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

Eye-Fixation-Related Potentials (EFRPs) As a Predictor of Human Error Occurrences During a Visual Inspection Task

Hiroki Watanabe[1], Yuichiro Higashi[2], Takuma Saga[1], Masanori Hashizaki[2], Yusuke Yokota[1], Hirotaka Kataoka[2], Hiroshi Nakajima[2] and Yasushi Naruse[1]

***Abstract*— Estimation of human attentional states using an electroencephalogram (EEG) has been demonstrated to help prevent human errors associated with the degradation. Since the use of the lambda response –one of eye-fixation-related potentials time-locked to the saccade offset– enables such estimation without external triggers, the measurements are compatible for an application in a real-world environment. With aiming to apply the lambda response as an index of human errors during the visual inspection, the current research elucidated whether the mean amplitude of the lambda response was a predictor of the number of inspection errors. EEGs were measured from 50 participants while inspecting the differences between two images of the circuit board. Twenty percent of the total number of image pairs included differences. The lambda response was obtained relative to a saccade offset starting a fixation of the inspection image. Participants conducted four sessions over two days (625 trials/ session, 2 sessions/ day). A Poisson regression of the number of inspection errors using a generalized linear mixed model showed that a coefficient of the mean amplitude of the lambda response was significant ($\hat{\beta} = 0.24$, $p < 0.01$), suggesting that the response has a role in the prediction of the number of human error occurrences in the visual inspection.**

I. INTRODUCTION

Neurophysiological studies have researched how human mental states such as a mental workload or sustained attention are correlated with modulation of electroencephalogram (EEG) [1], [2]. Online measurements of such EEG correlates have been known to help prevent human errors induced by a high mental workload or degraded attention before it happens in an ecologically-valid environment. It, for example, is demonstrated that EEG measurements enable to predict occurrences of misperception of auditory alarms due to an inattention state during piloting an aircraft [3].

For such EEG measurements in the real-world environment, eye-fixation-related potentials (EFRPs) have attracted attention [4]. EFRPs are obtained relative to the offset of saccadic movements (i.e., a time point of staring a fixation of an object). The applications of EFRPs would be beneficial for measurements of mental states in a real-world environment because, to extract EFRPs, external triggers associated with events are not necessary. To date, studies successfully showed the effectiveness of EFRPs as a means of investigating a human cognitive process during free viewing [5]. As a related component to mental state estimation, a positive potential observed at around 80 ms relative to saccade offset in an occipital region –reflecting an afferent inflow of visual information at fixation– is called a lambda response [6], [7]. The amplitude is known to be modulated by an amount of attention to visual objects [8], and thus, can be applied to monitor a mental workload in a real-world environment [9].

The current research aims to extend research on the EFRP-based estimation of a human mental state to the visual inspection task by investigating whether the number of human error occurrences during the task can be predicted from the measurements of the lambda responses. A continuous inspection of the details of the products over a long time requires inspectors to keep sustained attention, and a decrease in the attention is one of the causes of inspection errors such as judgments of defective products as normal ones. To date, whether EFRPs have a role as a predictor of human error occurrences during the visual inspection is, however, less investigated. To this end, we measured EEGs during a visual inspection task over two days (two sessions/ day, four sessions in total). The number of inspection errors in each session was regressed based on a generalized linear mixed model (GLMM) using the mean amplitudes of the lambda responses and a log power in the alpha frequency band at occipital regions –the alpha oscillations are related to a response miss to the target stimuli in a sustained attention task [10]. The current research set to answer the question: Whether a mean amplitude of the lambda response is a significant predictor of the number of inspection errors during a visual inspection?

II. METHODS

A. Participants

The data were collected from 50 participants (25 males, age range: 20 – 39). All participants had normal or corrected-to-normal vision and reported no history of neurological disorders. The current research was approved by the Ethical Committee for Human and Animal Research of the National Institute of Information and Communications Technology. All participants provided written informed consent before participating in the experiment. Data was collected under the ethical standards in the Declaration of Helsinki.

B. Data Collection

EEG data were acquired from FCz, Pz, O1, and O2 positions according to the International 10–20 system using dry electrodes (Unique Medical Co., Ltd., Tokyo, Japan) with

*The current research was funded by OMRON Corporation.

[1] H. Watanabe, T. Saga, Y. Yokota and Y. Naruse are with the National Institute of Information and Communications Technology, Hyogo, Japan. (phone: +81-78-969-2225; e-mail: y_naruse@nict.go.jp)

[2] Y. Higashi, M. Hashizaki, H. Kataoka and H. Nakajima are with OMRON Corporation, Kyoto, Japan.

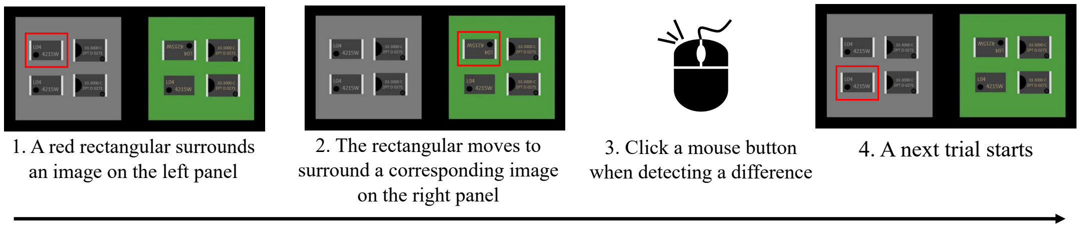

Fig. 1: A schematic figure of an experimental procedure.

a 500 Hz sampling rate. The ground and the reference electrodes were placed at the left and right mastoid, respectively. Electrooculogram (EOG) was monitored above and the side of the lateral canthus of the left eyes, respectively. A portable EEG device (PolymateMini AP108; Miyuki Giken Co., Ltd., Tokyo, Japan) was used for measurements.

Participants performed a visual inspection task to detect a difference between two images of the circuit boards. A procedure of a trial is summarized in Fig. 1. Images of a circuit board were arranged on the left and right panels in the display. First, a red rectangular frame surrounded an upper left image on the left panel for 1,200 ms. Participants memorized the appearance of the image while the rectangular were surrounding the image. Second, the rectangular moved to surround the image in the corresponding location on the right panel for 1,200 ms. Participants moved their gazes to the image and judge whether the image was the same as the one in the left panel. Participants clicked a mouse when detecting a difference between them. They were forbidden to (1) move their gazes before the red rectangular started to move to the right panel and (2) to gaze back to the image on the left panel once the red rectangular moved to the image on the right panel. The same procedure was repeated for all images on the display in the order from top to bottom and from left to right. After all images in the display were inspected, the next images were presented. The number of trials in a session was 625 trials. The order of images was randomized per participant. Participants performed two sessions in a day over two days (2,500 trials in total). The data collection in a single day lasted approximately 1.5 hours including preparation. A short break was inserted between the sessions.

Two types of differences were prepared: an upside-down difference, where the top and bottom of the image are upsidedown, and an edge difference, where both or either paintings of the edge of the images were missing (Fig. 2). Twenty percent of the total number of image pairs were included as an upside-down difference or an edge difference (i.e., 250 trials per difference type in total).

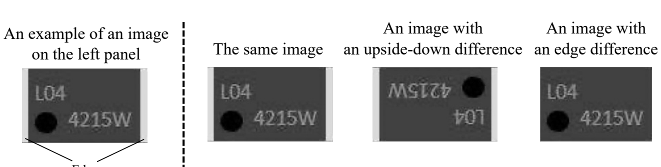

Fig. 2: Examples of different images in the task.

C. EEG data analysis

The preprocessing of raw EEG data was performed using MATLAB (The MathWorks, Inc., U.S.A) and the FieldTrip Toolbox [11] per session. Due to the mechanical troubles, six participants' data on day 1 and five participants' data on day 2 were not recorded. In the two participant data, only the second session on day 1 and day 2 were recorded, respectively. To calculate the mean amplitudes of the lambda response, time points starting to fixate the inspection images were manually determined by visual detection of saccade offset points from the horizontal EOG data. The trials that the saccade offset point was not visually determined were removed from the further analysis. The trials that participants gazed back to the image on the left panel immediately after fixating the image on the right panel (*gaze-back* trials) were also removed since there was a possibility that the lambda response was contaminated with the ocular artifact observed before a saccadic movement. The judgment of the *gaze-back* trials was based on whether three consecutive saccadic movements exist (i.e., gaze at the image on the right panel, gaze back at the image on the left panel, and gaze at the image on the right panel again). Besides, trials that participants' responses were not recorded correctly due to mechanical troubles also were removed.

The EEG and EOG data were epoched in a range [$-5{,}000$, 5,000 ms] relative to the saccade offset point. The epoched data were band-pass filtered from 0.3 Hz to 40 Hz (-6dB cutoff) using a 3020th zero-phase Kaiser-windowed sinc finite impulse response filter. To correct eye blinks, the concatenated EEG and EOG data across trials were decomposed into independent components. The component of eye blinks was determined based on visual inspection of the waveforms of components. The trials were further epoched in a range [-100, 350 ms]. Since the prestimulus region is contaminated with ocular artifacts related to saccadic movements [4], the trials were baseline-corrected using a mean value of an entire epoch. Trials including a data point surpassing ± 50 μv in O1 and O2 electrodes were removed from the further analysis. The mean percent of rejected trials among the existing trials was 0.11% (SD = 0.26) across participants.

The mean amplitudes of the lambda response were calculated in a 40-ms time window centered on a peak of the response, which is determined by taking a peak observed in a $\pm$30-ms time window centered on the peak latency of the lambda response in grand-average EFRPs. The mean amplitudes were obtained per session in each participant's EFRPs. To use the same number of trials across sessions, the minimum number of trials for all sessions was used for averaging per participant data. The trials used for averaging were randomly chosen per session. Five participants were removed from the further analysis since peaks of the lambda response were not observed in more than one session. To calculate a log power in the alpha frequency range, fast Fourier transformation was applied to a Hanning-windowed entire epoch, which was padded with trailing zeros to the next power of 2. A log of the power was averaged from 8 to 12 Hz. To improve the signal-to-noise ratio, the mean amplitudes of the lambda response and the log power in the alpha frequency range were calculated using data averaged between O1 and O2 electrodes.

D. Statistical analysis

To determine whether the mean amplitudes of the lambda response significantly explain the number of inspection errors, GLMM was constructed using the log link function and a Poisson distribution. We used a *lme4* package [12]. A response variable was a number of the inspection errors (N_{errors}): judgments of the same images as different ones or different images as the same ones. The model included the mean amplitudes of the lambda response ($lambda.amp$) and a log power in the alpha frequency band ($alpha.log.pow$) as fixed effects. The days and sessions were also added in the model as fixed effects since it is conceivable that the number of inspection errors might be changed over time. They were treated as ordered factors. The model fitted a random intercept for each participant. To take into account that the number of inspection errors is proportional to the total number of the trials used for the analysis, the model specified a log of the total number of trials (N_{trials}) as an offset in the model. The model expression in the Wilkinson–Rogers notation was $N_{errors} \sim lambda.amp + alpha.log.pow + day + session + offset(logN_{trials}) + (1|participants)$.

We confirmed overdispersion of the model by a overdispertion paramter $\hat{\phi}$ [13] using equation 1:

$$\hat{\phi} = \frac{\sum_i (y_i - \hat{\lambda}_i)^2 / \hat{\lambda}_i}{residual\ degrees\ of\ freedom} \tag{1}$$

where y_i and $\hat{\lambda}_i$ are the observed data and the estimated expectation of the Poisson distribution of i-th sample, respectively. The significance of the estimated coefficient $\hat{\beta}$ of each fixed effect was determined using the Wald test. To take the overdispersion into account, the calculation of p value was corrected by multiplying the standard errors of the estimated coefficients $\hat{\beta}$ by the square root of the overdispersion parameter $\hat{\phi}$. To avoid multicollinearity, variance inflation factors (VIFs) were confirmed. The alpha level was 0.05.

III. Results

Table I represents the mean percent of participants' judgement types per day and session. Fig. 3A depicts boxplots of the percent of inspection errors per day and session, indicating a trend to decrease over day and session. The grand-average EFRPs (N = 45) are shown in Fig. 3B. The peak latency in the mean grand-average EFRPs across O1 and O2 electrodes was 80 ms.

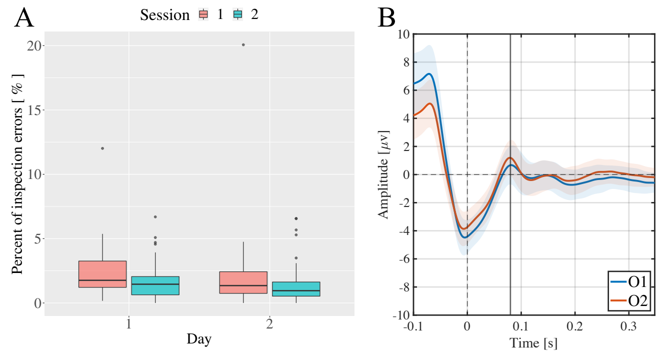

Fig. 3: (A) The percent of inspection errors per session and day. (B) The grand-average EFRPs with standard errors. The solid vertical line represents a peak of the lambda response.

Fig. 4 depicts scatterplots between the percent of inspection errors and (A) the mean amplitudes of the lambda response, and (B) the log power in the alpha frequency band, respectively. In the GLMM analysis, the overdispersion parameter of the model was 2.08. The VIF of every fixed effect was < 1.5. The coefficients of Day ($\hat{\beta} = -0.18$) and Session ($\hat{\beta} = -0.23$) reached the significance ($p < 0.01$) in the Wald test. These significant negative coefficients suggest that the inspection errors decrease over days and sessions.

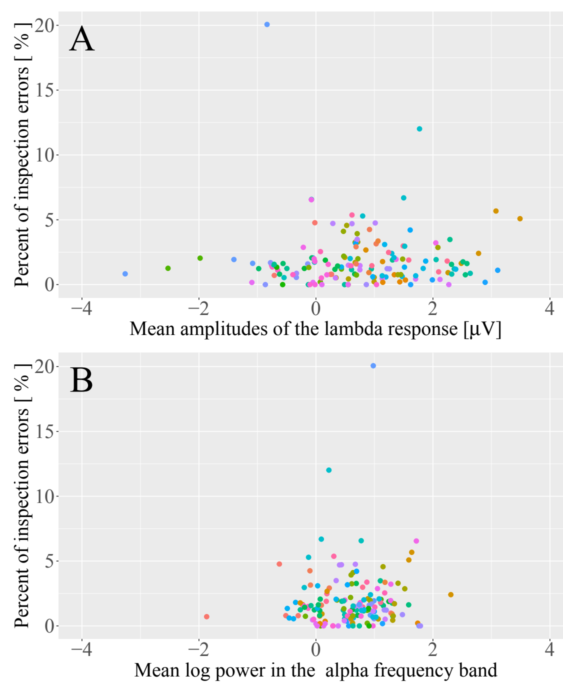

Fig. 4: (A) Scatterplots of percent of inspection errors and mean amplitudes of the lambda response and (B) the log power in the alpha frequency band. Participants were distinguished by color.

Notably, the coefficient of the mean amplitudes of the

TABLE I: The mean percent of participants' judgements per day and session.

	Day 1				Day 2			
	Session 1		Session 2		Session 1		Session 2	
Actual image	Same	Different	Same	Different	Same	Different	Same	Different
Same	81.1(2.2)	0.7(1.0)	80.7(1.7)	0.4(0.7)	80.7(2.0)	0.4(0.8)	80.4(2.3)	0.4(0.7)
Different	1.7(1.6)	16.5(2.3)	1.4(1.3)	17.5(2.2)	1.5(3.1)	17.4(3.5)	1.1(1.4)	18.1(2.7)

NOTE. SD is given in parenthesis.

lambda response was significant ($\hat{\beta} = 0.24$, $p < 0.01$), indicating that an increase in the inspection errors with an increase in the mean amplitudes of the lambda response. The coefficient of a log power in the alpha frequency band, on the other hand, was not significant ($\hat{\beta} = -0.07$, $p = 0.71$). Table II summarises the results of the GLMM analysis.

TABLE II: Summaries of GLMM analysis

Fixed effects	Estimates	S.E.	z-value	p-value
Lambda.amp	0.24	0.09	2.67	$< 0.01**$
Alpha.log.pow	-0.07	0.17	-0.38	0.71
Day	-0.18	0.07	-2.67	$< 0.01**$
Session	-0.23	0.06	-3.83	$< 0.01**$

NOTE. S.E. denotes standard error. ** indicates $p < 0.01$.

IV. Discussion

A GLMM analysis revealed that the mean amplitude of the lambda response plays a role in a predictor of the number of inspection errors –an increase in the number of the errors as an increase in the mean amplitude of the response.

Given that a decrease in the mean amplitude of the lambda response correlates with degraded attention to a visual target [8], we had expected that a negative relationship between the mean amplitudes and the number of inspection errors: The number of errors increased as a decrease in the mean amplitudes because degraded attention to inspection images can be considered as one of the factors to induce the errors. Contrary to the prediction, the positive relationship between the lambda responses and the number of errors might suggest that excessive attention to inspection images rather induces the errors. In the scenario, the session where the mean amplitudes were relatively high indicates that participants attempted to keep a high degree of attention to inspection images. However, in such a session, localized attention decrease might have happened more frequently than a session where the mean amplitudes were relatively low because it is not easy to maintain a high degree of attention to the inspection images throughout a session. To elucidate the validity of the current explanation and the detailed relationship between the lambda responses and inspection errors, further investigation is necessary for the future.

V. Conclusions

The current research suggests measurements of the lambda response during the visual inspection would be beneficial in terms of estimating the number of human error occurrences. For practical usage, future research will focus on whether the use of the EEG features can be applied to a real visual inspection at a production site.

VI. Conflict of Interest Statement

The authors declare that the current research was funded by OMRON Corporation. The funder was engaged in the design of data collection, data analysis, discussion of the results, and preparation of the manuscripts.

References

[1] S. Makeig and T. P. Jung, "Tonic, phasic, and transient EEG correlates of auditory awareness in drowsiness," *Cogn. Brain Res.*, vol. 4, pp. 15–25, 1996.

[2] Y. Yokota, S. Tanaka, A. Miyamoto, and Y. Naruse, "Estimation of human workload from the auditory steady-state response recorded via a wearable electroencephalography system during walking," *Front. Hum. Neurosci.*, vol. 11:314, 2017.

[3] F. Dehais, I. Rida, R. N. Roy, J. Iversen, T. Mullen, and D. Callan, "A pBCI to predict attentional error before it happens in real flight conditions," *Conf. Proc. IEEE Int. Conf. Syst. Man. Cybern.*, pp. 4155–4160, 2019.

[4] F. Hutzler, M. Braun, M. L. H. Võ, V. Engl, M. Hofmann, M. Dambacher, H. Leder, and A. M. Jacobs, "Welcome to the real world: Validating fixation-related brain potentials for ecologically valid settings," *Brain Res.*, vol. 1172, pp. 124–129, 2007.

[5] A. R. Nikolaev, C. Nakatani, G. Plomp, P. Jurica, and C. van Leeuwen, "Eye fixation-related potentials in free viewing identify encoding failures in change detection," *NeuroImage*, vol. 56, pp. 1598–1607, 2011.

[6] A. Yagi, "Saccade size and lambda complex in man," *Physiol. Psychol.*, vol. 7, pp. 370–376, 1979.

[7] K. Kazai and A. Yagi, "Comparison between the lambda response of eye-fixation-related potentials and the P100 component of pattern-reversal visual evoked potentials," *Cogn. Affect. Behav. Neurosci.*, vol. 3, pp. 46–56, 2003.

[8] A. Yagi, "Visual signal detection and lambda responses," *Electroencephalogr. Clin. Neurophysiol.*, vol. 52, pp. 604–610, 1981.

[9] Y. Takeda, N. Yoshitsugu, K. Itoh, and N. Kanamori, "Assessment of attentional workload while driving by eye-fixation-related potentials," *Kansei Eng. Int. J.*, vol. 11, pp. 121–126, 2012.

[10] A. Martel, S. Dähne, and B. Blankertz, "EEG predictors of covert vigilant attention," *J. Neural Eng.*, vol. 11, 2014.

[11] R. Oostenveld, P. Fries, E. Maris, and J. M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, 2011.

[12] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *J. Stat. Softw.*, vol. 67, pp. 1–48, 2015.

[13] J. J. Faraway, *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models, 2nd edition.* Chapman and Hall/CRC, 2016.