2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Investigation of Weighted Scales for Measuring Visual Fatigue in Screening Tasks

Yong Feng and Fei Chen *Senior Member, IEEE*

Abstract— **The "screening" trend of modern society has been a progressively increasing burden on the human visual system, and visual fatigue problems are attracting growing attention. Nowadays, subjective testing is the most widely used measure for visual fatigue; however, the low accuracy of subjective testing has been hindering its further improvement. Motivated by the idea of weighted scoring, this study investigated the effects of two weighted scales for measuring visual fatigue in screening tasks. Specifically, a questionnaire with 10 items collected from the classic scales was performed with eye-tracking testing in two typical screen visual fatigue experiments, i.e., searching and watching. Then the subjective scores were factor-analyzed into three subscales before attempting linear regression analyses, which set the dependents to two previously validated eye-tracking parameters, i.e., fixation frequency and saccade amplitude. Finally, two weighted scales were obtained in assessing visual fatigue of varying levels, which demonstrated the potential to improve testing accuracy of visual fatigue with the calibration of objective measurement.**

I. Introduction

The recent outbreaks of the Coronavirus Disease 2019 (COVID-19) pandemic have been rising globally, locking people with more time at home as well as on digital devices than ever. The suffering event expedites the "screening" (means screens everywhere) of our life with the booming of online learning, working and shopping, etc. Actually, we had been flooded with multifarious electronic interfaces before the pandemic, e.g., the overflow of screens of TVs, computers, phones and outdoor LED advertisements, which are traditionally referred to as visual display terminals (VDTs). The ubiquity of VDT aggravates the status of visual fatigue (i.e., VF, aka eye fatigue, strain or discomfort), as VF is often triggered after two hours of screen tasks [1]. As one of the most common vision disorders, VF is defined as a group of subjective symptoms encountered in visual tasks, such as eye dryness, pain, blurred vision, even headache and lack of concentration [2]. Apart from degrading our visual performance in work and living quality, VF may also lead to some sub-health or ill states, such as ophthalmic diseases like asthenopia [3], and mental problems, like anxiety, insomnia and depression [4]. Therefore, research focusing on detecting and relieving VF has attracted increasing attention in the fields of medicine and engineering along with the "screening" trend of our life [5], which is essential for promoting the vision hygiene of online courses and telecommuting [6].

Yong Feng is with Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China.

Fei Chen is with University Key Laboratory of Advanced Wireless Communications of Guangdong Province, Department of Electrical and Electronic Engineering, Southern University of Science and Technology, 1088 Xueyuan Avenue, Shenzhen 518055, China (phone: +86-755-88018582; e-mail: fchen@sustech.edu.cn).

Currently, the measurement of VF can be mainly categorized into two types: subjective and objective tests [7]. The objective methods are popular now with the advances of related equipment, which are generally based on physiological measurements, such as electroencephalograph (EEG), electrocardiograph (ECG), eye-tracking and clinical visual function tests [8]. Generally speaking, objective assessments are more reliable and accurate than subjective ones. However, these techniques are initially explored in recent years and far from practical applications, since they often require special devices that are either unavailable or cumbersome to ordinary users. Even worse, most of them are intrusive and restricted to controlled conditions in laboratories [5].

On the contrary, the subjective tests for assessing VF have a long history, which originated from the questionnaire surveys in the early research since the 1970s [9]. Generally in these tests, the subjects are required to select one of several options of varying degrees or quantities for several questions, based on their feelings at the moment or in a past period. For example, typical options are five levels (i.e., none, a little, medium, strong, severe) associated with scores 1-5 for several questions related to eye drying, tearing, pain, etc., and the overall score of the subjective scale is computed accumulatively for evaluating VF in a quantitative way.

For the past half century, various subjective scales emerged in either classic (usually simple combination of symptoms) or professional (mainly tailored and complicated) paths and some of them have been even adopted by clinicians [7]. Compared with the recently developed objective methods, the subjective approaches are simple, mature and quantitative, and have no interference with VF results and no need for professional equipment and expertise, thus being widely employed in research and clinic. However, it has challenging problems in standards and accuracy [7, 10-11].

Aiming at the accuracy improvement of classic scales for practical integration in this study, it is worth noting that almost all the typical classic scales for VF are Likert scales to our knowledge [7, 12-13]. That is, the total score is obtained accumulatively with equal weights. However, there are actually a few professional scales in adjacent areas of fatigue that are weighted, such as the SSQ (simulator sickness questionnaire) [14]. This motivated the present work to establish weighted scales for VF measurement in different screening tasks, aiming to improve their accuracy and increase the adaption for different scenarios, compared to the widely used Likert scales. Moreover, a probe into simpler as well as effective scales for fast detection also falls into the scope of this investigation.

II. METHOD

A. Experimental Design

The experiment was conducted with a subjective scale and an eye-tracker, which were measured several times for each subject. During the measurement, VF was artificially stimulated by visual tasks. The experimental materials involved a VF task program (VF trigger) with several pictures and movie clips, a subjective scale with many eye discomfort symptoms and an eye-tracking VF monitoring program using searching and gazing tasks.

Specifically, the VF trigger contained two fundamental scenarios: searching and watching, which were simulated with a static random digit search task (see a sample picture in Fig. 1) and a movie watching task in this experiment, respectively. The movie material was clipped from the animation Zootopia by Disney with 5 mins × 5 continuous pieces (2D video with 720P).

The questionnaire employed in this study had a collection of 10 subjective symptoms (as shown in Table I), disassembled from classic VF scales, such as the Heuer scale [12], Hayes's scale [15], our previous self-designed VF scale [8] and the Karolinska sleepiness scale (KSS), which had also been reported to work in VF detecting recently [16].

The eye-tracking program was similar to the previous one [8], where two parameters fixation frequency (FF) and saccade amplitude (SA) were measured in the static random digit search and fixed-point gaze tasks. Note that FF is the frequency of fixation events in the task and SA is the distance between two fixations. These two indicators have been validated to be effective in detecting VF previously [8].

6 5 4 5 0 5 7 8 2 4 4 7 5 4 3 2 5 4 4 1
6 5 3 1 2 6 4 6 6 5 6 2 9 6 2 3 5 4 1 9
8 3 7 3 0 2 4 8 7 8 5 8 9 2 2 6 2 8 8 1
8 1 6 5 8 8 8 3 4 1 2 2 6 3 1 9 6 2 5 0
0 1 8 2 9 8 1 1 0 5 6 1 3 8 1 1 1 3 7 1
2 2 5 8 1 0 6 6 7 1 3 3 3 2 5 7 0 2 4 4
3 0 3 2 3 7 5 5 0 7 7 8 0 5 3 2 2 0 5 6
7 7 1 4 3 6 7 7 7 8 6 2 7 0 8 0 5 5 2 6
7 5 5 6 9 1 2 8 1 4 8 4 3 0 2 5 8 6 0 7
4 3 1 8 2 7 7 4 1 9 0 5 0 5 5 5 1 3 7 4
3 2 6 6 8 2 2 1 6 1 6 0 7 7 2 3 1 4 0 5
0 9 2 6 2 5 1 2 5 6 4 7 7 3 9 8 3 1 3 6
1 5 6 1 2 0 7 6 4 5 1 0 2 7 4 1 1 8 5 8
0 0 7 4 9 4 2 4 3 5 2 1 1 4 8 3 8 3 1 4
4 6 1 8 1 3 7 3 1 6 8 7 6 4 4 7 8 9 6 6
5 6 6 6 4 3 5 3 9 1 7 5 7 1 7 5 7 1 7 3

Figure 1. A sample picture of random digit search with a 16 rows × 20 columns matrix with digits 0-9 randomly displayed at each cell.

B. Subjects

Ten participants were recruited in this experiment, who were 5 male and 5 female students aged 18-24 years (M = 20.6 years, SD = 1.7 years) from the Southern University of Science and Technology. All subjects claimed to be physically healthy (no ocular or neurological disorders) with corrected or normal visual acuity ⩾ 1.0. They all had adequate night-sleep and nap time, and took no stimulants, such as alcohol, coffee, or even tea, in 24 hours before the tests. All subjects signed in the informed consent form and received informative instructions before the tests, and were paid with remuneration after the tests. In addition, the experiment protocol was approved by the ethics committee of the university.

TABLE I. SUBJECTIVE SCALE IN THE EXPERIMENT

Symptoms Scores	1'	2'	3'	4'	5'
General Visual Fatigue					
Blinking					
Dry Eyes		(1')	None		
Itchy Eyes		(2')	A little		
Hot Eyes		(3')	Obvious		
Foreign Body Sensation		(4')	Strong		
Blur		(5')	Severe		
Tearing					
Eye Pain					
Sleepy					

C. Apparatus

The eye-tracking parameters were measured with the EyeLink 1000 Plus desktop eye tracker with a 1000 Hz sampling rate. The whole experiment was conducted in a soundproof under the condition of temperature 20-24℃, humidity 30%-45% and illuminance 150 lx. There was a display screen on the desk (19 inches, 1280 × 1024 pixel resolution and 60 Hz refresh rate) for digit searching and movie watching as well as an eye-tracking program. All subjects fixed their heads on the adjustable chin rest, and the perpendicular distance between eyes and the center of the screen was always kept at 45-55 cm during the eye-tracking activity.

D. Procedure

The experiment had two separate sub-experiments for each subject corresponding to the two tasks, as shown in Fig. 2, i.e., the digit search task ("search task") and the movie watching task ("watch task"). Each sub-experiment followed the "Relaxation -Test-Task-Test-Task-...-Test" paradigm with 6 tests and 5 tasks in total. And in each test part, the subjects were asked to fill in the subjective scale in 1 min and then perform the eye-tracking test in 2 mins. Each subject participated in each sub-experiment in the same period of two different days to achieve equal states before experiments. In addition, every subject had an eye relaxation for at least 20 mins before the sub-experiments.

E. Data Analysis

The data was processed by SR Research EyeLink DataViewer 3.1 and IBM SPSS 26 Statistics software.

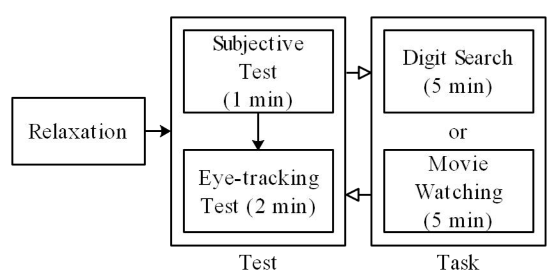

Figure 2. Flowchart of the experiment.

III. RESULTS AND DISCUSSION

A. Comparison of Subjective and Eye-tracking Measurements

The average subjective scores (SSs) by the simple and effective scale in [8] (the same below) of all subjects in two sub-experiments for every 5 mins were analyzed, respectively. As shown in Fig. 3, the difference of SSs (i.e., SSs after 5, 10, 15, 20, 25 mins minus that of the initial 0 min) in the search task rises rapidly in the first 10 mins and reaches a relatively high and steady state later (probably in medium to severe VF), while the data for movie watching almost keep unchanged with the state of the beginning (likely in mild to medium VF). They are significantly different in a paired t-test ($p<0.05$), which suggests that the subjective VF in the search task is much higher than that in the watch task, in a short period of 25 mins.

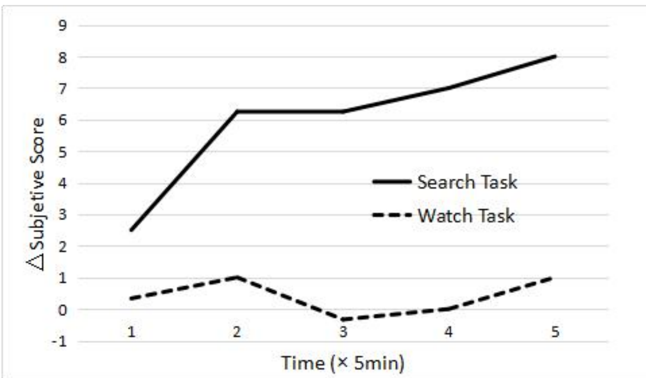

Figure 3. Results of differences of subject scores in the search and watch tasks, measured every 5 mins.

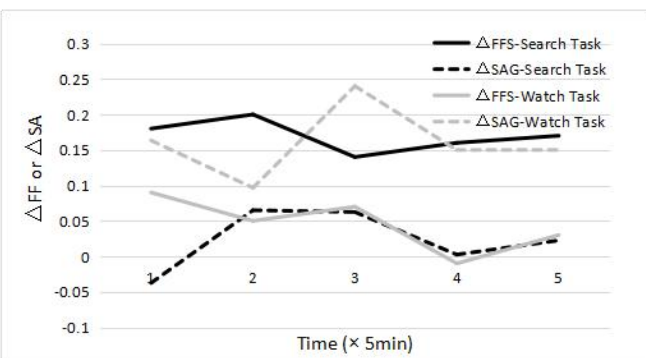

Figure 4. Results of differences of FF and SA in eye-tracking tests, measured every 5 mins. FFS means FF detected with search task in the eye-tracking program while SAG refers to SA related to the gazing task in the program [8].

For the eye-tracking test, Figure 4 shows that the differences of average FFS (i.e., FF measured by the random digit search tasks inside the EyeLink program. ΔFFS denotes FFS after 5, 10, 15, 20, 25 mins minus that of the initial 0 min) and SAG (SA measured by the gazing tasks in the program) [8] are changing along with time in two sub-experiments. FFS for the search task increases in the testing period (4.71% on average) while SAG for the watch task increases more significantly (21.33% on average) with huge fluctuations (13.33%~32.00%), both of which suggest more fatigue than the initial state. On the contrary, FFS for the watch task has a slight downward trend after 5 mins while SAG for the search task just fluctuates around 0. Although conflicting, the change of SSs is quite different from that of eye-tracking parameters FF or SA for the search task, which obviously burdens the eye in the experiment according to the general VF feeling item of the scale (e.g., $p = 0.23$ between SS and FF in the search task). This suggests that the traditional classic Likert visual scales might be improved with the aid of the advantage of accuracy from objective approaches such as eye-tracking tests.

B. Factor Analysis of Subjective Scale

For the benefit of practical integration and frequent monitoring, a questionnaire with 10 items would bring too much disturbance to users as well as cost too much time. Thus, a factor analysis with varimax rotation was performed to determine which items trended to cluster together and accordingly simplified the scale with less key items left, which would be given weights later. Before the analysis, some abnormal results such as "Sleepy" and "Tearing" were eliminated at first, since the corresponding values were almost kept unchanged in this experiment. Moreover, the "General VF" item was intended as a checking option in the scale and removed for the following procedure.

For the search task in our experiment settings, factor analysis for the remaining 7 items disclosed that there were 3 subscales, which might be interpreted as Type I symptoms (3 questions on "Blinking," "Dry Eyes" and "Hot Eyes") (T1) that seemed mild and often occurred among the group of VDT workers, Type II symptoms (3 questions on "Iitchy Eyes", "Blur" and "Foreign body sensation") (T2) which might suggest medium fatigue, and Type III symptom (Only 1 question on "Eye Pain") (T3) that might be severe in eye fatigue and related to mental fatigue, as shown in Fig. 5 and Table II. In the meanwhile, the factor analysis for the watch task was similar but the "Itchy Eyes" item was categorized into Type I.

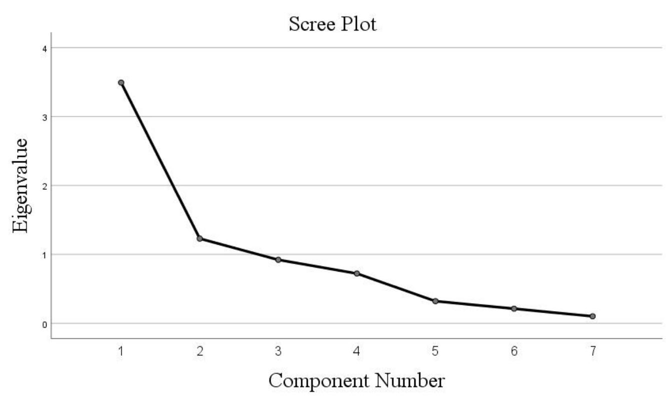

Figure 5. Scree plot of factor analysis on 7 items in the search task.

C. Linear Regressions for the Weighted Scale

Finally, the linear regressions were conducted to obtain weights for the two tasks, with three independents T1, T2, T3 set to be the rounded average values of their corresponding items, and the dependent variable chosen to be the more effective eye-tracking parameter, i.e., FFS for the search task and SAG for the movie watching. Thus, we obtained $FFS = 4.826 - 0.385\times T2 - 0.673\times T3$ for the search task, where T1 was excluded, although not all the hypotheses were satisfied for multiple linear regression. Similarly, we also had SAG = 1.261 − 0.448×T2 − 0.035×T3 for the movie task, where T1 was also excluded.

TABLE II. ROTATED COMPONENT MATRIX OF FACTOR ANALYSIS IN THE SEARCH TASK

Rotated Component Matrix			
	Component		
	1	2	3
Foreign Body Sensation	0.937	0.204	0.121
Blur	0.896	0.236	-0.117
Itchy Eyes	0.862	0.175	0.281
Blinking	0.277	0.864	-0.008
Dry Eyes	0.220	0.763	0.194
Hot Eyes	0.078	0.714	0.097
Eye Pain	0.112	0.165	0.963

* Three items were removed, i.e., "Tearing", "Sleepy" and "General Visual Fatigue".

As shown by the above two tentatively exploratory formulas, T3 is associated with a greater weight compared with T2 in the search tasks, suggesting that Type III, which is mainly involved with symptoms of severe fatigue, plays a more important role in the subjective evaluation of medium to severe VF. Contrarily, T2 (Type II symptoms, generally stand for medium fatigue) is supposed to be crucial while assessing slight to medium VF in the watch task, and thus has a greater weight. As for T1, it is excluded in both regressions, which might indicate that those symptoms, such as blinking and dry eye problems (Type I), are too common and too weak to trigger the VF states, especially under the conditions that Type II or II symptoms scores are higher than one. Therefore, different weighted scales are appropriate for scenarios with VF of different degrees.

The weighted scales tailored to the two common screen tasks following the above analysis are expected to increase the accuracy and flexibility of subjective VF measurement as well as reduce testing time for frequent detection. However, there are some limitations in this pilot study apart from the small sample size. Firstly and most importantly, the specific relationships of subjective scores with eye-tracking parameters (or perhaps all the objective indicators) have not been revealed quantitatively yet (very likely to be nonlinear). Moreover, the weights or coefficients in the two exploratory equations may vary with sample size, subjective preferences, specific tasks and ocular physiological factors (e.g. age, gender, diseases, etc.). Therefore, future work may concentrate on increasing the sample size to obtain stable empirical weights for more scenarios and different groups of participants, and verifying the reliability and validity of the scales for practical applications.

IV. CONCLUSION

Aiming to improve the accuracy of subjective testing in visual fatigue and promote its adaptive integration in practical devices, this work established two weighted subjective scales that were simpler, and more suitable for scenarios with visual fatigue of varying degrees. A questionnaire of collected items from typical classic Likert scales was performed along with the eye-tracking tests in the searching and watching tasks. The subjective scores were simplified to three items only by factor analysis, and two different weighting solutions were finally obtained by linear regressions, which were set with three items as independents and previously validated eye-tracking parameters of fixation frequency or saccade amplitude as dependent in different tasks. The tailored weighted scales showed potential in the subjective measurement of visual fatigue with higher accuracy and scenario adaption as well as simpler and faster monitoring.

REFERENCES

[1] American Optometric Association, "Computer vision syndrome," https://www.aoa.org/patients-and-public/caring-for-your-vision/protecting-your-vision/computer-vision-syndrome, accessed on 2020/11/5.

[2] C. Blehm, S. Vishnu, A. Khattak, S. Mitra and R. W. Ye, "Computer vision syndrome: a review," *Surv. Ophthalmol.*, vol. 50, no. 3, pp. 253–262, May–Jun. 2005.

[3] H. S. Thorud, M. Helland, A. Aarås, T. M. Kvikstad, L.G. Lindberg and G. Horgen, "Eye-related pain induced by visually demanding computer work," *Optom. Vis. Sci.*, vol. 89, no. 4, pp. E452–E464, Apr. 2012.

[4] C. C. Han, R. Liu, R. Liu, Z. Zhu, R. Yu and L. Ma, "Prevalence of asthenopia and its risk factors in Chinese college students," *Int. J. Ophthalmol.*, vol. 6, no. 5, pp. 718–722, Oct. 2013.

[5] A. L. Sheppard and J. S. Wolffsohn, "Digital eye strain: Prevalence, measurement and amelioration," *BMJ Open Ophthalmol.*, vol. 3, no. 1, pp. e000146, Apr. 2018.

[6] S. Bhattacharya, S. M. Saleem and A. Singh, "Digital eye strain in the era of COVID-19 pandemic: An emerging public health threat," *Indian J. Ophthalmol.,* vol. 68, no. 8, pp. 1709-1710, Aug. 2020.

[7] Y. Feng, S.F. Zhu and F. Chen, "A review on the development of subjective evaluation of visual fatigue," *Chinese J. Biomed. Eng.*, to be published.

[8] Y. Feng, L. Wang and F. Chen, "An eye-tracking based evaluation on the effect of far-infrared therapy for relieving visual fatigue," in *2019 41th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC*, Berlin, 2019, pp. 313–316.

[9] S.Y. Liang, C. Y. Yang and J. J. Li, "Experimental scheme exploration on visual fatigue in multimedia classroom light environment," *Chin. Illum. Eng. J.*, vol. 30, no. 3, pp. 43–47, Mar. 2019.

[10] Y. H. Lin, C. Y. Chen, S. Y. Lu and Y. C. Lin, "Visual fatigue during VDT work: effects of time-based and environment-based conditions," Displays, vol. 29, no. 5, pp. 487–492, Dec. 2008.

[11] Y.H. Zhang, H.B. Yang, Y.F. Xu and L. Feng, "Comparison of visual comfort and fatigue between watching different types of 3D TVs as measured by eye tracking," in *Int. Conf. Appl. Human Factors Ergon*, Los Angeles, 2017, pp. 175-186.

[12] H. Heuer, G. Hollendiek, H. Kröger and T. Römer, "Rest position of the eyes and its effect on viewing distance and visual fatigue in computer display work," *Zeitschrift fur Experimentelle und Angewandte Psychologie*, vol. 36, no. 4, pp. 538–566, Jan. 1989.

[13] J. E. Sheedy, J. N. Hayes and J. Engle, "Is all asthenopia the same?" *Optom. Vis. Sci.*, vol. 80, no. 11, pp. 732–739, Nov. 2003.

[14] R.S. Kennedy, N.E. Lane, K.S. Berbaum and G.L. Michael, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *Int. J. Aviat. Psychol*, vol. 3, no. 3, pp. 203–220, Nov. 2009.

[15] J.R. Hayes, J.E. Sheedy, J.A. Stelmack, and C.A. Heaney, "Computer use, symptoms, and quality of life," *Optom. Vis. Sci.*, vol. 84, no. 8, pp. E738–E755, Aug. 2007.