the experiment for a number of reasons. Using the illuminometer would lengthen the experiment time because it and the lighting controls would be separately located. Another reason is that the participants would have to relocate for each change in lighting level so that the experimenters could read the illuminometer from behind the booth. The illuminometer could not be read from in front of the booth because the experimenter would lower the actual lighting level by absorbing light. Therefore, a calibration procedure was necessary. The following equipment was used in the calibration procedure: a Minolta illuminometer, a 18% Kodak gray card, and a hand-held photometer (Minolta CS-100). The illuminometer was taped to the front of each monitor (i.e. lighting levels were set separately for each monitor). The gray card was placed on a wall next to the monitors. The photometer was aimed at the gray card and located behind the light chamber (viewing booth) next to the lighting controls. For each monitor and each lighting level, the corresponding luminance readings were made. For example, 200 lux on the Sony was equivalent to 26.4 cd/m^2 read from the photometer. Thus, during the experiment, the lighting levels were changed using only the hand held photometer rather than the illuminometer.

Figure 7. Viewing-booth/light-chamber used to control glare in experiment.

The last independent variable was the polarity of the screen and it had two levels: positive and negative. The positive polarity screen had dark characters on a bright background and the negative polarity screen had bright characters on a dark background.

Procedure

The participants were first asked to read and sign an informed consent form (Appendix B). Then, the participants' near field acuity was tested using the Ortho-Rater Device. After the participants passed the visual acuity test, they were seated in the viewing booth and the experimenter read the instructions to them. A copy of the instructions is listed in Appendix D. Next, the participants were allowed to practice reading passages from the Tinker reading task. The practice session allowed the participants to become familiar with the passages and the computer program. After completing 6 practice trials, the actual experiment began.

First, the participants were asked to sit in the dark for three minutes to allow them time to dark adapt. They were then asked to read one of the Tinker reading passages displayed on the screen in the dark. The participants were asked to rate the image quality of the screen. The participants repeated this process for all of the ambient lighting levels and polarities of the screen.

After the participants received all levels for one monitor, they were asked to sit in the dark for another three minutes. Meanwhile, the experimenter changed the monitor-filter combination. Once the next monitor was in place, the participants received all of the levels for lighting and polarity once again. This procedure was repeated until all of the monitor/filter combinations were given to the participants. Upon completion of the last level, the participants were paid and thanked for their time.

Experimental Design

The Phase 2 experiment was a human performance study. Specifically, it was a 7 x 5 x 2 factorial, within-subjects design. Independent variables were monitor/filter combination, ambient illumination, and polarity of the computer. The independent variables were blocked according to monitor/filter combination. In other words, subjects received all levels of ambient illumination and polarity of the screen for one monitor at a time. A Balanced Latin Square was used to order the way in which the participants received the monitor/filter combinations. The ambient illumination and polarity of the screen independent variables were randomized for each monitor with one exception. The dark ambient light condition was always given first to the observers so that they have time to dark adapt.

RESULTS

The image quality values were subjected to a three factor, within-subjects analysis of variance (ANOVA) procedure. The main effects of Monitor and Lighting were significant, as were the interactions of Monitor x Lighting, Lighting x Polarity, and Monitor x Lighting x Polarity. No other ANOVA effects were significant (Table 8).

Table 9. Analysis of Variance Subjective Rating

Source	**DF**	**SS**	**MS**	**F**	***p***
Between					
Subject	13	274.869	21.144		
Within					
Monitor	6	75.069	12.512	3.39	0.005
Monitor Subject*	78	288.045	3.693		
Lighting	4	2311.506	577.877	99.55	0.000
Lighting Subject*	52	301.865	5.805		
Polarity	1	19.433	19.433	3.44	0.086
Polarity Subject*	13	73.399	5.641		
Monitor*Lighting	24	52.237	2.177	1.98	0.005
Monitor*Lighting Subject*	312	343.792	1.102		
Monitor*Polarity	6	8.910	1.485	1.77	0.117
Monitor*Polarity Subject*	78	65.518	0.840		
Lighting*Polarity	4	21.894	5.473	4.65	0.003
Lighting*Polarity Subject*	52	61.192	1.177		
Monitor*Lighting*Polarity	24	13.763	0.573	1.50	0.63
Monitor*Lighting*Polarity*subject	312	118.951	0.381		
Total	**979**	**4030.384**			

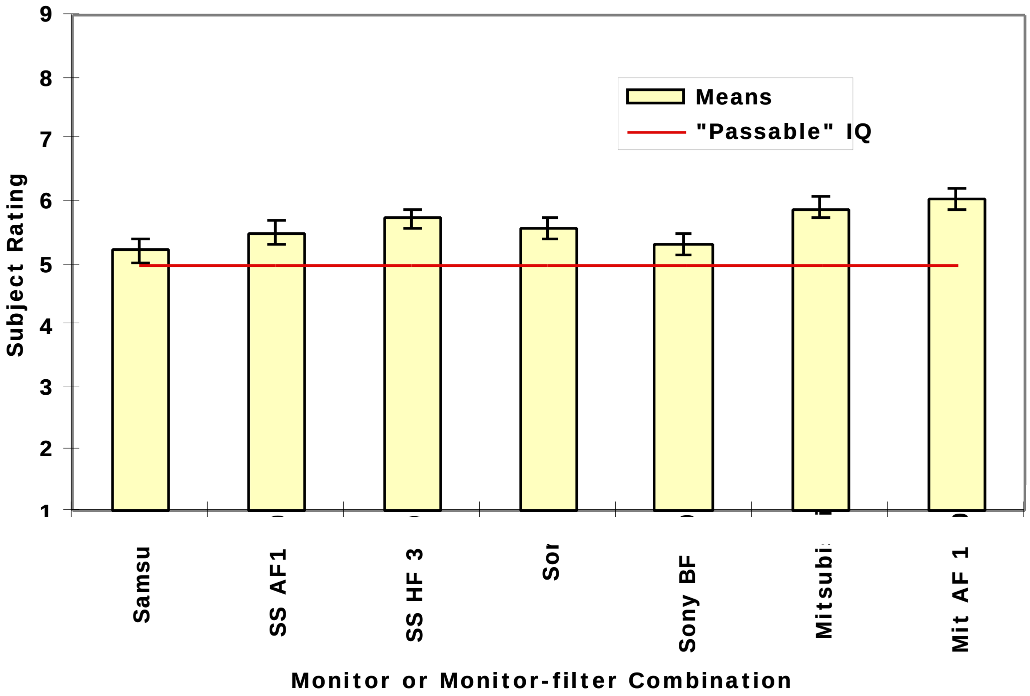


Figure 8. Main-effect of monitor-filter combination on subjective image quality judgments. Error bars show +/- 1 standard error of the mean.

Monitor

Figure 8 shows the main effect of Monitor on subjective image quality judgments. A Newman-Keuls *post-hoc* test (Table 10) indicated that the SamSung monitor was rated significantly lower in image quality than the Mitsubishi and Mitsubishi/AF 150. Moreover, the *post-hoc* test indicated that the Sony monitor was degraded in image quality with the addition of the BF10 anti-reflection filter. Finally, the *post-hoc* test indicated that the Mitsubishi monitor (with or without the add-on glare filter) was rated highest in overall image quality. Figure 8 also shows the mid-scale value for the subjective image quality judgments, which corresponded to the descriptive label of “passable.” It is apparent from the data trends that all monitors were rated as at least passable in image quality and that no monitor condition was rated higher than ok. This indicates that across all conditions of lighting in polarity, the test conditions represented acceptable levels of quality.

Table 10. Newman-Keuls Results for the Main Effect of Lighting on Subjective Ratings

(Note: Means with the same letter are not significantly different.)

Monitor	Mean	N	SNK Grouping		
Mitsubishi/AF 150	6.0429	140		A	
Mitsubishi	5.8929	140	B	A	
SamSung/HF300	5.7071	140	B	A	C
Sony	5.5714	140	B	A	C
SamSung/AF 150	5.4857	140	B	A	C
Sony/BF 10	5.3071	140	B		C
SamSung	5.2214	140			C

.

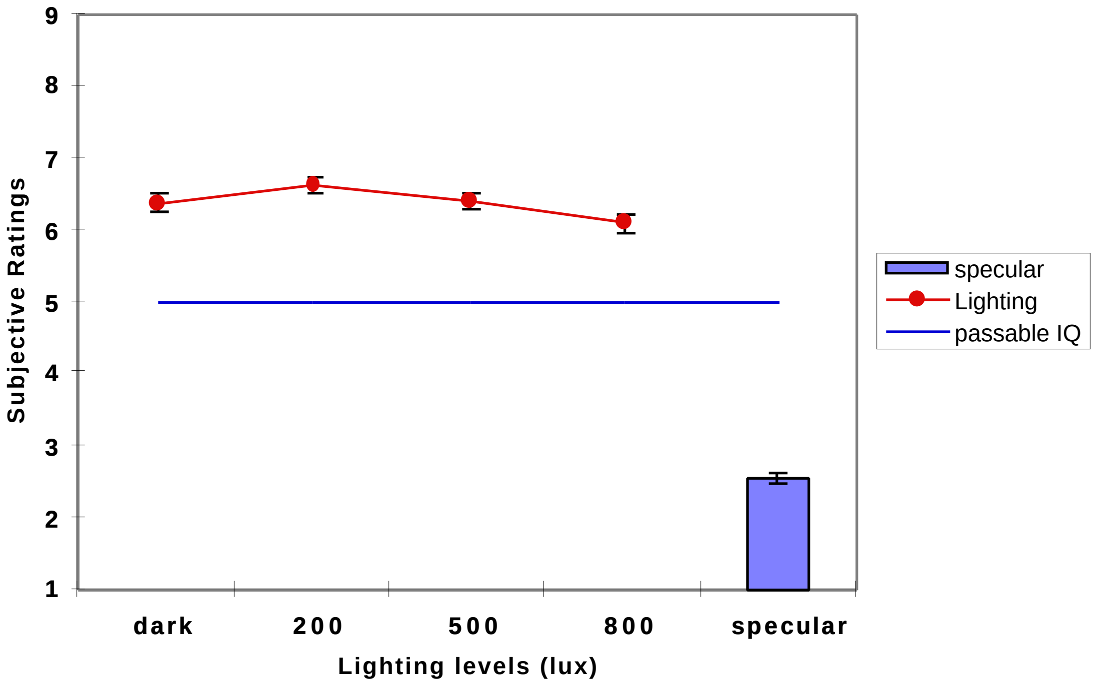


Figure 9. Main-effect of ambient lighting on subjective image quality judgments. Error bars show +/- 1 standard error of the mean.

Lighting

Figure 9 shows the main effect of lighting levels on subjective image quality judgments. This graph combines the dark, diffuse, and specular ambient illumination conditions. A Newman-Keuls *post hoc* analysis is shown in Table 11. The *post-hoc* test indicated that the image quality judgments for the specular condition were significantly lower than for the dark and bright diffuse conditions ($p < 0.05$). This data trend indicates that subjective judgments of image quality were not influenced by levels of diffuse glare. Image quality judgments were determined by the presence of specular glare instead. Given the large range of conditions assessed in this experiment, these data suggest that the presence or absence of specular glare was a primary determinant for the image quality judgments. In other words, these data suggests that observers rated image quality as low whenever specular glare was present. Observers did, however, rate image quality as at least passable or ok in the other ambient glare conditions.

Table 11. Newman-Keuls Results for the Main Effect of Lighting on Subjective Rating (Note: Means with the same letter are not significantly different.)

LIGHTING	Mean	N	SNK Grouping	
200	6.6173	196	A	
500	6.3929	196	A	
Dark	6.3724	196	A	
800	6.0867	196	A	
Specular	2.5510	196		B

Interactions

The significant ANOVA interactions follow main effect data trends. Two two-way and one three-way interaction were significant: monitor x lighting, lighting x polarity and monitor x lighting x polarity.

Monitor x Lighting

Figure 10 shows the two-factor interaction between monitor-filter combination and ambient lighting. These data do not change the conclusions reached by inspection of the main effects in Figures 8 and 9. It is apparent from Figure 10 that the image quality judgments for each monitor remained relatively constant across the dark and diffuse ambient conditions, with the notable exceptions of the SamSung monitor and the Sony/F10 monitor-filter conditions. More importantly, Figure 7 illustrates the strong effect of specular reflections on image quality judgments, as observed in the main effect of lighting level. Additionally, it is noted that image quality judgments for all dark and diffuse glare conditions exceeded the passable level; whereas, image quality judgments for all of

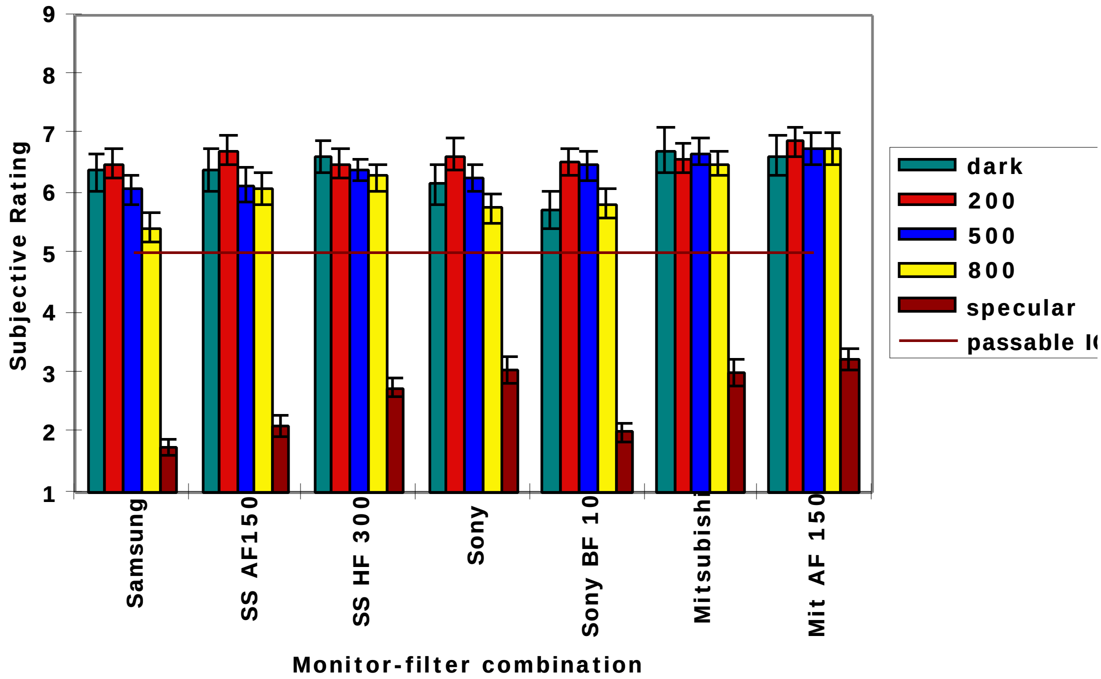


Figure 10. Interaction effect between monitor and ambient lighting on subjective image quality judgments. Error bars show +/- 1 standard error of the mean.

the specular glare conditions were located at low image quality values.

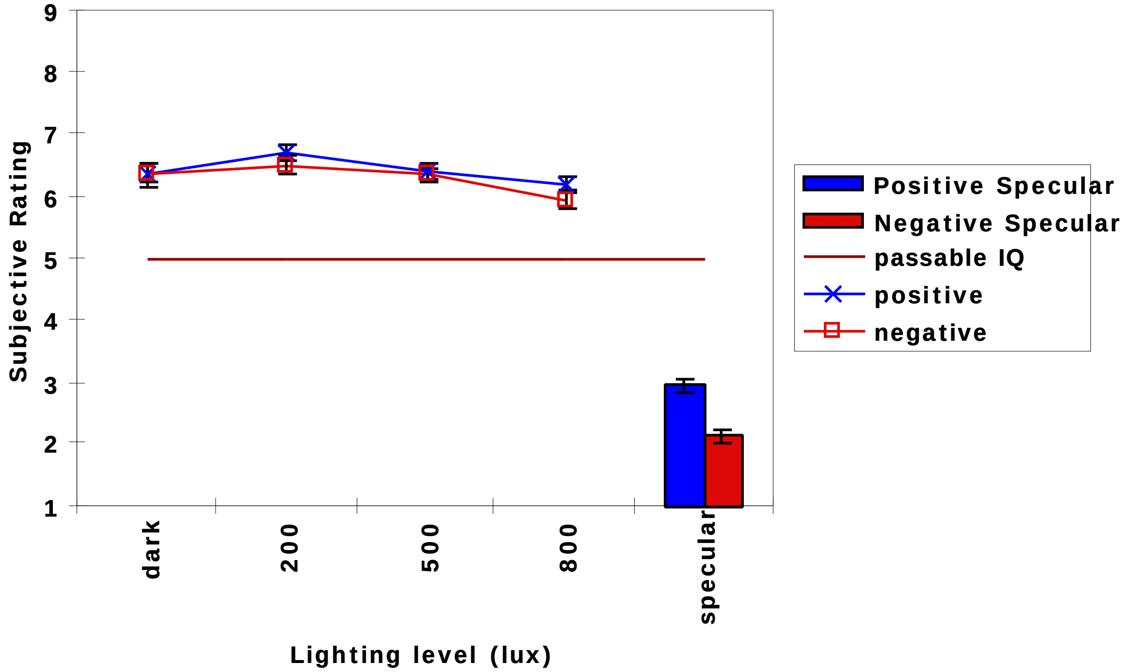


Figure 11. Interaction effect between ambient lighting and screen contrast polarity on subjective image quality judgments. Error bars show +/- 1 standard error of the mean.

Lighting x Polarity

Figure 11 shows the two-factor interaction between lighting levels and screen contrast polarity. Newman-Keuls post-hoc test indicated that screen polarity did not effect image quality judgments under the dark and diffuse ambient glare. Although the main effect of screen polarity was not significant, it is apparent that screen polarity effects image quality judgments when considered in the context of particular types of ambient light. Specifically, Figure 11 shows that image quality judgments for negative contrast (i.e., bright text on dark background) were significantly lower under the specular glare condition than were the judgments for positive polarity (i.e., dark text on bright background) under specular conditions.

Figure 12 shows the significant three-factor interaction among monitor, ambient lighting levels, and display polarity. Inspection of this graph indicates that the data trends observed in the lower order effects (discussed above) are not modified in any substantial manner. Specifically, the effects of ambient lighting, notably the effect of specular reflections, are apparent in the three-factor interaction. And the effects of monitor type are consistent across both positive and negative polarity conditions. Therefore, no further interpretation is necessary for this interaction.

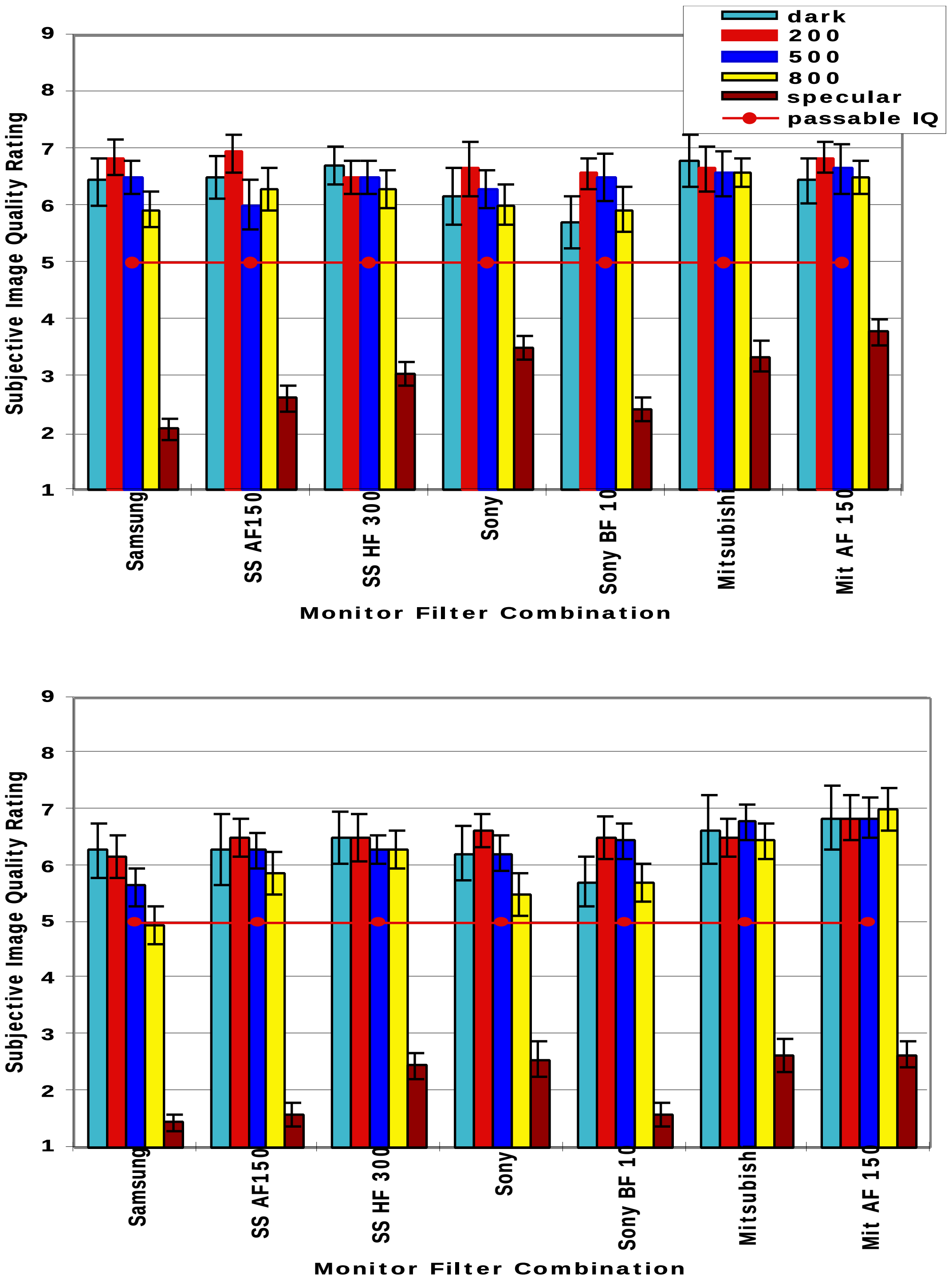


Figure 12. Three-factor interaction among monitor, ambient lighting, and contrast polarity on subjective image quality judgments. Top panel show data for positive polarity screen; bottom panel shows data for negative polarity.

Summary of ANOVA Results

The results of the ANOVA indicate that there were differences among the monitors, ambient lighting conditions, and screen polarity conditions examined. Although the three-factor interaction was significant, inspection of the lower-order effects indicates relatively clear data trends. Specifically, the SamSung monitor was rated lower in quality than the other monitors that included anti-reflection filters. Similarly, positive polarity screens were rated higher (approximately 4 percent) than negative polarity screens. And specular glare substantially degraded image quality judgments in comparison to the diffuse ambient glare levels. Inspection of Figure 12 clearly shows the magnitude of these main effect data trends. It should be noted that the subjective judgments were performed in the context of a large number of glare conditions, and that the observed data trends suggest that the presence of specular glare determined the extent of degradation rated by the observers. That is, the presence of specular glare appeared to determine low ratings of image quality, whereas all other glare levels were rated as at least passable in image quality.

Correlational Analysis

ISO Metrics. Subjective scale values were evaluated against two ISO metrics: *screen image luminance ratio* (Diffuse, 200 lux) and *specular reflection image luminance ratio* by correlational techniquesSpecifically, the subjective ratings were compared with ISO metrics obtained for Monitor under each of the polarity test conditions. Pearson Product-Moment Correlation coefficients are reported in Table 11.

Table 12. Correlations Between Subjective Image Quality Judgments and the ISO 9241-7 Image Quality Metrics

	Screen Image Luminance Ratio	**Specular Reflection Luminance Ratio**
	r , p	r , p
Positive Polarity, Large Source	-0.180, p=0.700	-0.833, p=0.020
Positive Polarity, Small Source	0.143, p=0.760	-0.825, p=0.022
Negative Polarity, Large Source	0.700, p=0.080	-0.767, p=0.044
Negative Polarity, Small Source	0.747, p=0.054	-0.690, p=0.086

The correlation coefficients show that the *specular reflection ratio* correlates substantially more with subjective ratings than the *screen image luminance ratio* does. The data suggest that subjective ratings do not significantly decrease as screen contrast is reduced (i.e., *screen image luminance ratio* is reduced). However, the data indicate that the *specular reflection luminance ratio* did index subjective quality ratings.

Class Correlation . In addition to the luminance ratio correlation, subjective judgments also were evaluated against ISO classifications. Specifically, subjective ratings in positive and negative polarity also were correlated with positive and negative polarity ISO classifications. A spearman rank correlation was used. Reading times in positive polarity had a correlation coefficient of -0.5774 (p=0.175) and in negative polarity the correlation coefficient was -0.9449 (p=0.001).

These results indicate that subjective ratings were not siginifcantly higher for a Class I monitor than a Class II in positive polarity. Subjective judgments did, however, systematically decrease as the Class of the monitor increased to II and III in negative polarity. Therefore, subjective ratings times appear to be affected by the Class of monitor/filter combination used. The ANOVA results also support this conclusion, because the three-way interaction showed the strong effect of polarity.

The results from both correlation analyses suggest that participants paid less attention to the contrast of text on the screen, reduced by diffuse glare, than to the presence of specular glare. In fact, the ANOVA results for the Monitor x Lighting interaction demonstrate this effect. The trend for lighting levels across all displays is not consistent because the subjective ratings are effected about the same for the no glare condition (dark) as all three of the diffuse glare conditions (200, 500, and 800 lux). The author contends that this finding is the result of an experimental context effect. In other words, specular glare was so influential on subjective ratings of image quality that subjects paid very little attention to reductions in contrast from the diffuse lighting conditions.

DISCUSSION

Psychophysics Glare Research

Perception of displayed information can be considered a direct measure of how objectionable display images are perceived in the presence of glare. A number of studies have investigated this issue. Beaton et al. (1985) and Hunter (1988) represent two directly relevant studies to the present study. Both studies used a perceived image quality rating scale: phychophysical measurement technique

Three findings of Beaton, Murch, and Knox (1985) study had far-reaching implications for this thesis. First, glare filter investigations must have a dark or low ambient lighting condition for a baseline. Secondly, both diffuse and specular glare sources must be used to test whether the filter can both enhance contrast and reduce specular (patterned) reflections. Third, the validity of the human performance measures used was found to be high. In other words, VDT operators can discriminate and detect even small differences between lighting filter conditions regardless of the psychophysical data collection procedure employed.

Hunter’s dissertations also was directly relevant to the design of this research. One objective of his study was to investigate the effects that glare filters with different surface treatments and transmissivities have on subjective image quality ratings. Another objective was to quantitatively model the relationship between measured image quality of display/filter combinations and recorded human performance data (perceived image quality data). His models were based on signal to noise measurements similar to the SNRs used by Beaton and Snyder (1984). The most important finding of this study for this thesis was that good fit models for perceived image quality and physical image quality (e.g. RTF metric) were achieved. This finding underlies the rational for this thesis: using perceived image quality ratings to establish a human factors basis to justify measurement requirements in ISO 9241-7.

Beaton et al.(1985) and Hunter (1988) both agree that in order for a CRT anti-reflection treatment or glare reduction method to be effective it must do two things. It must first enhance the display’s contrast by reducing diffuse reflections from the phosphor surface. Additionally, it must reduce the intensity and it must attenuate sharpness of specular reflections from the glass face-plate of the CRT. The findings of this research fall in line with Beaton et al. (1985) and Hunter (1988) in terms of the relative importance of these three criteria. Specifically, the findings of this study indicate that specular glare significantly degrades image quality ratings more than diffuse glare.

ISO Justification

As mentioned in the introduction, there are two main problems with the ISO standard that this thesis set out to investigate. First, the ISO 9241-7 does not address an important component of specular reflections: reflected image structure or sharpness. This is unfortunate because researchers have noted that sharply defined images are perceived as more annoying and noticeable than a less sharply defined images of the same intensity (Rancourt and Grenewalt, 1986). The findings of this thesis indicate that specular glare significantly degrade image quality ratings more than diffuse glare, and therefore, should be weighted more strongly in the determination of ISO classifications.

The second problem with the ISO standard is that it is, at best, only loosely based on human performance data. It is entirely unclear how the standard developers arrived at the specific luminance ratio requirements. Therefore, the question of whether the three ISO classifications truly represent perceivable display performance differences, was addressed.

This thesis establishes a human factors basis to justify the measurement requirements in the ISO 9241-7 standard. Specifically, the findings show that it underemphasizes the specular reflection luminance ratios contribution and overemphasizes the screen image luminance ratios contribution to compliance classification calculations, because the procedure gives equal weight to both ratios. The author contends that this finding is the result of an experimental context effect. In other words, the specular glare was so influential on subjective ratings of image quality that subjects paid little or no attention to reductions in contrast from the diffuse lighting conditions.

REFERENCES

Beaton, R. J. (1993). *Optical specifications for anti-glare filters*. Blacksburg, VA: Displays and Controls Laboratory, Virginia Polytechnic Institute and State University.

Beaton, R. J., Murch, G. M., and Knox, S. T. (1985). *On the selection of glare filters for CRT-based workstation: Part I. Human performance studies* (Tech. Report HFL-604-02). Beaverton, OR: Human Factors Research Laboratory, Tektronix, Inc.

Beaton, R. J. and Snyder, H. L. (1984). The display quality of glare filters for CRT terminals. *In Society for Information Display 1984 Digest of Technical Papers* (pp. 298-301). Blacksburg, VA.

Cakir, A., Hart, D., and Stewart, T. (1980). *Visual display terminals: A manual covering ergonomics, workplace design, health and safety, task organization*. New York: Wiley. Displays International Symposium Digest of Technical Papers (pp. 420-422).

Cone, J. (1997). Email correspondence. Regarding ISO 9241-7 Specular Source Rational.

Sanders, M.S. and McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). New York: McGraw-Hill Book Company.

Garcia, K. D., and Wierwille, W. W. (1995). The effect of glare on performance of a VDT reading-comprehension task. *Human Factors*, 27(2), 163-173.

Grandjean, E. (1987). *Ergonomics in computerized offices*. London: Taylor & Francis.

Habinek, J. K., Jacobsen, P. M., Miller, W., and Suther, T. W. (1982). A comparison of VDT anti-reflection treatments. *In Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 285-289). Santa Monica, CA: Human Factors Society.

Human Factors and Ergonomics Society, Inc. (1988). *American national standard for human factors engineering of visual display terminal workstations*. (ANSI/HFES 100-1988). Santa Monica, CA: Author.

Hunter, M. W. (1988*). CRT anti-glare treatments, image quality, and human performance*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg,VA.

Illuminating Engineering Society (1987). VDT office lighting (IES RP-24-1987). New York.

International Organization for Standardization (1997). Ergonomic requirements for office work with visual display terminals (VDTs) Part 7: Display requirements with Reflections

Isensee, S. H. and Bennett, C. A. (1983). The perception of flicker and glare on computer CRT displays. *Human Factors*, *25* (2), 177-184.

Kubota, S. (1994) *Effects of surface reflection properties of VDT screens on subjective ratings of the reflected glare disturbances*. Unpublished study contained in Japanese comments on ISO/CD 9241-7.3, ISO/TC159/SC4 N-289.

Laubli, T., Hunting, W., and Grandjean, E. (1980). *Visual impairments in VDU operators related to environmental conditions*. In E. Grandjean and E. Vilgliani (Eds.), Ergonomic aspects of visual display terminals. London: Taylor & Francis, (pp. 85-94).

Lloyd, C. J., Mizukami, M., and Boyce, P. R. (1996). A preliminary model of lighting-display interaction. *Journal of the Illuminating Engineering Society*, 59-69.

Morse, R. S. (1985). Glare filter preference: influence of subjective and objective indices of glare, sharpness, brightness, contrast and color. *In Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 782-786). Santa Monica, CA: Human Factors Society.

Rancourt, J. and Grenawalt, W. (1986). Approaches to enhancing VDT viewability and methods of assessing the improvements. Advances in Display Technology VI: Proceedings of SPIE - The International Society for Optical Engineering, 624, 8-13.

Rancourt, J., Grenewalt, W., Hunter, M. W., and Snyder, H. L. (1986). Quantitative evaluation of the effect of an anti-reflection filter. *In Society for Information Displays International Symposium Digest of Technical Papers* (pp. 420-422).

Stammerjohn, L. W., Smith, M. J., and Cohen, B. G. F. (1986). Evaluation of workstation design factors in VDT operations. *Human Factors*, 23(4), 401-412.

Stammerjohn, L. W., Smith, M. J., and Cohen, B. G. G. (1981). Evaluation of workstation design factors in VDT operations. *Human Factors*, 23(4), 401-412.

Stevenson, J. R., Fendley, J. R., and Wallner, J. (1986). Optical characterization of anti-glare surfaces on cathode ray tubes. *In Society for Information Display International Symposium Digest of Technical Papers* (pp. 424-427).

APPENDIX A

ISO Measuraement Procedure Observations

Diffuse measurements

- The ISO standard does not specify which optical surface to focus the photometer for the diffuse measurements: the glass face plate or the phosphor surface. The author of this paper focused on the phosphor surface because that is where the diffuse reflections act to reduce contrast.

Small-Source Specular Measurements @ 15 degrees

- For these measurements, the standard states that the reflections from the glass surface and the reflections from the phosphor layer should be superimposed. Additionally, the photometer should be focused on the virtual image. Unfortunately, the reflections could not be superimposed in this study. In order to superimpose the reflections, the display needed to be rotated out of position: 15 degrees from normal. The author believes that this is a flaw in the design of the standard because this measurement is nearly impossible to make. Thus, the author decided to take one scan of the phosphor reflection and then add it to a scan of the glass surface reflection, as these luminance values are additive quantities.
- When a filter is placed over the monitor, it creates a third reflection that should be superimposed with the two from the display. Again, at a 15 degree angle, these images could not be superimposed. However, a concerted effort was made to superimpose the reflections as best as possible. The author, thus, decided to change the angle of the filter slightly so that these two images would superimpose. Therefore, only two scans were required for these measurements: one for the phosphor layer of the screen and one for the superimposed first surface reflections.

Extended-Source Specular Measurements at 15 degrees

- Unlike the small source measurements, the ISO standard does not specify whether the photometer should be focused on the virtual image or the phosphor layer, nor does it state that the images should be superimposed for the extended-source measurements. The author did, however, focus on the virtual image. In this case, the reflections from the phosphor layer and the first surface did superimpose because of the aperture of the sphere was so large. Thus, only one scan was needed.
- Adding a filter to a display introduced yet another optical surface reflection which need to be superimposed for these measurements. Thus, the filter was moved until the images superimposed, but again the problem was that the filter was not rotated at a 15 degree angle.

Points of Confusion

Determining Class

- The notion that a monitor can have two separate and different classifications, one for positive polarity and one for negative polarity ,is confusing. For example, a monitor can be Class I for positive polarity and Class II for negative polarity. This means that the monitor is suitable for general office use in positive polarity but only suitable for "most" office environments in negative polarity. If the monitor only had one classification, then the standard would be much easier to understand and interpret. Furthermore, the author never found a monitor that had the same Class for both positive and negative polarity. However, some monitor/filter combinations kept the same Class in both polarities.
- The requirements to determine classification seem arbitrary and confounding. The author has already described the requirements in the section titled "Methods of Physically Measuring Glare" of this report. Table 12 lists some examples of which luminance ratios need to pass to be a Class I, II, or III. Examples 3 and 5 clearly show the confusion. The standard can't say if only one or two ratios fail, then it is a Class II, because classification depends on which ratios fail.
- Table 12 also shows that knowing a monitor is Class II in a polarity does not give much information about the monitor. For example, one monitor could be Class II in positive polarity because it failed the Screen Image Luminance Ratio for a small source, but another monitor may fail at both the Screen Image and the Specular Reflection Luminance Ratios for the small source.

Table 13. Examples of Luminance Ratios that need to pass for Class I, II, or III

	Example 1	**Example 2**	**Example 3**	**Example 4**	**Example 5**	**Example 6**
Large Source						
Screen Image Luminance Ratio	pass	pass	pass	pass	pass	pass
Specular Reflection Luminance Ratio	pass	pass	pass	fail	fail	fail
Small Source						
Screen Image Luminance Ratio	pass	pass	fail	pass	pass	fail
Specular Reflection Luminance Ratio	pass	fail	fail	pass	fail	fail
	Class I	Class II	Class II	Class II	Need to do calculation to see if a Class III	Need to do calculation to see if a Class III

Calculating the Equations

- When calculating the reflectivity equations and the luminance ratio equations, it is difficult to understand what numbers should be used for the L_A. For example, the reflectivity equations use the luminance level used for the integrating spheres (i.e. 2000 cd/m2). However, the luminance ratio equations use the luminance levels found in Table 1 of this report. These are the same luminance levels that are used to determine classification.
- Furthermore, the standard does not explain why these luminance levels were chosen and intuitively they do not make much sense. For example, the screen image luminance ratio equation is shown below.

$$\text{Screen Image Luminance Ratio} = \frac{L_{\text{background}} + L_{\text{diffuse}} + L_{\text{specular}}}{L_{\text{foreground}} + L_{\text{diffuse}} + L_{\text{specular}}} \geq 3 \qquad \text{(positive polarity)}$$

The L_{specular} component is calculated by the formula $L_{\text{specular}} = L_{A(\text{Table 1})}$ * Specular Reflectivity. For the large source, the L_A would be 200 cd/m^2 but for the small source the L_A would be 2000 cd/m^2. The standard does not clarify why the different sources use the different lighting levels in the equations. Remember that the large and small aperture integrating spheres are set to the same level (2000 cd/m^2) in the measurement procedure.

ISO 9241-7 Recommendations

The first time a person tries to classify a monitor they notice an exorbitant number of measurements and subsequent calculations are needed. Specifically, there are ten photometric measurements to take and 12 algebraic equations to calculate to determine the class of a monitor. Two sets of measurements that appear to be redundant are the two small source specular ($L_{DS(SML,15)}$) for negative/positive polarity and the two large source specular ($L_{DS(EXT,15)}$) for positive/negative polarity. The only difference between these two measurements is the exit port size of the specular light source used (1 degree for small-area source and 15 degrees for large-area source at 500 mm). “Optically, the large-area and small-area reflection characteristics will be equivalent for specular reflections” (Cone, 1997). In other words, measured reflected luminances values will be the same for these two sets of measurements. When John Cone, one of the creators of the document, was contacted he described the rational behind the small source specular measurements.

Cone claims that the main reason for the small-area source measurement is that the large-area source measurement is not “fair” to mechanically or chemically etched (surface-textured) screens. In fact, “these treatments do not, in general, actually reduce the total number of quanta reflected... they may actually add some diffuse reflection to the first optical surface" (Cone, 1997). The ISO standard should not lower its requirements just so mechanically and chemically etched screens can pass the standard. In fact, researchers have noted that mechanically etched or chemically abraded screens can actually reduce the quality of the displayed information, regardless of ambient lighting conditions. Beaton and Snyder found that average MTF values for an etched CRT were lower than average MTF values for a polished CRT (1984). Therefore, the author believes that these two small source measurements should be eliminated from the standard.

A more fundamental problem with the specular measurements of the standard is they do not address an important component of specular reflections: the reflected image sharpness or structure. This is unfortunate because researchers have noted that a more sharply defined or structured image is perceived as more annoying and noticeable than a less structured image of the same intensity (Rancourt and Grenewalt, 1986). Moreover, specular reflections are annoying and distracting to the operator because the eye will have difficulty focusing on the display images, since the characters on the screen and the reflections are in different focal lengths. When the eye is focused on the characters, the reflected image is blurred and out of focus. Because the eye naturally wants to bring the blurred image into focus, it will shift its attention to the specular reflection: thereby blurring the character images. Therefore, the eye's focus is constantly oscillating between screen characters and the reflected image (Cakir et al 1980). This oversight may be due to the inability of the investigators and researchers in the field to agree upon a single quantitative physical measurement method for characterizing display specular reflectance. The RTF, however, has shown promise in this arena. Hunter (1988) found that good fit models were produced for perceived image quality and the RTF physical image quality metric. Therefore, the author recommends that the standard committee considers implementing the RTF metric instead of the four specular measurements outlined in the standard. The benefits of implementing the RTF would not be restricted to an accurate indexing of perceived image quality, but would also provide substantial time savings (one measurement as opposed to four).

Additionally, affixing an after-market stand-alone filter to a monitor will increase the number of specular images. This is due to light rays reflecting from the first surface of the monitor and the first surface of the filter. Additionally, these reflections will be at different focal lengths because there is a gap between the front surface of the monitor and the front surface of the filter. The additional specular image presents yet another image on the screen that the eye will naturally try to pull into focus: compete with the character images on the screen for attention. This problem has a significant implication for the ISO standard. First, the ISO standard requires that all images be superimposed so it does not account for this problem. Furthermore, the limited human performance data from which the ISO standard is based did not test monitors coupled with stand alone filters (Kubota 1994. This increase in the number of specular reflections represents another reason why investigators and researchers in the field should work together to develop an improved quantitative physical measurement method for characterizing display specular reflectance.

APPENDIX B

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Informed Consent for Participants of Investigative Projects

Title of Project: An Assessment of the Validity of the International Standard, ISO 9241-7

Investigators: Gary Olacsi, Industrial and Systems Engineering Graduate Student
Joy Kempic, Industrial and Systems Engineering Graduate Student
Dr. Robert Beaton, Industrial and Systems Engineering Professor and Director of the Displays and Controls Laboratory

I. The Purpose of this Research

The purpose of this research is to critically assess the validity of the international standard, ISO 9241-7. This standard specifies ergonomic display requirements with reflections for office work with VDTs. In particular, it measures display hardware design attributes that produce screen reflections and stipulates the type of office ambient lighting environment for which a display is suitable.

Although the authors of this standard claim to have based it upon human performance data, it is entirely unclear how they arrived at specific ergonomic display requirements. Therefore, this experiment will collect subjective and objective human performance data and correlate it with ISO 9241-7 measurement results in order to determine whether the standard truly captures perceivable display performance differences.

II. Procedures

After you have passed the visual acuity tests using the Ortho-Rator device, you will be seated in the viewing booth and asked to place your head on a chin rest. Next, you will be allowed to practice reading passages from the Tinker reading task. The Tinker Speed of Reading Test was developed by Miles A. Tinker and it consists of 500 paragraphs (Feldt 1955). All of the paragraphs are kept at the same comprehension level so that accuracy and speed are the only factors under test. Each paragraph consists of about 3 sentences. Within each paragraph, one word is out of context and you will be asked to identify this word. The word is usually located in the last sentence of each paragraph. The following is an example of a Tinker reading passage. The word out of context is ‘car’.

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

Upon completion of 6 practice trials, you will be asked to sit in the dark for three minutes to allow time for your eyes to dark adapt. You will then be asked to read one of the Tinker reading passages displayed on the screen in the dark. You will be asked to type the word out of context and your reading speed will be timed. You will also be asked to rate the image quality of the text passage on a nine point scale ranging from worst imaginable to best imaginable. You will repeat this process for different ambient lighting levels and different polarities of the screen. (Note: Polarities can be of two types: positive or negative. Positive polarity refers to dark characters on a light background and negative polarity refers to light characters on a dark background.)

After you receive all levels for one monitor, you will be asked to sit in the dark for another three minutes. Meanwhile, the experimenters will change the monitor or monitor-filter combination. Once the monitor is in place, you will receive all of the levels for lighting and polarity once again. This procedure will be repeated until you have received all of the monitors or monitor filter combinations. Upon completion of the last level, you will be paid and thanked for your time.
The total time commitment is one session with a maximum length of 2 and 1/2 hours.

III. Risks

There is minimal risk involved in participating in this experiment. In simplified terms, minimal risk means that the risks involved in participating are no greater than those encountered in everyday life.

IV. Benefits of this Research

You will receive no benefits from participating in this experiment, other than whatever satisfaction you may derive from contributing to the knowledge of glare on computer screens. No promise or guarantee of benefits has been made to encourage you to participate.

V. Extent of Anonymity and Confidentiality

The results of this study will be kept strictly confidential. The information that you provide will have your name removed and only a participant number will identify you during analysis of the research.

VI. Compensation

Upon completing the experiment, you will receive five dollars per hour for a maximum of $12.50 (for 2 1/2 hours) and a $25 gift certificate to a local restaurant for your participation.

VII. Freedom to Withdraw

You are free to withdraw from this study at any time without penalty. If you choose to withdraw, then you will be compensated for the portion of the time you participated in the study.

VIII. Approval of Research

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Participants at Virginia Polytechnic Institute and State University.

IX. Participant's Responsibilities

I voluntarily agree to participate in this study. I have the following responsibilities:

1.) I should not volunteer for participation in this research if I am younger than 18 years of age, wear tinted lenses, or can not speak English as a Native language

2.) I will not discuss this experiment with anyone for at least 30 days after participating in the experiment.

X. Participant's Permission

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in the project.

If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

________________________________ ________________________________

Signature Date

Should I have any questions about this research or its conduct, I may contact:

Gary Olacsi Investigator	552-3470
Joy Kempic Investigator	951-4276
Dr. Robert Beaton Faculty Advisor	231-5986
H.T. Hurd Chair, IRB Research Division	231-5281

APPENDIX C

Subjective Image Quality Scale

** Please rate the image quality of the Computer screen.

Worst Imaginable	Awful	Poor	Marginal	Passable	OK	Good	Excellent	Best Imaginable
1	2	3	4	5	6	7	8	9

APPENDIX D

INSTRUCTIONS

Task Overview

For this experiment, there are two tasks which you will be asked to perform. The first task requires that you read text passages which are 4 lines in length and contain one out of context word. The out of context word will usually be found in the last line of these text passages. The second task of this experiment requires that you rate the overall image quality of the display using a nine point subjective image quality scale. Prior to the experiment, you will be asked to complete a demo of the computer program. This demo will allow you to gain familiarity with performing these tasks. The demo will also expose you to the range of ambient lighting conditions which will influence your opinion of the screens' subjective image quality.

Please note: There will be times during the experiment that we will ask you to sit in the dark for 3 minutes in order to allow your eyes a chance to adapt to the dark.

Examples

It is true that "fingers were made before
forks," but forks were invented many years
ago and today there is no excuse for our not
knowing how to use shovels.

Margaret's father gave her a large box of
writing paper for her birthday because she
has a great many friends in every state to
whom she frequently sings.

Procedure

You will see 3 main screens during the experiment. The first screen asks if you are ready to read a passage. The next screen will be the passage. The last screen will ask you to rate the image quality.

For the third screen of the , you will be asked to rate the general image quality of the entire screen. During the demo, the experimenters will show you examples of the extremes of the image quality will be encountering to aid you in using this subjective image quality scale. Your accurate opinion of the image quality of each trial and your ability to use all the categories of the scale are essential to the success of this experiment.

APPENDIX E

X-Sender: bobb@mail.vt.edu
Mime-Version: 1.0
Date: Tue, 26 Aug 1997 18:39:30 -0400
To: golacsi (Gary Olacsi), jkempic (Joy Kempic) From: john cone
<jcone@dante.nmsu.edu> (by way of Bob Beaton) Subject: Re: Are you still there?

Hi, Gary & Joy.

I received the following from John Cone today.

Thx,
Bob

Bob -

The root of the issue was a desire to make part 7 a performance standard rather than a product standard. Actual methods of controlling specular reflections (at the time the stanard was under development) were pretty much surface etch (mechanical and chemical), surface silica coating, and single-and multi-layer antireflection coating. On flat panels, molded texture on the cover plastic was (and is) standard. Most of the surface textures can be combined with simple anti-reflection coating for extremely effective reflection-control. The Europeans (especially Swedes and Germans) wanted to make multilayer coating a requirement. Other countries did not want to exclude surface texturing, as it had (and has) a much lower production cost. And, of course, under controlled environmental conditions, no treatment might be required.

The solution was to propose some image quality measures, to be met under specified surrogates of reasonably valid workplace conditions. The requirement could be met with any combination of high image luminance (my favorite), and/or one or more means of diffuse and direct reflection control.

The next problem arises in the "fair" measurement of reflections from screens treated to diffuse the direct reflection. These treatments do not, in general, actually reduce the total number of quanta reflected. In fact, they may actually increase the total, by adding some diffuse reflection to the first optical surface.

The problem is complicated by the existence of multiple optical surfaces (up to 7 in an LCD) and by curved, non-parallel reflecting surfaces (in CRTs). In the case of specular surfaces, the solution is simple: Use a large luminous source. Then all the multiple reflections can be superimposed over most of the reflected virtual image of the source, and measurement can proceed routinely.

But a large source penalizes the surface-textured screens, and overstates their reflectance for small sources. So we were back to institutionalizing multilayer coatings.

Surface texture can be thought of as applying a point-spread function to the first reflective surface (and the second, in the case of a CRT: The rear glass surface is usually "orange-peeled."). The total amount of light reflected stays about the same, or may actually increase with an increase in diffuse reflectance. However, from a point-source, the intensity of reflection is reduced in any given direction.

The actual luminous intensity of the reflected ray(s) in one direction is the sum (actually, the convolution) of the point spread function(s) and the spatial distribution of luminances about the specular angle - plus any purely diffuse reflected components.

I say point spread functions, rather than function, because - At the first optical surface, some methods of surface texturing, such as chemical etching, often leave some of the original smooth glass surface intact.
- There are usually subsequent optical surfaces or interfaces that also can support specular reflection, and whose reflections are diffused, if at all, by transmission through the surface texture. This will impose a different point-spread function.
- Surface textures are often combined with other coatings, for anti-reflection, anti-static, emission control, etc.

The optical effects of each of these factors is usually dependent on the angle of incidence, and is not fixed. On CRTs, the faceplate glass surfaces are not parallel. Specular reflections from the first and second surfaces do not superimpose, for any given angle of incidence, at more than one point, tending to be near the center of the screen. Therefore, it is (in my opinion) not practical to attempt complete or physical characterization of the optical reflection transfer functions of the display.

As an alternative, the committee decided to try simple characterizations at two points, each having some face validity in the use of displays: - A large-area source, emulating the

walls of the office, or papers on the desk. ISO 9241 part 6 (as I recall) specifies the maximum wall luminance at 200 cd/m2. Simple analysis shows that papers on the desk, under 500 lux, have a diffuse reflected luminance of somewhere around 125 cd/m2. - A near-point source, emulating an unavoidable high luminance located nearly behind the user. Examples would be distant windows, luminaires in adjoining offices, etc.
- To avoid the specular reflection of the high luminance, the screen would have to be angled such that a specular reflection would be formed of walls, papers on the table, etc. Therefore, compliance with the standard would be an either/or thing. The screen has to be able to stand up to either small, high-luminance sources, or large, moderate luminance ones (or both).

Optically, the large-area and small-area reflection characteristics will be equivalent for specular reflections. If, however, there is a significant point-spread function, the small-area source will produce a smaller refllection coefficient, in proportion to the amount of the reflection that falls outside the field of view of the photometer. That light is lost from the measurement.

Initially, the geometry of the small-source measurement was set up with a CRT in mind. Assuming a convex sphere, the specular reflection forms a virtual image of the source somewhat behind the faceplate, with a magnification of less than one. We looked at current CRT practice, and found a radius of curvature of arount 600 mm was the smallest in common use.

The measurement geometry and photometer field were selected so that the photometer field would fit comfortably within the reflected virtual image of the small source. Therefore, for a purely specular surface, having a radius of curvature of 600 mm or more, the total specular reflection coefficient would be measured accurately, provided the measurement was taken at the point on the screen where the first and second surface specular rays (and therefore the virtual images) coincided.

The photometer should be focused on the center of the virtual image of the source. The correct aim and focus distance can be difficult to establish if the virtual image is not clearly visible. It can be made visible on a textured surface by moistening the surface. Soapy water, alcohol, glycerin, even light (non-silicone) oil might be used, followed by the cleaning spray of your choice.

In the limiting case, where a true point source is measured by a photometer with an infinitessimal angle of acceptance, what would be measured is the maximum amplitude of the reflection point-spread function. This is impractical. The photometer has a finite aperture*. Since the luminance of the diffused reflection of a point source has no central area of constant luminance, the measured reflected luminance depends critically on the photometer measurement aperture. Different instruments, and different measuring distances, will produce different results. This is unacceptable in a standard, so finite source and aperture angles are specified. The Japanese standard spot photometer has a minimum aperture of 20' (1/3 degree). That was the largest found, so that was taken as the maximum measuring aperture angle.

* Actually, several apertures are potentially significant. Using the Pritchard photometer as a reference, the lens diameter and focal length determine to some extent how much reflected light is accepted for measurement. The focal plane aperture determines the extent of the virtual (specular) reflection measured. The problem is that the scattered light does not come from the same plane as the virtual image. The virtual image is behind the screen surface; the diffused light is redirected from the screen surface. One or the other must always be out of focus in the measurement. By choosing the first-surface specular reflection (virtual image of the source) for focus, the lens aperture (both physical and numerical) exerts an effect on how much out-of-focus light from the screen surface will enter the measurement. Extended discussion of this issue finally came down to weariness and a consensus that most spot photometers out there use about a 50 mm diameter lens and about a 100 mm focal length, so initially the problem would be ignored. It may be revisited if labs get really divergent results on highly textured screens. A contrary view of this may be obtained from at least one of the Dutch members.

Initially, the small-source size was the smallest possible. Practically, this was simply the size which, when massaged through the reflection optics of the source-screen-photometer lens location (magnification, distance, etc) gave a virtual image whose apparent diameter (visual angle) was 30' (= 1.5 times the 20' photometer aperture).

Later, the Japanese did some psychophysics, relating reported acceptability of screen reflections to specular reflections measured with different source sizes. They found best correlation with source sizes of 1 and 3 degrees diameter, as measured from the screen surface. They preferred the 3 degree source, at least in part because measurements are easier. I liked the 1 degree, because interpretation is easier, and it measures moderate diffusion with more sensitivity. I think the consensus swung to 1 degree, but it is possible that this may be revisited in the future.

Optically, the bottom lines are that each method measures the convolution of the spatial distribution of a luminous source of finite size with the reflection point-spread function of the screen: - The large-area source is large enough that further increases in its size do not result in further increases in the luminance of its virtual image, measured in the center. This, in effect, measures total direct (specular plus partially diffuse) reflection coefficient. - The small-area source measures a different point on the convolution. Its size is a compromise between a true point source, which would make practical measurement totally photometer-dependent, and a practical source having validity in the workplace. It is small enough that, for most screen textures, a significant part of the direct reflection falls outside the virtual (specular) image. Therefore, the luminous intensity of the specular rays is reduced relative to a purely specular reflection. In measurement (and view by the user) the luminance of the center of the reflection is reduced relative to the equivalent specular reflection. (Of course, the area or visual solid angle over which the reflection is perceptible increases in proportion.)

As to practical measurements, I would point out that most CRTs today seem to use either silica coating or chemical etch. Both leave substantial areas of flat (or little-changed) glass. (This is a technical way of saying that they often don't do much.) (Silica coating, when it is heavy enough to be effective, often causes "sparkle." Ask me about "sparkle" some day.) The reduction in the coefficient of specular reflection may be hard to measure, relative to the large source.

Some LCDs have a fairly heavy texture on the cover sheet, and may be easier to use for demonstration measurements.

Good luck. Hope this wasn't too much more than you wanted to know about reflections.

- John

At 08:54 AM 8/26/97 -0400, you wrote:
>At 10:45 PM -0400 8/25/97, John Cone wrote:

>>Bob -

>>Yep - I still keep this account active, although I don't check it every da Best bet is email to jcone@nmsu.edu.

>>- John

>Hi, John.

>I'm very glad to find you. I hope all is going well for you. When you get
a moment, I'd be interested in knowing what you are up to at NMSU.

>My main reason for trying to contact you was to ask a question or two
about the ISO 9241-7 standard. Specifically, I'd like to understand the rationale underlying the use of the two sizes of specular luminaires. I understand the effect reported by Kubota (i.e., increasing specular reflection with increasing luminaire diameter). However, Kubota's data seem to be dependent on his measurement system--that is, a detector field of view larger than the luminaire diameter.

>If been making the ISO 9241-7 measurements with a 2-D CCD detector system.
When I setup the 30 deg off-axis condition, the small- and large-area luminaire readings are equivalent. This occurs since the required detector FOV are smaller than the projected image of the luminaires.

>Can you help me understand the intention of the ISO requirements?

>Many thanks,
>Bob

>----------------------------

>Robert J. Beaton, Ph.D., CPE
>Director, Displays and Controls Laboratory Human Factors Engineering Center, 549 Whittemore Hall Virginia Tech, Blacksburg, VA 24061-0118 USA

>TEL 540.231.8748 * FAX 540.231.3322 * EMAIL bobb@vt.edu TEL 540.953.3100 * FAX 540.953.3104 * EMAIL BobBeaton@aol.com WEB http://bobb.dcl.vt.edu, http://dcl.vt.edu, http://ergonorms.dcl.vt.edu

VITA

Gary S. Olacsi was born on May 24, 1970, in Pacoima, California. He received his Bachelor of Arts degree in Psychology from California State University Northridge in 1994. His interest in the human factors field led him to Virginia Polytechnic Institute and State University (Blacksburg, Virginia) where he studied Human Factors Engineering. Gary joined the Displays and Controls Laboratory in October 1996 and was a Virginia Tech Intern for the Image Quality Group of the Human Factors Lab at Eastman Kodak, Rochester N.Y. While an intern for Kodak, he conducted display reflectance measurement studies on a contemporary flat panel display technology. Currently, Gary is a Graduate Research Assistant in the Displays and Control Laboratory working towards a Ph.D. degree in Industrial and Systems Engineering (Human Factors option) at Virginia Tech. He is also an intern in the Human Research and Engineering Directorate of the Army Research Laboratory where he is working to support the Visual and Auditory Processes Branch. Gary is an active member of the Human Factors and Ergonomics Society student chapter at Virginia.