A Comparison of Seven Visual Fatigue Assessment Techniques In Three Data-Acquisition VDT Tasks

Chia-Fen Chi and **Fang-Tsan Lin**, National Taiwan University of Science and Technology, Taipei, Taiwan

We compared 7 methods of measuring visual fatigue – accommodation power, visual acuity, pupil diameter, critical fusion frequency (CFF), eye movement velocity, subjective rating of visual fatigue, and task performance – for their sensitivity to visual load. In the experiment, 10 participants performed a monitoring task at 2 viewing distances, read articles under 2 levels of screen contrast, and tracked visual targets at 2 different speeds. The same measurement techniques, excluding pupil diameter and eye movement velocity, were compared by extending the task time from 20 to 60 min with the same VDT tasks to test for possible improvement in sensitivity. The results indicated that sensitivities of accommodation power, visual acuity, and CFF were greatly improved by a longer task period, but these 3 measurement techniques did not distinguish among tasks. Pupil diameter, eye movement velocity, and subjective rating of visual fatigue were sensitive in differentiating tracking from reading and monitoring tasks. Eye movement velocity and subjective rating were sensitive to the changes in target velocity of the tracking task. Although task performance was not directly comparable to other measurement techniques, it helped to ensure that participants maintained the same performance level by devoting more resources to the high-load conditions. Actual or potential applications of this research include using some of these assessment techniques for the design of adaptive displays.

INTRODUCTION

Visual display terminals (VDTs) can be used to store and convey static and dynamic visual information, and thus they have inevitably become interfaces in both the office and the industrial workplace. However, VDTs have brought with them numerous complaints of visual fatigue, mental load, and musculoskeletal pains. Of these complaints, visual fatigue is the most pronounced and prevalent (Kamienska-Zyta, 1993; Knave, Wibom, Voss, Hedstrom, & Bergqvist, 1985). Although complaints of visual fatigue have not been completely resolved, it has been found to be intensified by the following characteristics of VDT tasks: low luminance contrast, near viewing distance, and fast-moving targets. The American National Standards Institute (1988) suggested that the minimum luminance contrast of character details within or between characters should be 3:1, and preferably 7:1 and greater. Zhu and Wu (1990) recommended the luminance contrast range of 9:1 to 11:1 for acceptable task performance and visual fatigue.

Jaschinski-Kruza (1988) stated that one reason for reports of visual fatigue from VDT workers is that the viewing distance is usually less than their dark-focus point. During performance of a VDT task, which is typically near work, a person's ciliary accommodation muscle changes the optical power of the lens to form a sharp image on the retina, and the horizontal extraocular muscles converge the

Requests for reprints should be sent to Chia-Fen Chi, Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan 106. **HUMAN FACTORS**, Vol. 40, No. 4, December 1998, pp. 577–590.

axes of the eyes to fuse the two retinal images. These oculomotor mechanisms of accommodation and convergence are increasingly strained as the viewing distance shortens. Jaschinski-Kruza (1988) showed that visual strain was greater at a viewing distance of 50 cm than at 100 cm, irrespective of the individual's dark-focus point.

Saito, Taptagaporn, and Salvendy (1993) and Iwasaki and Kurimoto (1987) measured the degree of eyestrain caused by visual work by measuring the temporary change in accommodation power following VDT work. Haider, Kundi, and Weissenbock (1982) measured visual acuity before and after the working period in order to indicate functional changes in the accommodation mechanism, and they reported that the decrease of visual acuity during work correlated well with different lengths of working periods.

An increase in pupil size adversely affects the depth of focus and the precision required of the accommodative responses; thus pupil size has been used as an indicator of visual discomfort. For example, Saito et al. (1993) considered positive CRT displays to be better than negative ones because the pupil diameter was 10% smaller when viewing a positive-type CRT. Taptagaporn and Saito (1990) concluded that for all lighting conditions, a positive display caused less adaptive strain on the eye than a negative display because there were smaller pupil diameter differences among viewing a positive CRT display, a manuscript, and a keyboard. However, the relationship between pupil diameter and feelings of visual comfort has not been confirmed (Taptagaporn & Saito, 1990).

Kanaya (1990) stated that one of the causes of VDT-related complaints was the high density of office work created by the introduction of the VDT. For example, VDT workers are paced by the processing speed of the VDT; they have to input commands and continuously react to the information consequently displayed on the screen, forcing them to engage in very intense and stressful work as the processing speed of the VDT increases. Ishihara, Miyao, Tamura, Iguchi, and Furukawa (as cited in Nishiyama, 1990) found that the faster figures were presented on a CRT screen, the more the pupil area increased. When luminance was adequately controlled, however, task-evoked pupillary responses were shown to reflect information processing demands within capacity limits (Backs & Walrath, 1992; Beatty, 1982). Therefore, besides lighting and screen parameters, both information processing load and eye movement velocity can create an increase in the pupil diameter that adversely affects the depth of focus.

From the measurement of eye movements, Saito et al. (1993) discovered that both the amplitude and frequency of eye movements during VDT work were relatively high. VDT operators had to move their eyes 2.5 times faster than traditional clerical workers who did not use VDTs. Hallett (1986) indicated that the extraocular muscle forces are a function of fixation position and angle of eye movement (saccadic amplitude). Extreme torsion will stress the optic nerve or lead to conjunctivitis and will therefore lead to damage and pain. Thus visual fatigue could be partially induced by the action of the eyeball and eye muscles, in particular when the operation of internal and external muscles of the eye is in excess of that required for normal levels of eye movement.

Additional Measures of Visual Fatigue

Iwasaki, Kurimoto, and Noro (1989), citing Osaka's (1985) research, found that green and yellow critical fusion frequency (CFF) deteriorated significantly 30 min after loading; however, the red CFF decreased significantly after performing the visual task for only 15 min. This decrease of CFF, which might indicate deterioration of the retinal function, was confirmed in Iwasaki and Akiya (1991). They presented a simple mathematical addition task to one eye (the loaded eye) while the other eye acted as the control. Because the relative decrease in CFF for the loaded eye compared with the CFF for the control eye was found to be proportional to the time for performing the visual task, they suggested that the decrease in CFF reflected a decline in the activity of the retina or the optic nerve.

Subjective rating scales are easy to administer and can at times be more sensitive than objective measurements (Hwang, Wang, & Her, 1988; Saito, Sotoyama, Saito, & Taptagaporn, 1994). However, subjective measures are not very diagnostic; they are only global indicators of workload. Hence, subjective rating scales have often been used to cross-validate more objective measurements of visual fatigue. Primary task measurement, which measures the actual performance on the task, is one major technique for evaluating mental workload (Wickens, 1992). In the current experiment, task performance data were collected to ensure that participants maintained the same level of performance by devoting more resources to high-load conditions.

Accommodation power, visual acuity, pupil diameter, eye movement velocity, CFF, subjective rating scales, and task performance are all potentially useful measures of visual fatigue. These measures have seldom been compared on their sensitivity to the visual load of VDT tasks. Some have been criticized as being sensitive only to strenuous visual (unnatural) work performed over a long task period. Therefore, the purpose of the current study was to compare these commonly used visual fatigue measurement techniques based on their sensitivities to changes in visual load. A sensitive measurement technique was defined as one that can reliably discriminate between two visual load levels for performing each VDT task within a relatively short period.

Given that degree of visual discomfort is related to the amount of time spent looking at VDT screens (Knave et al., 1985), three screen-intensive tasks (monitoring, reading, and tracking) were implemented. The primary consideration was to determine which measurement technique was best for assessing visual load created by various task factors (e.g., short viewing distance, low luminance contrast, and fast moving targets). For example, short viewing distance may increase the muscle strain in the oculomotor mechanisms of accommodation and convergence, and this change in muscle strain can best be measured by the change in accommodation power. Low luminance contrast may reduce the visual resolving ability in the retina, and this change in visual resolving ability can be best indicated by the decrease in CFF. Tracking targets moving at high velocity may create eye movement load, and this load can be best measured by the eye movement velocity, the pupil diameter, or both.

The current experiment assessed how viewing distance, luminance contrast, and target velocity affected each of the measures and the interrelationships among the measures. Viewing distance, luminance contrast, and target velocity were manipulated in the monitoring, reading, and tracking tasks, respectively. Three loading factors were manipulated using three different tasks because task characteristics could have a direct impact on the sensitivity of each measurement technique, and particular tasks were expected to highlight certain load manipulations (e.g., a tracking task might highlight the velocity of a moving target).

EXPERIMENT 1: 20-MIN VDT EXPOSURE

Method

Participants. There were 10 participants (5 women and 5 men) aged 24–32 with a mean age of 28.5 years in Experiment 1. All participants were pretested for at least 0.8 visual acuity (with corrective lenses if needed).

Experimental design. A three factor within-subject design was used for the sensitivity analysis. Task, load, and measurement technique were the factors with three, two, and seven levels, respectively. Given that the loading factor was qualitatively different for each of the three tasks, a hierarchical analysis of variance (ANOVA) model in which the load manipulation was nested within each performed task was used for the data analysis. Each participant was tested individually on three days; consecutive sessions were separated by at least 30 min. The order of presentation of task and load levels was completely randomized across participants. A 30-min training course was provided for the participants before the experiment.

Seven measurement techniques were compared within the shortest possible time required to generate visual fatigue symptoms. Given that Iwasaki et al. (1989) discovered a statistically significant decline in the red CFF after performing a visual task for 15 min, participants were required to perform monitoring and tracking tasks for 20 min, and the length of the reading material was designed to take approximately 20 min. General illumination provided by indirect lighting was fixed at 250 lx throughout the experiment to minimize the influences of extraneous variables on pupil diameter. The display was adjusted to obtain a screen contrast of 11:1 (background = 62.8 cd/m^2, character = 5.7 cd/m^2) for all task conditions except for the low-contrast reading task. Viewing distance was fixed at 80 cm for all task conditions except for monitoring (40 cm).

Before each task was performed, baseline values of accommodation power, CFF, and visual acuity were measured. CFF was the mean of two ascending trials, whereas accommodation power was the mean of two near point accommodation trials approaching from near to far. If two repetitive measurements differed by more than 10%, a third measurement was taken to derive the mean. For all experimental conditions, two extra near-point accommodation trials were taken for the determination of accommodation power.

After each experimental task was performed, a rating scale for seven descriptive items was shown on the screen to collect a subjective rating of visual discomfort (Laubli, Hunting, & Grandjean, 1981; Saito et al., 1994): (a) "My eyes feel tired (visual fatigue)"; (b) "Eye dry, irritated or burning"; (c) "Eye pain"; (d) "Hard to focus my vision"; (e) "Double vision on screen"; (f) "Flicker vision"; and (g) "Headache." Each item was rated on a five-point scale for severity of discomfort. After the participant completed the subjective rating scale, accommodation power, CFF, and visual acuity were measured again. Pupil size, eye movement amplitude, and frequency data were collected during the task performance. Subjective rating, accommodation power, CFF, and visual acuity measurements took about 30 s, 45 s, 30 s, and 25 s, respectively. Thus the measurement sequence was not considered to have a significant effect on the measurement sensitivity.

Apparatus. A 386 PC was used to generate the displays, to control and time the experimental events, and to collect and analyze the eye movement data, pupillary response data,

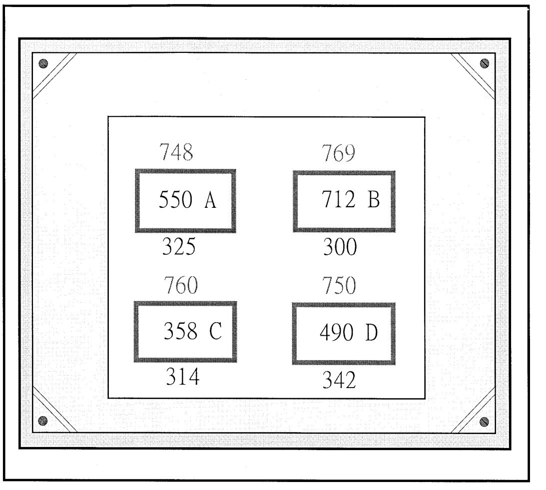
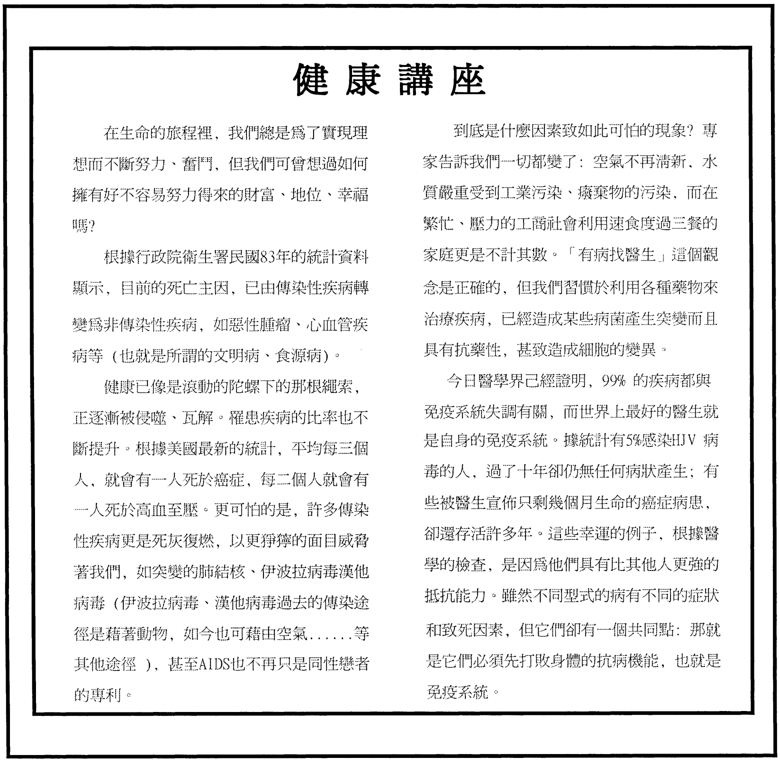


Figure 1. VDT screen for the monitoring task.

and other experimental data. Display screens were presented on a 30-cm wide × 22-cm tall high-resolution color graphics terminal. Eye movement and pupillary response data were collected at a rate of 60 Hz using an Applied Science Laboratories Eye View Monitor and TV Pupillometer (system model 1994-S). In order to maintain a fixed viewing distance between the participant's eyes and the screen, the participants used a chin rest. All photometry to calibrate the luminance of the displays was performed with a Minolta CRT color analyzer CA-100. A critical flicker frequency tester (Model: 12023A) manufactured by Lafayette Instrument Company and a vision tester (OPTEC 2000) manufactured by Stereo Optical were used to measure the CFF and visual acuity, respectively. A VDT near-point tester (TOMEY NP-200) manufactured by Toyo Physical corporation in Japan was used to measure the accommodative power.

Monitoring task. Participants were required to monitor a screen that contained four digital displays (see Figure 1). Four three-digit numbers were shown on four displays. These values varied every 500 ms and had to be kept between an upper bound and a lower bound presented above and below the boxed numbers, respectively. When the inspected numbers on these displays were out of bounds, participants had to respond by hitting a corresponding key within 1 s to rectify the situation. The number shown on that display would then be reset to be within



健康講座

在生命的旅程裡，我們總是為了實現理想而不斷努力、奮鬥，但我們可曾想過如何擁有好不容易努力得來的財富、地位、幸福嗎？

根據行政院衛生署民國83年的統計資料顯示，目前的死亡主因，已由傳染性疾病轉變為非傳染性疾病，如惡性腫瘤、心血管疾病等（也就是所謂的文明病、食源病）。

健康已像是滾動的陀螺下的那根繩索，正逐漸被侵噬、瓦解。罹患疾病的比率也不斷提升。根據美國最新的統計，平均每三個人，就會有一人死於癌症，每二個人就會有一人死於高血至壓。更可怕的是，許多傳染性疾病更是死灰復燃，以更猙獰的面目威脅著我們，如突變的肺結核、伊波拉病毒漢他病毒（伊波拉病毒、漢他病毒過去的傳染途徑是藉著動物，如今也可藉由空氣......等其他途徑 ），甚至AIDS也不再只是同性戀者的專利。

到底是什麼因素致如此可怕的現象？專家告訴我們一切都變了：空氣不再清新、水質嚴重受到工業污染、廢棄物的污染，而在繁忙、壓力的工商社會利用速食度過三餐的家庭更是不計其數。「有病找醫生」這個觀念是正確的，但我們習慣於利用各種藥物來治療疾病，已經造成某些病菌產生突變而且具有抗藥性，甚致造成細胞的變異。

今日醫學界已經證明，99% 的疾病都與免疫系統失調有關，而世界上最好的醫生就是自身的免疫系統。據統計有5%感染HJV 病毒的人，過了十年卻仍無任何病狀產生；有些被醫生宣佈只剩幾個月生命的癌症病患，卻還存活許多年。這些幸運的例子，根據醫學的檢查，是因為他們具有比其他人更強的抵抗能力。雖然不同型式的病有不同的症狀和致死因素，但它們卻有一個共同點：那就是它們必須先打敗身體的抗病機能，也就是免疫系統。

Figure 2. VDT screen for the reading task.

limits. Approximately every 20 s, one and only one of the digital displays went out of bounds.

Viewing distances of 80 and 40 cm were tested to investigate whether the visual fatigue symptoms primarily related to the increase in oculomotor strain at a short viewing distance. All numbers were about 2.5 mm in height and width and subtended a visual angle of 0.358° from the 40-cm viewing distance. To make the visual targets subtend the same visual angle at 80 cm, all numbers were doubled in height and width. Task performance was measured by the combination of hit and correct rejection responses – that is, $d' = Z(P_1) + Z(P_2)$, where $Z(P_1)$ and $Z(P_2)$ are the normal deviates corresponding to correct rejection rate and hit rate (Chi & Drury, 1998).

Reading task. Participants were required to read a composition of about 15 000 Chinese words from the screen under two luminance contrast ratios (see Figure 2). Each page contained about 836 Chinese characters presented in a double-spaced, double-column style. Each character was about 5 mm in height and width, subtending a visual angle of 0.358° from 80 cm. Luminance of the screen characters was increased from 5.7 cd/m^2 to 10.4 cd/m^2 to reduce the luminance contrast ratio from 11:1 to 6.03:1 with the background luminance fixed at 62.8 cd/m^2. Performance was measured by the correct response rate out of 10 true/false questions administered after each reading task was performed.

Tracking task. Participants were required to track a scanning line running clockwise around a radar screen to detect the sudden appearance of the visual targets (see Figure 3). Two visual targets – a donut (signal) or a circle (nonsignal) – were present for 450 ms approximately every 2 s on the scanning line. Given that the signal to nonsignal ratio was about 1:1, participants were required to respond as quickly as possible only if the detected target was a donut. The diameters of the inner and outer circles on the donut were 4 mm and 8 mm, whereas the diameter of the circle was 8 mm. The scanning line (radius = 90 mm) was designed to sweep at two different

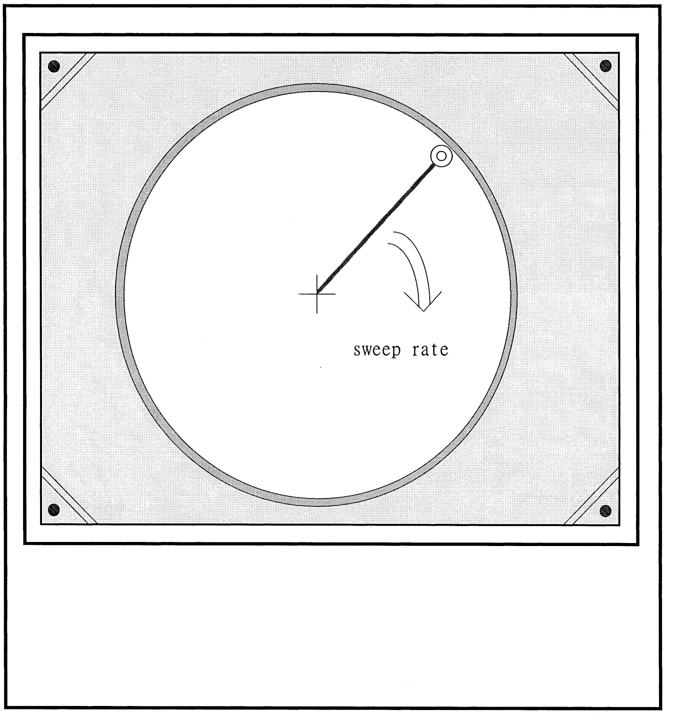


Figure 3. VDT screen for the tracking task. The two sweep rates were 0.2 Hz (72°/s) and 0.4 Hz (144°/s).

velocities of 0.2 cycle/s and 0.4 cycle/s and generate different rates of motion for the visual targets (8.1°/s and 16.2°/s). As with the monitoring task, performance was measured by the combination of hit and correct rejection responses into a detectability measure.

Results

All subjective ratings were subjected to a nonparametric Kruskal-Wallis analysis. It was discovered that subjective ratings were significantly different only on Question 1 (visual fatigue) and Question 2 (eye dry, irritated, or burning) among the six task conditions, $\chi^2(5) = 13.8$, $p = .031$; $\chi^2(5) = 14.2$, $p = .027$. Given that subjective ratings on these two items were highly correlated, $r^2 = .74$, $p = .0001$, the rating of visual fatigue was chosen to be compared with the other objective measurements.

Table 1 summarizes the mean and standard deviation values of the accommodation power, visual acuity, and CFF before and after each designated task, and the pupil diameter, eye movement velocity, subjective rating of visual fatigue, and task performance measures (correct response rate in reading and d' in monitoring and tracking). Eye movement velocity (°/s) was determined by the product of the mean saccadic amplitude and the frequency (Saito et al., 1993). The differences in accommodation power, visual acuity, and CFF measured before and after each performed task were compared with the other measurement techniques.

Individual hierarchical ANOVAs. All seven measurement techniques were considered as within-subject measures in the analysis. Initially, each of the dependent measures (except task performance) was separately transformed to *Z* scores (Wierwille & Connor, 1983) across all six conditions (3 tasks × 2 loads) for direct comparisons among measures, regardless of the original scale of measurement. Task performance was excluded from the standardized transformation because of a lack of comparability across performance measures for the different tasks. The results of Bartlett's test (Winer, 1971) indicated a strong inhomogeneity of variance among the different measurement techniques, $\chi^2(5) = 579.3$, $p < .005$; thus an overall ANOVA containing

TABLE 1: Means and Standard Deviations for Performance Measures in Experiment 1

Condition	Accom. Power (Diopters) Before/After	Δ Accom. Power	Visual Acuity (s) Before/After	Δ Visual Acuity	CFF (Hz) Before/After	Δ CFF	Pupil Diameter (mm)	Movement Velocity (deg/s)	Visual Fatigue	Task Performance
Monitoring (80 cm)	8.24/7.62	0.62	1.02/0.85	0.17	35.2/34.0	1.2	4.00	7.59	2.6	5.04
	(3.10)/(3.10)	(0.97)	(0.19)/(0.29)	(0.22)	(3.1)/(2.8)	(1.6)	(0.66)	(1.46)	(1.1)	(0.54)
Monitoring (40 cm)	8.26/7.67	0.59	1.01/0.75	0.26	35.0/33.1	1.9	3.87	7.64	3.3	5.11
	(3.09)/(3.17)	(1.11)	(0.19)/(0.30)	(0.13)	(2.8)/(2.9)	(2.8)	(0.47)	(1.36)	(1.1)	(0.55)
Reading (baseline)	8.22/7.81	0.41	1.02/0.87	0.15	35.1/33.9	1.2	3.71	6.68	2.3	0.96
	(3.15)/(3.27)	(0.92)	(0.13)/(0.23)	(0.18)	(3.1)/(2.8)	(2.0)	(0.78)	(1.53)	(0.9)	(0.10)
Reading (low-contrast)	8.21/7.59	0.62	1.02/0.78	0.24	35.5/32.9	2.6	3.92	6.89	2.8	0.91
	(3.09)/(3.45)	(0.85)	(0.18)/(0.32)	(0.23)	(3.0)/(2.8)	(2.4)	(0.88)	(1.67)	(1.1)	(0.10)
Tracking (0.2 Hz)	8.24/7.75	0.49	1.03/0.83	0.20	35.0/33.8	1.2	4.68	5.84	3.4	1.88
	(3.08)/(3.50)	(1.08)	(0.21)/(0.25)	(0.08)	(3.1)/(2.1)	(1.2)	(0.74)	(0.96)	(1.0)	(0.55)
Tracking (0.4 Hz)	8.18/7.39	0.79	1.01/0.87	0.14	34.9/33.7	1.2	4.93	12.98	4.6	1.60
	(3.12)/(3.46)	(0.70)	(0.18)/(0.22)	(0.10)	(2.9)/(2.2)	(1.5)	(0.84)	(1.32)	(0.5)	(0.76)

Note: Standard deviations in parentheses.

TABLE 2: Summary of the Individual ANOVAs of Experiment 1

	Task Effect		Load (Task) Effect	
Measurement Technique	*F*	*p*	*F*	*p*
Accommodation power difference	0.05	.9518	0.77	.5191
Visual acuity difference	0.31	.7383	2.86	.0552
CFF difference	0.73	.4946	1.93	.1491
Pupil diameter	10.22	.0011* [a]	2.06	.1297
Eye movement speed	17.90	.0001** [a]	65.18	.0001**
Rating of visual fatigue	6.47	.0076* [b]	11.85	.0001**

[a]Tracking > monitoring, reading. [b]Tracking > reading.
*$p < .01$. **$p < .001$.

the six measurement techniques was considered inappropriate for the analysis.

Individual hierarchical ANOVAs with the load manipulation nested within each performed task were performed for each of the measurement techniques to test the main effects of the task and load. Given that the *Z* transformation is a linear one, the *F* and *p* values of the individual ANOVAs for each measurement technique were identical to what would be obtained with the untransformed data. The results of the *F* and *p* values from the individual ANOVAs are provided in Table 2.

Based on the individualized ANOVA in Table 2, accommodation power, visual acuity, and CFF were found not to be significant in differentiating between the three tasks and between the two load levels. Pupil diameter was sensitive in distinguishing among the three tasks, $F(2, 18) = 10.22$, $p = .0011$, but was not sensitive in distinguishing between the two load levels, $F(3, 27) = 2.06$, $p = .1297$. Eye movement velocity was sensitive in distinguishing among the three tasks, $F(2, 18) = 17.90$, $p = .0001$, and between the two load levels, $F(3, 27) = 65.18$, $p = .0001$. Similarly, subjective rating was sensitive to the task differences, $F(2, 18) = 6.47$, $p = .0076$, and the load differences, $F(3, 27) = 11.85$, $p = .0001$.

A series of Newman-Keuls tests among the three tasks ($p = .05$) was performed on those measurement techniques (pupil diameter, eye movement velocity, and subjective rating) that had a significant task effect. Results from the Newman-Keuls tests indicated that pupil diameter and eye movement velocity for tracking were significantly greater than for monitoring and reading. Participants reported a significantly greater degree of visual fatigue during tracking than during reading. The differences in the subjective ratings between tracking and monitoring or between monitoring and reading were not significant.

In addition, *t* tests for comparing the two load levels nested in the three VDT tasks were conducted on eye movement velocity and subjective rating (which had significant load effects) and the task performance measure for lack of comparability between tasks. Given that the loading factor was qualitatively different for each of the three tasks, a significance level appropriate for protection against Type I error would be to divide the usual significance level, .05, by 3 (i.e., approximately .0167) for three pairs of the follow-up *t* tests. The results indicated that only the change in scanning rate in tracking generated significantly different eye movement velocities ($p = .0001$) and subjective ratings ($p = .0039$). The nonsignificance of eye movement velocity in differentiating between the two load levels for performing two other tasks indicated that eye movement velocity was sensitive only to changes in target velocity for dynamic visual tasks. The nonsignificant load effect on task performance indicated that participants maintained the same level of performance across two load levels and that more resources were apparently devoted to the high-load conditions.

Correlation among visual fatigue measurements. The Pearson correlation among all (n = 60) pairs of measurements for 10 participants

TABLE 3: Correlation between Performance Measures in Experiment 1

	Visual Acuity	CFF Difference	Pupil Diameter	Eye Movement Velocity	Rating of Visual Fatigue
Accom. power difference	–.04	.14	–.03	.13	–.04
Visual acuity difference		.27*	.03	–.26*	.07
CFF difference			.05	–.16	.01
Pupil diameter				.21	.25*
Eye movement velocity					.39**

*$p < .05$. **$p < .01$

in six conditions was calculated (p = .01). Task performance was excluded from the correlation analysis because of a lack of comparability across performance measures for the different tasks. Table 3 shows that eye movement velocity was significantly related to subjective rating of visual fatigue, $r = .39$, $p = .002$. Other correlations that approached significance were between decreases in visual acuity and CFF, $r = .27$, $p = .04$, between eye movement velocity and a decrease in visual acuity, $r = –.26$, $p = .05$, and between pupil diameter and subjective rating of visual fatigue, $r = .25$, $p = .05$.

EXPERIMENT 2: 60-MIN VDT EXPOSURE

In Experiment 1, changes in accommodation power, visual acuity, and CFF were all found to be insensitive. It was suspected that the 20-min task was too short and that sensitivity of these measurement techniques could be improved if the same tasks were performed for a longer period.

Participants. The 10 participants (5 women and 5 men) in Experiment 2 were different from those in Experiment 1. Their ages ranged from 18 to 25 years (mean, 23.5 years).

Experimental design. Experiment 2 was identical to Experiment 1: a three-factor within-subject design with task, load, and measurement technique varied at three, two, and five levels, respectively. The same monitoring, reading, and tracking tasks were performed for approximately 60 min with the same loading factors manipulated for the three VDT tasks. The same measurement techniques were compared in Experiment 2, except for pupil diameter and eye movement velocity, which were omitted.

Results

Table 4 summarizes the mean and standard deviation values of the accommodation power, visual acuity, and CFF before and after each designated task, and the subjective rating of visual fatigue and performance measures collected after each task was performed.

Individual hierarchical ANOVAs. Following the procedures used in Experiment 1, the four measurement techniques, excluding task performance, were standardized and subjected to a hierarchical ANOVA procedure. Again, the results of Bartlett's test indicated a strong inhomogeneity of variance among the different techniques, $\chi^2(3) = 216.7$, $p < .01$. Thus individual hierarchical ANOVAs with the load manipulation nested within each performed task were performed for each of the measurement techniques to test the main effects of task and load (results are summarized in Table 5).

From the individual ANOVAs in Table 5, accommodation power, $F(3, 27) = 10.44$, $p = .0001$, visual acuity, $F(3, 27) = 22.11$, $p = .0001$, and CFF, $F(3, 27) = 7.16$, $p = .0011$, were found to be sensitive in distinguishing between the two load levels but were not sensitive in distinguishing among the three tasks ($p > .05$). Thus the sensitivity of these measurement techniques was significantly improved by using a longer task period. Subjective rating was sensitive to both

TABLE 4: Means and Standard Deviations for Performance Measures in Experiment 2

Condition	Accom. Power (Diopters) Before/After	Δ Accom. Power	Visual Acuity (s) Before/After	Δ Visual Acuity	CFF (Hz) Before/After	Δ CFF	Visual Fatigue	Task Performance
Monitoring (80 cm)	11.1/10.4	0.7	1.11/0.98	0.13	35.5/34.3	1.2	2.5	4.28
	(1.4)/(1.2)	(0.7)	(0.15)/(0.17)	(0.14)	(1.3)/(1.8)	(1.5)	(0.8)	(0.24)
Monitoring (40 cm)	11.2/9.8	1.4	1.11/.089	0.22	35.6/33.6	2.0	3.2	4.22
	(1.2)/(1.1)	(0.6)	(0.13)/(0.12)	0.08	(1.2)/(1.5)	(1.3)	(0.9)	(0.29)
Reading (baseline)	10.9/9.8	1.1	1.14/1.07	0.07	35.6/33.9	1.7	2.5	0.9
	(1.8)/(1.0)	(1.4)	(0.10)/(0.15)	(0.09)	(1.3)/(1.5)	(1.3)	(1.0)	(0.1)
Reading	10.9/9.6	1.3	1.15/ 0.94	0.21	36.0/33.3	2.7	2.9	0.9
(low-contrast)	(1.7)/(0.8)	(1.2)	(0.08)/(0.13)	(0.10)	(1.9)/(1.0)	(1.1)	(0.6)	(0.1)
Tracking (0.2 Hz)	10.7/10.0	0.7	1.10/ 0.99	0.11	34.9/33.0	1.9	2.6	1.70
	(1.6)/(1.3)	(0.7)	(0.18)/(0.19)	(0.11)	(1.4)/(1.4)	(0.9)	(0.5)	(0.38)
Tracking (0.4 Hz)	10.8/9.4	1.4	1.10/ 0.88	0.22	34.7/32.5	2.2	4.7	1.48
	(1.7)/(1.4)	(0.8)	(0.15)/(0.17)	(0.11)	(1.3)/(1.4)	(0.7)	(0.5)	(0.59)

Note: Standard deviations in parentheses.

TABLE 5: Summary of the Individual ANOVAs of Experiment 2

Measurement Technique	Task Effect		Load (Task) Effect	
	F	*p*	*F*	*p*
Accommodation power difference	0.23	.7971	10.44	.0001**
Visual acuity difference	0.41	.6681	22.11	.0001**
CFF difference	0.97	.3991	7.16	.0011*
Rating of visual fatigue	13.61	.0003** [a]	33.99	.0001**

[a]Tracking > monitoring, reading.
*$p < .01$. **$p < .001$.

task differences, $F(2, 18) = 13.61$, $p = .0003$, and load differences, $F(3, 27) = 33.99$, $p = .0001$. Newman-Keuls tests were performed on the three tasks on the subjective rating measures that had significant task effects. Participants reported significantly more visual fatigue during tracking than during reading and monitoring. No significant difference in subjective rating was found between monitoring and reading.

With the appropriate significance level for protection against Type I error ($\alpha = .0167$), *t* tests comparing the two load levels nested under the three tasks on all measurement techniques indicated that only visual acuity was sensitive to the luminance contrast ($p = .005$) in reading. A number of other results approaching significance should be noted: Accommodation power was sensitive to the viewing distance in monitoring ($p = .043$) and tracking speed in tracking ($p = .049$), and visual acuity was sensitive to the tracking speed in tracking ($p = .041$). The *t* tests on CFF did not approach significance for the three VDT tasks, though the individual ANOVA showed significant load effects. As with Experiment 1, the subjective rating of visual fatigue was sensitive to the target velocity manipulated in the tracking task ($p = .0001$). A nonsignificant load effect on task performance indicated that participants maintained the same performance level across two load levels by devoting more resources to the high-load conditions.

In Experiment 2, significant correlations were found between the decreases in visual acuity and accommodation power, $r = .45$, $p = .0004$, and between the decreases in visual acuity and CFF, $r = .32$, $p = .01$. No other significant correlations were found.

DISCUSSION

The lack of significant results in Experiment 1 for accommodation power, visual acuity, and CFF appear to contradict other research findings (Haider et al., 1982; Iwasaki & Kurimoto, 1987; Jaschinski-Kruza, 1988; Saito et al., 1993). Experiment 2 showed that the sensitivities of accommodation power, visual acuity, and CFF to load differences were greatly improved by extending the task performance time from 20 to 60 min. Therefore, visual fatigue comparisons between studies must take differences in task time into account. In Experiment 2, differences in accommodation power found by manipulating viewing distance and target velocity suggested that a short viewing distance and a fast-moving target were harmful to the oculomotor mechanisms of accommodation. Differences in visual acuity from manipulating luminance contrast and tracking speed indicated that low luminance contrast and a fast-moving target can lead to an overall decrease in visual function. However, except for the decrease in visual acuity that was significantly changed by the luminance contrast, these other results only approached significance.

The luminance contrast manipulation in reading most nearly approached significance among *t* tests involving CFF, but even this did not reach statistical significance to demonstrate sensitivity to deterioration in retinal function.

From the results of various measurement techniques, the significant load effect arising from a short viewing distance, low luminance contrast, and a fast-moving target confirmed that these were the significant stressors in VDT tasks (Jaschinski-Kruza, 1988; Saito et al., 1993; Zhu & Wu, 1990). The VDT task environment should be redesigned to prevent these task conditions.

For the three interactive VDT tasks tested, the major differences in measures of visual fatigue were between tracking and the other two tasks (i.e., between dynamic and static visual tasks). The absence of significant task effects for accommodation power, visual acuity, and CFF indicated that these commonly used measurement techniques are not sensitive enough to compare tasks that differ in one of the important characteristics of VDT tasks: target velocity (Saito et al., 1993). However, eye movement velocity and subjective rating of visual fatigue were both very sensitive to the changes in target velocity for this dynamic visual task.

The scan path analysis of eye movement showed that participants tracked targets of low and high velocities somewhat differently. Participants used smooth pursuit eye movement along the scanning circle to track the low-velocity target and used large saccadic movements to traverse across the scanning circle to follow the fast-moving target. Eye movement velocity was not sensitive to differences between load levels for our static visual tasks (reading and monitoring), and its application is somewhat limited by the requirement for complex measurement equipment and a trained experimenter. We once suspected that the use of eye movement velocity based only on saccades quantified visual load, and we proposed a more sophisticated model (Chi & Lin, 1997) that could describe eye movement pattern and quantify visual load. However, complexity and lack of diagnosability greatly reduced its usefulness.

The negative relationship between visual acuity and eye movement velocity in Experiment 1 was expected because movement of a target (or observer) decreases visual acuity, and a faster target causes an increase in the pupil diameter, which affects both depth of focus and visual acuity (Ishihara, Miyao, Tamura, Iguchi, & Furukawa, as cited in Nishiyama, 1990). In contrast to Taptagaporn and Saito's (1990) finding, pupil diameter was significantly related to the subjective rating of visual fatigue. It was suspected that the subjective rating could be affected by accommodation power, which was in turn adversely affected by an increase in pupil diameter. However, the correlation supporting this interpretation was not strong.

Significant correlations between decreases in visual acuity and accommodation power and between decreases in visual acuity and CFF in Experiment 2 indicated that decreases in visual acuity, accommodation power, and CFF all could have been caused by a general decrease in eye function. The highly intercorrelated results between these measurement techniques confirmed that extraocular muscles for controlling eye movements, the oculomotor mechanism of accommodation, and the pupil control system cannot be easily isolated (Megaw, 1995).

Malmstrom, Randle, Murphy, Reed, and Weber (1981) suggested the need for an integrated model for studying visual fatigue. The lack of a simple one-to-one correspondence between task characteristics and measurable visual symptoms is somewhat disappointing to human factors engineers seeking diagnostic tools to improve VDT task conditions. However, as visual fatigue has become a great concern in the design of human-computer interfaces, some of these assessment techniques could potentially be applied to the design of adaptive displays. That is, the human-computer interface (display) could be adapted to the level of visual fatigue evaluated by reliable indicators, such as pupil diameter and eye movement velocity. Perhaps the wording and tone of help messages could be adapted depending on the visual fatigue of the users (Eberts, 1994).

The current experiment compared the sensitivities of visual fatigue measurement techniques with the commonly seen stressors in VDT tasks: short viewing distance, low luminance contrast, and fast-moving targets. In contrast to most other research, we chose to manipulate the load levels differently for three different VDT tasks in order to compare how the sensitivities of measurement techniques varied among tasks. However, such a design could possibly confound the loading manipulations with the differences among the three VDT tasks in some unanticipated manner. The significant findings of the current study suggest the need for further study to test all three loadings (viewing distance, luminance contrast, and target speed) using a dynamic visual task. Such a study could further reveal the effects of each of the loading manipulations and their connection with measurable visual fatigue symptoms.

CONCLUSIONS

In these experiments, 10 participants performed a monitoring task at two viewing distances, read articles under two levels of screen contrast, and tracked a visual target moving at two different velocities. The same set of measurement techniques (excluding pupil diameter and eye movement velocity) were further compared by extending the task time from 20 to 60 min to determine possible improvements in sensitivity. Task performance was measured to ensure that participants maintained the same level of performance across different load conditions. The results indicated that the sensitivities of visual fatigue assessment methods varied widely between tasks and between loading manipulations. Sensitivities of accommodation power, visual acuity, and CFF were greatly improved by a longer task period. However, these three measurement techniques were highly interrelated and were not sensitive enough to differentiate between the different VDT tasks. Eye movement velocity and subjective rating of visual fatigue were sensitive in differentiating tracking from reading and monitoring tasks and were also sensitive to target velocity changes in the tracking tasks.

ACKNOWLEDGMENTS

This research was supported by the National Science Council, Taiwan, Republic of China (NSC-86-2213-E011-013). The authors gratefully acknowledge Richard Jagacinski and Colin Drury for their editorial assistance and comments on earlier versions of this paper.

REFERENCES

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

Backs, R. W., & Walrath, L. C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Applied Ergonomics, 23*, 243–254.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*, 276–292.

Chi, C.-F., & Drury, C. (1998). Do people choose an optimal response criterion in an inspection task? *IIE Transactions, 30*, 257–266.

Chi, C.-F., & Lin, F.-T. (1997). A new method for describing search patterns and quantifying visual load using eye movement data. *International Journal of Industrial Ergonomics, 19,* 249–257.

Eberts, R. E. (1994). *User interface design*. Englewood Cliffs, NJ: Prentice-Hall.

Haider, M., Kundi, M., & Weissenbock, M. (1982). Worker strain related to VDUs with differently colored characters. In E. Grandjean & E. Vigliani (Eds.), *Ergonomic aspects of visual display terminals* (pp. 53–64). London: Taylor & Francis.

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 10.1–10.100). New York: Wiley.

Hwang, S. L., Wang, M. Y., & Her, C. C. (1988). An experimental study of Chinese information displays on VDTs. *Human Factors, 30*, 461–471.

Iwasaki, T., & Akiya, S. (1991). The significance of changes in CFF values during performance on a VDT-based visual task. In M. Kumashiro & E. D. Megaw (Eds.), *Towards human work: Solutions to problems in occupational health and safety* (pp. 352–357). London: Taylor & Francis.

Iwasaki, T., & Kurimoto, S. (1987). Objective evaluation of eye strain using measurements of accommodative oscillation. *Ergonomics, 30,* 581–587.

Iwasaki, T., Kurimoto, S., & Noro, K. (1989). The changes in color flicker fusion (CFF) values and accommodation times during experimental repetitive tasks with CRT display screens. *Ergonomics, 32*, 293–305.

Jaschinski-Kruza, W. (1988). Visual strain during VDT work: The effect of viewing distance and dark focus. *Ergonomics, 31,* 1449–1465.

Kamienska-Zyta, M. (1993). Ergonomics evaluation of the work of VDT operators in Poland. *Applied Ergonomics, 24*, 432–433.

Kanaya, S. (1990). Vision and visual environment for VDT work. *Ergonomics, 33,* 775–785.

Knave, B. G., Wibom, R. I., Voss, M., Hedstrom, L. D., & Bergqvist, U. (1985). UOV work with video display terminals among office employees: I. Subjective symptoms and discomfort. *Scandinavian Journal of Work, Environment, and Health, 11,* 457–466.

Laubli, T., Hunting, W., & Grandjean, E. (1981). Postural and visual load at VDT workplaces: II. Lighting conditions and visual impairments. *Ergonomics, 24,* 933–944.

Malmstrom, F. V., Randle, R. J., Murphy, M. R., Reed, L. E., & Weber, R. J. (1981). Visual fatigue: The need for an integrated model. *Bulletin of the Psychonomic Society, 17*(4), 183–186.

Megaw, E. D. (1995). The definition and measurement of visual fatigue. In J. R. Wilson & E. N. Corlett (Eds.), *Evaluation of human work* (pp. 840–863). London: Taylor & Francis.

Nishiyama, K. (1990). Ergonomic aspects of the health and safety of VDT work in Japan: A review. *Ergonomics, 33,* 659–685.

Osaka, N. (1985). Effect of VDT color, eccentricity, and adaptation upon visual fatigue. *Japanese Journal of Ergonomics, 21*(2), 89–95.

Saito, S., Sotoyama, M., Saito, S., & Taptagaporn, S. (1994). Physiological indices of visual fatigue due to VDT operations: Pupillary reflexes and accommodative responses. *Industrial Health, 32*, 57–66.

Saito, S., Taptagaporn, S., & Salvendy, G. (1993). Visual comfort in using different VDT screens. *International Journal of Human-Computer Interaction, 5*, 313–323.
Taptagaporn, S., & Saito, S. (1990). How display polarity and lighting conditions affect the pupil size of VDT operators. *Ergonomics, 33,* 201–208.
Wickens, C. (1992). *Engineering psychology and human performance.* New York: HarperCollins.
Wierwille, W. W., & Connor, S. A. (1983). Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. *Human Factors, 25*, 1–16.
Winer, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.
Zhu, Z., & Wu, J. (1990). On the standardization of VDT's proper and optimal contrast range. *Ergonomics, 33*, 925–932.

Chia-Fen Chi is a professor of industrial management at the National Taiwan University of Science and Technology. She received a Ph.D. in industrial engineering from the State University of New York at Buffalo in 1990.

Fang-Tsan Lin is an associate professor of industrial engineering at Da-Yeh University in Taiwan. He received a Ph.D. in 1996 in industrial management from the National Taiwan University of Science and Technology.

Date received: September 12, 1996
Date accepted: February 17, 1998