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

Yellow (Lens) Better: Bioelectrical and Biometrical Measures to Assess Arousing and Focusing Effects

R. Laureanti*, *Member, IEEE*, M. Bilucaglia*, *Member, IEEE*, M. Zito, R. Circi, A. Fici, F. Rivetti, R. Valesi, S. Wahl, L. T. Mainardi, *Member, IEEE* and V. Russo

Abstract— **Colours can induce several psychological effects, conditioning perceptions, cognitive/emotional states and human performances. In this exploratory study we investigated the effect of a yellow light exposure, obtained filtering the ambient light with coloured glasses, on the human's psychological functioning. In particular we wanted to assess if people are more able to focus when exposed to a yellow light. We recorded EEG, SC, HR and gaze-related data from 16 subjects (50% split in experimental and control group) during the execution of a reactivity test (the Hazard Perception Test, HPT). Compared with the control group, the experimental group showed increases in concentration, focus, visual attention and arousal, as measured by increases of first fixation duration and Beta over-Alpha ratio (BAR) as well as by decreases of distraction, workload, and number of gaze revisits.**

I. INTRODUCTION

The effect of colours on human psychology, performances and emotions have been demonstrated in literature [1].

In order to assess the affective and cognitive changes due to colours, different techniques, based on either a direct or an indirect approach, have been proposed. The direct approach is based on bioelectrical measures, such as Electroencephalography (EEG), Skin Conductance (SC) and Heart Rate (HR), while the indirect approach is based on psychometric tools, such as associations between colours and emotional terms (e.g., adjectives), semantic differentials and personal rating scales [2].

In a previous paper, we investigated the effect of the blue colour, showing its "wake-up" property [3]. In the present work we assessed the properties of the yellow. We produced a "yellow light perception" filtering the ambient light with a pair of custom-made glasses, as an effective alternative to direct light exposure [4].

By a psychological point of view, only few studies so far tested (directly or indirectly) the cognitive and emotional effects of the yellow light, however, ambivalent results were reported. Yellow colour showed a focusing effect, measured as a decrease in reaction times, an enhance in concentration [5] and an increase in reading speed [6]. An arousing effect, measured by SC (but not by HR), was also reported [7]. By contrast, no effects in modulating emotional responses to affective pictures, as measured by late positive potentials (EEG recordings), were found [8].

In this exploratory study we investigated the physiological responses to the exposure of a yellow-filtered light during the execution of a reactivity test. We selected several EEG metrics to assess cognitive states [9], cognitive workload [9] and arousal [10]. Additionally, HR, Skin Conductance Response (SCR) and Skin Conductance Level (SCL) were used to measure arousal [7]. Finally, fixation related measures allowed the assessment of visual attention [11].

II. METHODS

A. Instrumentation

EEG data were recorded using a B-Alert ×10 (ABM, Inc.), a portable wireless EEG headset with a sample frequency of 256 Hz and a resolution of 12 bits. The device has 9 Ag/AgCl electrodes, located at the F3, Fz, F4, C3, Cz, C4, P3, POz, P4 sites of the 10–20 system. Two Ag/AgCl adhesive patches, placed on the mastoids (M1, M2), were used as ground and reference.

EEG data were collected using iMotions (iMotions A/S, Copenhagen, Denmark) software, an integrated research platform that supports study design, stimuli presentation and real-time synchronization of various devices. Data stream from B-Alert to iMotions, established by means of a Software Development Kit (SDK), consisted of raw EEG data, as well as several cognitive metrics.

The SC signal was recorded placing 2 Ag/AgCl electrodes on the index and ring finger from the non-dominant hand.

In order to have an accurate and minimally invasive HR measure [12], we recorded the Blood Volume Pulse (BVP) signal using a photopletismographic sensor placed on the middle finger from the same hand.

BVP and SC sensors were connected to the FlexComp System (Thought Technology, Inc.), a 10 channels general-purpose bio-signal acquisition device. The sample frequency was set at 256 Hz and the resolution at 14 bits. FlexComp data were collected using BioGraph Infiniti (Thought Technology, Inc.) software.

In order to ensure a synchronization between iMotions and BioGraph Infinity data, a photosensor connected to the FlexComp device was attached to the stimulus monitor to discriminate between high and low luminance level. A synchronization sequence, consisting of alternated black (B)

*These authors have contributed equally to the work.
M. Bilucaglia, A. Fici, F. Rivetti, R. Circi, R. Valesi are with the Behavior and Brain Lab IULM – Neuromarketing Research Centre,Università IULM, Milan, Italy (corresponding author: marco.bilucaglia@iulm.it).
R. Laureanti and L. T. Mainardi are with the Department of Electronics, Information and Bioengineering (DEIB), Politecnico di Milano, Milan, Italy.
M. Zito and V. Russo are with the Department of Business, Law, Economics and Consumer Behaviour "Carlo A. Ricciardi" and Behavior and Brain Lab IULM – Neuromarketing Research Centre, Università IULM, Milan, Italy.
S. Wahl is with the ZEISS Vision Science Lab, Institute for Ophthalmic Research, University Tuebingen, Tuebingen, Germany and Carl Zeiss Vision International GmbH.

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

and white (W) patterns (500 ms each, in the order W-B-W-B-W-B), was showed at the beginning of the experiment.

Finally, gaze data were recorded through iMotions with a ProSpectrum ×60 (Tobii, LLC), a 150 Hz EyeTracker bar with 30° max gaze angle, 0.3° of accuracy and 0.06° RMS of precision.

B. Study Population and Experimental Protocol

We used a pair of custom-made glasses in order to filter the ambient light and produce a "yellow light" perception, as an alternative to direct light exposure. Yellow lenses specifications were: 2 mm thickness, 50% of visible light transmission and cut-off wavelength at 610 nm [8].

The experiment involved 16 healthy subjects (8 men), with age ranging from 30 to 51 years ($M = 40.2,\ SD = 5.3$). We randomly assigned the subjects to 2 balanced groups: 8 wearing the lenses (L group) and 8 not wearing any lenses (nL group) as a control group.

The study protocol was approved by the ethical committee of Università IULM and informed written consent was obtained from each participant before starting the experiment.

Each subject sat on a chair placed in front of a 23.8 inches monitor (FlexScan EV2451, EIZO KK) located in a 7 m × 3 m experimental room, artificially lit by florescence lights and in absence of any natural light. The experimenter positioned SC and EEG sensors and checked the quality of the signals before starting the recording. The contact impedance of the EEG sensors was measured and ensured to be less than 10 $k\Omega$ [13].

After the synchronization sequence, the subject performed the B-Alert Benchmark in order to extract the cognitive metrics (see the "B-Alert Metrics" paragraph). A 60 s neutral stimulus was then presented to record the subject baseline level of activity. A 60 s eye-closed baseline (EYC) was administrated in order to estimate the individual alpha frequency (IAF). After EYC, L group wore the lenses.

Finally, subjects performed a web-based version (available at `http://hazardperceptiontest.net`) of the Hazard Perception Test (HPT), to measure their reactivity during a driving scenario. HPT consists of videos of real driving scenes recorded from the driver's point of view in which another road user creates a potential traffic conflict (e.g., a dangerous and accident-prone situation). Once the offending road user appeared on the scene, the participant was asked to press the keyboard's spacebar as quick as possible [5]. The HPT consisted in 3 videos with a similar assignment (namely Task 1, Task 2 and Task 3) presented in a randomized order. At the end, each Task was automatically scored according to how fast the subject recognized the hazards.

B-Alert metrics

B-Alert SDK provides several cognitive metrics, related to discrete cognitive states and and cognitive workload. Cognitive state metrics include High and Low Engagement, Distraction and Drowsiness, while cognitive workload metrics include Workload BDS, Workload FBDS and Workload Average.

Engagement is a cognitive state linked to vigilance which include arousal, intrinsic motivation, motivation for success and concentration. The related metrics reflect the engagement level (high or low). The Distraction metric assesses the condition of being side-tracked during a cognitive task, while Drowsiness measures the somnolence level, generally associated to sleep deprivation [14]. Cognitive state metrics are built on a generalized classification model (previously validated on a large population) and individualized (i.e., adapted to the specific subject) by means of a benchmark. In particular, the metrics express the posterior class-probability of the 4 classes [9].

Cognitive workload is related to the mental effort required to complete a cognitive task and expresses the relationship between the resources required to carry out the task and the resources available to, and hence supplied by, the operator [14]. Cognitive workload metrics derive from 2 generalized classification models previously built on 2 different tasks (namely the forward and backward digit span, FBDS and the backward digit span, BDS). Workload Average is the mean probability between Workload BDS and Workload FBDS [9].

Cognitive state and cognitive workload metrics are provided as time signals with a temporal resolution of 1 s.

EEG processing

The EEG signal was processed using Matlab (The Mathworks, Inc.) and the EEGLab toolbox [15].

Raw data were band-pass filtered between 0.1 and 30 Hz and notch filtered on both 50 and 100 Hz in order to remove the line noise. Then, an Independent Component Analysis was performed. Independent Components (ICs) were automatically labelled using ICLabel in terms of posterior class-probability of being originated from a particular source, such as "brain", "eye-movement" or "EMG" [16]. ICs with a "brain" class probability lower than 70% were marked as "artifacts" and removed. Not-marked ICs were thus back projected to the original EEG space. Finally, EEG data were re-referenced to the common average.

The Individual Alpha Frequency (IAF) was estimated as the peak frequency of the Power Spectral Density in the extended alpha band ($7 \sim 12\ Hz$) during the EYC baseline. Alpha (α) and beta (β) bands were defined from the estimated IAF as, respectively, $\alpha = [IAF - 2; IAF + 2]$ and $\beta = [IAF + 2; IAF + 16]$ [17].

As an estimation of the arousal, we used the Beta-over-Alpha Ratio (BAR), defined as the ratio between the average of β-filtered and α-filtered channels [18].

SC and BVP processing

The SC signal was band-pass filtered using a zero-phase 4th order FIR filter ($0.001 \sim 0.35\ Hz$); then, a threshold for SC extreme values ($0.05 \sim 60\ \mu S$) and extreme rate of changes ($\pm 8\ \mu S/s$) was used in order to detect artifacts [19]. The artefactual points were replaced by a linear interpolation using adjacent points. From artifact-corrected SC, both the tonic SCL and the phasic SCR were extracted by means of the cvxEDA algorithm [20].

BVP signal was low pass filtered using a zero-phase 2nd order Butterworth filter; then, all peaks were identified using the Pan-Tompkins algorithm [21] and the instant HR was computed from the inverse of the peak-to-peak distance. Finally, the HR signal was linearly interpolated and filtered with a 2 s moving average filter in order to obtain a smoother signal.

Baseline normalization

The SC indices, the HR and BAR signals were epoched according to each HPT Task and z-score transformed with respect to the baseline epoch, according to the Equation:

$$x_z(t) = (x(t) - m_B)/s_B, \tag{1}$$

where $x_z(t)$ is the z-transformed signal, $x(t)$ is the original signal, m_B and s_B are, respectively, its temporal mean and standard deviation computed inside the baseline epoch.

Gaze data processing

Gaze data represent the temporal variations of gaze's position. In this study, gaze data were extracted from dynamic areas of interests (AOI), manually drawn and fitted frame-by-frame on the hazardous object. Only fixation-related metrics were considered, as they have been previously related to visual attention and perception [11].

From each AOI the following metrics (all expressed in milliseconds, except for revisit that is unitless) were extracted: first fixation duration (FFD), average fixation duration, time spent (the total fixation duration on the AOI), revisits (the number of outer fixations that return back on the AOI).

The gaze metrics of one subject were not considered due to a calibration issue. The final sample consisted, thus, of 15 subjects (7 L, 8 nL).

Statistical analysis

Statistical analyses were performed using Matlab. For every subject, all signals were time averaged across the tasks. Averaged signals, fixation-related metrics and the scores associated with each task were grouped for condition.

We set the significance level to $\alpha = 0.10$, as common practice in exploratory studies with small samples [22].

For each variable, the Kolmogorov-Smirnov test for normality was applied (with a significance level $\alpha = 0.05$): based on its results, the Wilcoxon signed-rank test or the two-samples t-test was used.

Additionally, the effect size was estimated by means of the Cohen's d, defined as [23]:

$$d = (m_1 - m_2)/[(s_1 + s_2)/2], \tag{2}$$

where m_1, m_2 are the means and s_1, s_2 are standard deviations of, respectively, L and nL groups. Large, medium, and small effect sizes are commonly placed at values of 0.8, 0.5, and 0.2, respectively. The significance of the effect size was assessed through its bootstrapped 95% confidence intervals, following the percentile approach [24].

In evaluating the results, we focused on the effect size, rather than on the p-value: if a metric showed a not-significant effect size, we labelled it as "not significant", independently from the p-value of the test. This is a common practice for studies with a low sample size, where the p-value should be interpreted tentatively at best [25].

III. RESULTS

SC related metrics and HR, as well as HPT scores, did not show any significant group difference in any task.

EEG and gaze-related metrics showed significant group differences, both in terms of a large and significant effect size and, in some cases, a p-value lower that 0.1, as summarized in Table I.

In Task 1, the L group showed a significant (in both effect size and p-value) lower Distraction, higher Drowsiness and lower Workload of all types. In Task 2, the L group showed significant (in both effect size and p-value) higher Drowsiness and lower Workload FBDS, as well as a higher BAR. Additionally, Distraction, Workload Average and Workload BDS were significantly (effect size only) lower for L group. In Task 3 L group showed a significant (effect size only) higher BAR and lower Distraction.

Regarding the gaze data, in Task 1, the L group showed a significant (in both the effect size and p-value) greater FFD and lower number of revisits. In Task 2 no significant group difference was found, while in Task 3, the L group showed a significant (effect size only) shorter time spent.

TABLE I

SIGNIFICANT DIFFERENCES BETWEEN GROUPS IN ALL TASKS.

	L	**nL**	**p-value**	**Cohen's d**
		Task 1		
Distraction	0.02 (0.02)	0.12 (0.10)	0.065*	-1.624§
Drowsiness	0.04 (0.03)	0.00 (0.01)	0.009*	1.931§
W AVG	0.41 (0.21)	0.59 (0.11)	0.050*	-1.144§
W BDS	0.39 (0.21)	0.56 (0.11)	0.050*	-1.053§
W FBDS	0.42 (0.21)	0.61 (0.11)	0.021*	-1.224§
Revisits	2.86 (2.27)	5.38 (1.92)	0.058*	-1.202§
FFD	512 (51)	297 (111)	0.001*	2.663§
		Task 2		
Distraction	0.02 (0.02)	0.10 (0.10)	0.195	-1.282§
Drowsiness	0.08 (0.05)	0.00 (0.00)	<0.001*	2.807§
W AVG	0.43 (0.21)	0.59 (0.12)	0.105	-0.916§
W BDS	0.42 (0.21)	0.56 (0.12)	0.195	-0.798§
W FBDS	0.44 (0.21)	0.61 (0.12)	0.065*	-1.018§
BAR	0.40 (0.45)	-0.00 (0.14)	0.045*	1.360§
		Task 3		
Distraction	0.02 (0.02)	0.10 (0.09)	0.161	-1.390§
BAR	0.52 (0.55)	0.09 (0.23)	0.105	1.096§
Time Spent	2676 (1820)	4032 (933)	0.161	-0.984§

Values are given as average (std). * marks a significance at test level (t-test or signed-rank test), while the symbol § indicates a significance for the Cohen's d estimate. W stands for Workload; AVG stands for Average.

IV. DISCUSSION

In this exploratory study we investigated the yellow lenses' effect in increasing attention, focus, arousal, using both indirect (HPT scores) and direct (physiological data) methods. In evaluating the results, we focused on the significance of the effect size, rather than on the test's p-value, as suggested for studies with a small sample size [25].

Similarly to [5], [6], we found in all the Tasks an increase in concentration for L group, measured by a decrease of the distraction values, as well as a decrease in workload for the L group in Tasks 1 and 2 only.

In line with [7], L group showed an arousing effect (increase of BAR values) in Tasks 2 and 3 only.

Compared with the nL group, L group showed an increase in FFD and a decrease in number of Revisits in Task 1, suggesting an increase of visual attention [11]. Furthermore, in Task 3 the total time spent on the AOI was lower in the L group, suggesting a higher "hazard detection efficiency".

L group's increase of drowsiness in Tasks 1 and 2 seems inconsistent with the other results, but this could be related to a limit of the classification model used for this metric. In fact, the drowsiness effectively tracks performance decrements associated with sleep deprivation, while under-performs for rested (i.e., not sleep deprived) subjects [9].

These results suggested focusing and arousing properties both at gaze and bioelectrical (EEG metrics) level. The magnitude of this "yellow lens effect" was large, as confirmed by the Cohen's d estimates of EEG-related metrics, but probably not large enough to produce measurable changes at indirect level: in fact, no group differences in HPT scores were found in any task, showing how direct measures seem to be more sensible to quantify cognitive state changes [26].

Overall, the results showed some variability among the different tasks. This could be related to the intrinsic task's differences, but this should be confirmed by a within group analysis (e.g. repeated measure ANOVA). Being an exploratory study, the results need to be further investigated by a future confirmatory study, using both a larger sample and more conservative statistical methods.

In conclusion, yellow lenses seem to be a promising useful tool to improve human attention in critical focusing-demanding circumstances, such as job/sport activities or driving.

ACKNOWLEDGMENT

The authors would like to thank Carl Zeiss Vision Italia S.p.A. for the given support and for providing the lenses.

REFERENCES

[1] A. J. Elliot and M. A. Maier, "Color psychology: Effects of perceiving color on psychological functioning in humans," *Annual Review of Psychology*, vol. 65, pp. 95–120, 2014.

[2] L. Wilms and D. Oberfeld, "Color and emotion: effects of hue, saturation, and brightness," *Psychological Research*, vol. 82, pp. 896–914, 2018.

[3] M. Bilucaglia, R. Laureanti, M. Zito, R. Circi, A. Fici, F. Rivetti, R. Valesi, S. Wahl, and V. Russo, "Looking through blue glasses: bioelectrical measures to assess the awakening after a calm situation," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 526–529.

[4] T. Schilling, A. Sipatchin, L. Chuang, and S. Wahl, "Tinted lenses affect our physiological responses to affective pictures: An EEG/ERP study," in *2nd International Neuroergonomics Conference: The brain at work and in everyday life*. Frontiers Research Foundation, 2018.

[5] P. Lacherez, A. K. Saeri, J. M. Wood, D. A. Atchison, and M. S. Horswill, "A yellow filter improves response times to low-contrast targets and traffic hazards," *Optometry and Vision Science*, vol. 90, no. 3, pp. 242–248, 2013.

[6] R. S. Hollingsworth, A. K. Ludlow, A. J. Wilkins, R. I. Calver, and P. M. Allen, "Visual performance and the use of colored filters in children who are deaf." *Optometry and Vision Science*, vol. 92, no. 6, pp. 690–9, 2015.

[7] R. Küller, B. Mikellides, and J. Janssens, "Color, arousal, and performance - a comparison of three experiments," *Color Research and Application*, vol. 34, no. 2, pp. 141–152, 2009.

[8] T. Schilling, A. Sipatchin, L. Chuang, and S. Wahl, "Looking through "rose-tinted" glasses: The influence of tint on visual affective processing," *Frontiers in Human Neuroscience*, vol. 13, p. 187, 2019.

[9] R. R. Johnson, D. P. Popovic, R. E. Olmstead, M. Stikic, D. J. Levendowski, and C. Berka, "Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model," *Biological Psychology*, vol. 87, no. 2, pp. 241–250, 2011.

[10] A. Mert and A. Akan, "Emotion recognition from EEG signals by using multivariate empirical mode decomposition," *Pattern Analysis and Applications*, vol. 21, no. 1, pp. 81–89, 2018.

[11] A. Ciceri, V. Russo, G. Songa, G. Gabrielli, and J. Clement, "A neuroscientific method for assessing effectiveness of digital vs. print ads," *Journal of Advertising Research*, pp. 10.2501/JAR–2019–015, 2019.

[12] V. D. Corino, R. Laureanti, L. Ferranti, G. Scarpini, F. Lombardi, and L. T. Mainardi, "Detection of atrial fibrillation episodes using a wristband device," *Physiological Measurement*, vol. 38, no. 5, pp. 787–799, 2017.

[13] S. R. Sinha, L. Sullivan, D. Sabau, D. San-Juan, K. E. Dombrowski, J. J. Halford, A. J. Hani, F. W. Drislane, and M. M. Stecker, "American clinical neurophysiology society guideline 1: Minimum technical requirements for performing clinical electroencephalography," *Journal of Clinical Neurophysiology*, vol. 33, no. 4, pp. 303–307, 2016.

[14] C. A. Mark, D. V. Poltavski, T. Petros, and A. King, "Differential executive functioning in young adulthood as a function of experienced child abuse," *International Journal of Psychophysiology*, vol. 135, pp. 126–135, 2019.

[15] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.

[16] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "Iclabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, 2019.

[17] G. Borghini, P. Aricò, G. Di Flumeri, N. Sciaraffa, and F. Babiloni, "Correlation and similarity between cerebral and non-cerebral electrical activity for user's states assessment," *Sensors*, vol. 19, no. 3, p. 704, 2019.

[18] M. Bilucaglia, G. M. Duma, G. Mento, L. Semenzato, and P. E. Tressoldi, "Applying machine learning eeg signal classification to emotion-related brain anticipatory activity," *F1000Research*, vol. 9, no. 173, p. 173, 2020.

[19] I. R. Kleckner, R. M. Jones, O. Wilder-Smith, J. B. Wormwood, M. Akcakaya, K. S. Quigley, C. Lord, and M. S. Goodwin, "Simple, transparent, and flexible automated quality assessment procedures for ambulatory electrodermal activity data," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 7, pp. 1460–1467, 2018.

[20] A. Greco, G. Valenza, A. Lanata, E. P. Scilingo, and L. Citi, "Cvxeda: A convex optimization approach to electrodermal activity processing," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 797–804, 2016.

[21] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.

[22] W. R. Schumm, K. K. Pratt, J. L. Hartenstein, B. A. Jenkins, and G. A. Johnson, "Determining statistical significance (alpha) and reporting statistical trends: controversies, issues, and facts," *Comprehensive Psychology*, vol. 2, no. Article 10, pp. 1–10, 2014.

[23] C. O. Fritz, P. E. Morris, J. J. Richler, and O. Fritz, "Effect size estimates: Current use, calculations, and interpretation," *Journal of Experimental Psychology: General*, vol. 141, no. 1, pp. 2–18, 2012.

[24] J. Algina, H. J. Keselman, and R. D. Penfield, "Confidence interval coverage for cohen's effect size statistic," *Educational and Psychological Measurement*, vol. 66, pp. 945–960, 2006.

[25] L. G. Halsey, D. Curran-Everett, S. L. Vowler, and G. B. Drummond, "The fickle p value generates irreproducible results," *Nature Methods*, vol. 12, no. 3, pp. 179–185, 2015.

[26] G. D. Flumeri, G. Borghini, P. Aricò, A. Colosimo, S. Pozzi, S. Bonelli, A. Golfetti, W. Kong, and F. Babiloni, "On the use of cognitive neurometric indexes in aeronautic and air traffic management environments," in *Symbiotic Interaction*, 2015, pp. 45–56.