

Tracking Vigilance Fluctuations with Aperiodic Semi-Transparent Visual Flickers

P. Cimarosto¹, F. Dehais¹

¹Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, France

E-mail: pietro.cimarosto@isae-supero.fr

ABSTRACT: We investigated whether aperiodic semi-transparent textured rapid visual stimuli (RVS) can provide an unobtrusive neural probe of vigilance fluctuations during a 60-min Mackworth clock vigilance task. Six healthy participants performed the task while EEG was recorded continuously and nearly imperceptible visual stimuli were superimposed on the interface. We analyzed event-related potentials (ERPs) and conventional band-power features at both the group and single-trial levels. Preliminary results showed significant hit-versus-miss differences in the N200 and P300 amplitudes elicited by the stimuli, whereas classical frontal theta and parieto-occipital alpha features did not significantly distinguish the two conditions. Single-trial classification yielded modest performance overall, although ERP-based features appeared more informative than conventional band-power markers, and combining evoked and spectral features slightly improved hit-versus-miss discrimination. These preliminary findings suggest that aperiodic semi-transparent RVS may offer a promising and unobtrusive approach for probing vigilance fluctuations.

INTRODUCTION

Sustained attention and vigilance are critical in many real-world monitoring situations, especially when operators must supervise highly automated systems for prolonged periods with few salient events. In such contexts, attentional lapses or hypovigilance can have serious operational consequences, including in aviation and air-traffic control, where vigilance degradation is linked to performance [1, 2]. Traditional passive BCI approaches for vigilance assessment have therefore mostly relied on spontaneous EEG markers, in particular spectral features such as increased theta/delta activity, alpha modulations, or combinations of frontal theta and parietal alpha, which have long been associated with fatigue, reduced alertness, and vigilance decrement [2–7]. However, while these endogenous frequency-based markers are informative, they often provide indirect estimates of attentional state and can be sensitive to inter-individual variability and contextual factors.

Visual evoked potentials (VEPs) offer a complementary route because they provide an externally time-locked neural probe of moment-to-moment information process-

ing. Several studies have employed the visual oddball paradigm as a secondary task in dual-task settings, reporting reduced ERP amplitudes when the primary task (e.g., driving, walking, or visually demanding activities) became more challenging [8–12]. Yet, conventional visual stimulation paradigms remain difficult to deploy continuously in ecological settings: repetitive flashes and flicker can be visually intrusive, distracting, and fatiguing, especially during prolonged use [13]. This limitation is particularly problematic when the goal is not to solicit an overt response, but rather to unobtrusively probe ongoing cognitive state in parallel with a primary task.

To address this issue, recent work has explored more comfortable forms for rapid visual stimulation (RVS) paradigms, including reduced-luminance and semi-transparent textured flickers that preserve decodable brain responses while limiting visual discomfort. These approaches have shown promise both in steady-state visual evoked potential (SSVEP)-based BCI applications [13] and, more recently, in probing attentional fluctuations [14]. Along this line, StAR (Stimuli for Augmented Response) stimuli are semi-transparent textured visual elements based on a Ricker-wavelet design that can be overlaid onto a task interface with minimal disruption. First introduced in the context of reactive BCIs [15], StAR stimuli have been shown to enable rapid calibration, elicit cortical responses with amplitudes up to twofold greater than those evoked by conventional flickers, and improve user comfort [16–18]. More recently, this family of transparent flickers has also been used in passive settings to tag task-related states such as cognitive workload, while remaining sufficiently unobtrusive to be embedded directly into the visual scene [19]. Such a constant visual probing strategy is attractive because, unlike oddball-style probes, it does not require adding a secondary detection task that could alter the cognitive process under study.

In the present study, we examine whether an aperiodic StAR-based probing approach can track moment-to-moment fluctuations in attention during the Mackworth clock task. We hypothesize that the amplitude of the evoked response elicited by these nearly imperceptible flashes decreases during attentional lapses. To benchmark the sensitivity of this measure, we compare these evoked features with conventional EEG band-power markers commonly used for vigilance assessment.

This comparison is conducted at two complementary levels: first, at the group level, through statistical contrasts between hits and misses; and second, at the single-trial level, within a BCI framework, by evaluating their ability to predict behavioral outcome.

MATERIALS AND METHODS

Participants: We collected experimental data from 6 healthy participants (3 men, 3 women; age: 24.3 ± 4.3 years) with normal or corrected-to-normal vision. All participants provided written informed consent before taking part in the experiment.

Task: The experimental task was based on the Mackworth clock paradigm. Participants observed an interface in which a red dot moved along a circular array of white rings, similarly to the hand of a clock. The dot advanced by one position every 0.8 s. Occasionally, with pseudo-random occurrence, it skipped one position and jumped to the following one. Participants were instructed to respond to these skip events by pressing the spacebar on a standard keyboard. Failure to respond within 3 s after a skip was classified as a miss.

A semi-transparent StAR stimulus based on a Ricker-wavelet textured patch was displayed over the main task interface and flickered throughout the task according to a cyclic aperiodic 2.2 s long code (132 frames on a 60 Hz display). Following the *pseudo-random* method from [18], we use 2-frame bursts (~ 33 ms), resulting in inter-burst intervals ranging from 8 frames (~ 133 ms) to 40 frames (660 ms).

Protocol: Each participant completed a 60-min computer-based vigilance task derived from the Mackworth clock paradigm [14]. Before the experiment, participants completed questionnaires assessing demographic characteristics and current mental state, including the Chalder Fatigue Scale, the Karolinska Sleepiness Scale (KSS), and a Digital Eye Strain questionnaire. At the end of the session, they also completed a questionnaire evaluating the comfort and distractiveness of the StAR visual stimuli. The analysis of these subjective measures falls beyond the scope of the present preliminary report and will be carried out after completion of the full data collection. EEG was recorded continuously throughout the task using a 32-channel semi-dry MBrainTrain system, while participants were simultaneously exposed to aperiodic StAR flickering stimuli superimposed on the task interface in order to probe visual cortical activity.

Group-Level Analysis: We first analyzed behavioral performance using reaction times to skip events and the number of missed responses. At the neurophysiological level, we extracted three classes of EEG features: (1) conventional band-power features, (2) event-related potentials (ERPs), and (3) event-related spectral perturbations (ERSPs).

For band-power analyses, we considered normalized frontal theta activity (4–8 Hz) and normalized parieto-

occipital alpha activity (8–12 Hz). For ERP analyses, we focused on the responses elicited by the StAR stimuli in visual cortical regions and computed the average amplitude in two time windows: an N200 window from 0.175 to 0.225 s and a P300 window from 0.25 to 0.3 s.

These features were analyzed at the group level by comparing their values in the period preceding a successful response (Hit) or a missed response (Miss). To this end, features measured within the 10 s preceding each event were compared using one-sample permutation t-tests on the distribution of subject-level average features. This analysis aimed to evaluate the reliability of the proposed markers over prolonged task duration and their sensitivity to fluctuations in attention.

Single-Trial Level Analyses: We evaluated the practical relevance of these features using single-trial classification to predict *Hit vs Miss*. Each skip event was characterized by four features grouped into two categories: ERP and Band. The ERP category included average N200 and P300 amplitudes; the Band category included normalized frontal theta and parieto-occipital alpha activity computed over the 10 s preceding each skip.

For the classification problem, trials were labeled according to behavioral outcome following the skip event (0: Hit, 1: Miss). Classification was performed using a linear discriminant analysis (LDA) classifier with eigenvalue decomposition and automatic Ledoit–Wolf shrinkage. For each problem, classification was evaluated using each feature category separately and then all features jointly. Performance was estimated using 3-fold cross-validation repeated over 1000 iterations of random undersampling of the majority class. Accuracy values were averaged across folds and iterations. To account for inter-subject variability, the full classification procedure was performed both separately for each participant (subject-specific model) and on the pooled multi-subject dataset (generalized model). In the subject-specific case, reported accuracy corresponds to the average across participants.

RESULTS

Analysis of preliminary data from six participants showed that, behaviorally, the mean number of missed responses per participant was 32.6 ± 19.1 , corresponding to 20.2% of the total number of events. The grand-average response time was 0.595 ± 0.135 s.

Regarding differences in feature states preceding a confirmed response (Hit) versus a missed response (Miss), parieto-occipital alpha and frontal theta showed no significant differences in band activity between conditions. In contrast, significant differences were observed for both the N200 amplitude ($p = 0.03$, Cohen's $d = -0.840$) and the P300 amplitude ($p = 0.03$, Cohen's $d = 1.002$).

Table 1 reports the results of the two classification tasks—Hit vs. Miss and time-on-task (Early vs. Late, corresponding to the beginning and end of the experiment, respectively)—for both subjective and generalized mod-

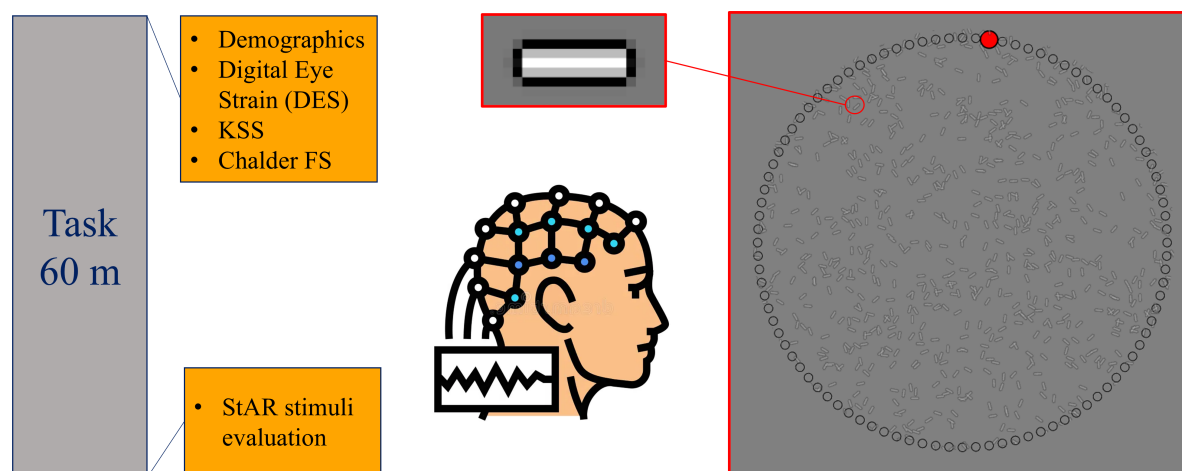


Figure 1: Structure of the experimental procedure (left) and interface (right). Participants were required to answer to questionnaires before and after the task, this included demographic questionnaires, Chalder Fatigue Scale, Karolinska Sleepiness Scale (KSS), Digital Eye Strain (DES) and StAR stimuli evaluation. The task is based on the Mackworth Clock task, in which a red dot moved across a circular path, jumping from one position to the following every 0.8 s. Occasionally, the red dot would skip a position, jumping to the one after (*skip*). Participants were instructed to respond to such event by pressing the spacebar key of a standard keyboard (*Hit*). Failing in doing so within 3 s from the *skip* occurrence would result in a missed response (*Miss*). EEG was collected continuously during the task while a Ricker-wavelet StAR stimuli semitransparent patch was flashed aperiodically over the main task interface following a Burst c-VEP pattern.

Table 1: Results for the hit vs miss using subjective (S) and generalized (G) models, evaluated on two feature categories (ERP and Band), both independently and jointly. For each combination of factors, we report the mean 3-fold cross-validation accuracy across 1000 dataset-shuffling iterations. We also report the dataset size before partitioning for each model and classification task; for the subjective model, this value corresponds to the average sample size across subjects.

Problem	Model	ERP	Band	All	Size
Hit/Miss	S	0.582	0.478	0.554	63.67
Hit/Miss	G	0.603	0.551	0.630	382

els, using ERP, Band, and ERSP features independently and jointly.

DISCUSSION

The objective of this study was to assess whether an aperiodic RVS approach based on semi-transparent textured flickers could be used as an unobtrusive probe of fluctuations in attention during prolonged monitoring. More specifically, we sought to benchmark the sensitivity of flicker-evoked responses against more classical EEG vigilance markers, first at the group level through statistical comparisons between conditions, and then at the single-trial level through classification analyses. In this respect, the Mackworth task provided an appropriate testbed, as it induced a substantial proportion of missed targets (20.2% on average), consistent with the vigilance decrement classically observed in prolonged low-demand monitoring tasks [14, 20].

At the group level, our main finding is that ERP responses elicited by the aperiodic StAR flickers were sen-

sitive to upcoming attentional lapses. Both the N200 and P300 amplitudes, measured in the 10 s preceding a target, were significantly reduced before misses relative to hits. By contrast, conventional oscillatory markers—namely frontal theta and parieto-occipital alpha band activity—did not significantly discriminate between the two conditions within the same analysis window, in contrast with some previous reports [20–22]. These results suggest that evoked responses to textured flickering stimuli may provide a more sensitive index of short-term attentional fluctuations than classical band-power measures in the present paradigm.

This interpretation is in line with the idea that externally evoked responses provide a more direct measure of moment-to-moment processing resources than spontaneous oscillatory activity alone. In this context, the attenuation of the N200 and P300 before misses may reflect a transient reduction in perceptual and attentional engagement preceding the behavioral lapse. Notably, a similar pattern was reported by Ladouce et al. [14], who found that steady-state responses elicited by nearly invisible flickers decreased prior to missed targets in a long-duration vigilance task, while conventional alpha and theta band features were less informative. Taken together, these findings further support the relevance of barely visible rapid visual stimulation paradigms—whether periodic or aperiodic—as visually unobtrusive tools for tracking fluctuations in attention and consolidate previous findings using similar unobtrusive StAR texture stimuli to probe mental workload [19].

The classification analyses provide a more cautious picture. Although ERP features showed statistically signif-

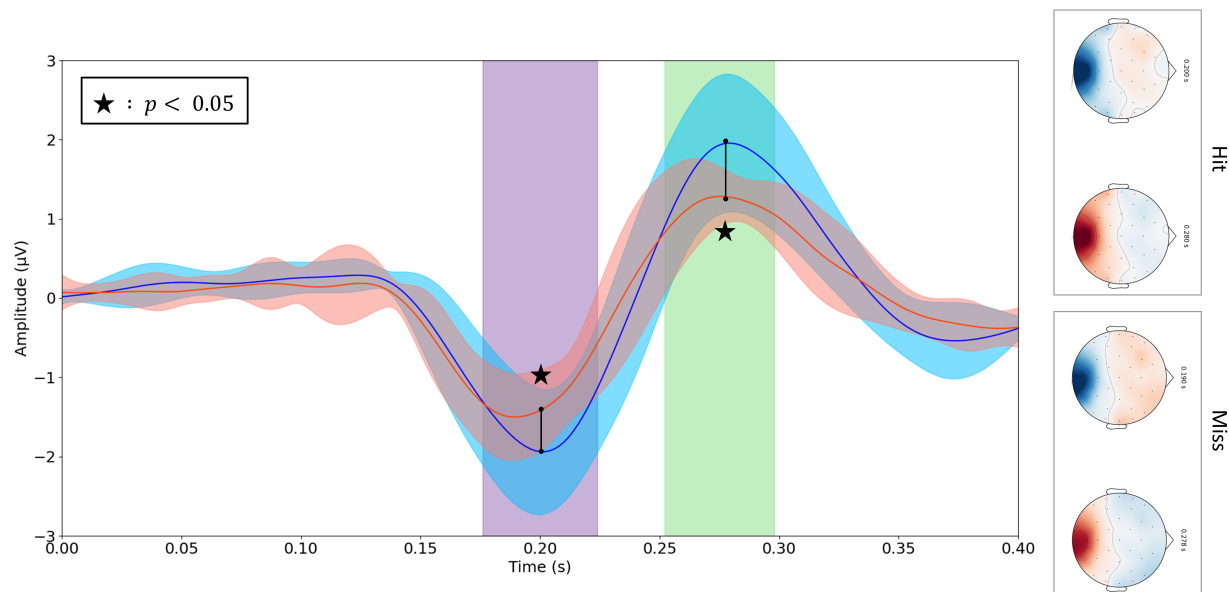


Figure 2: Event-related potentials (ERPs) in the Hit (blue) vs Miss (red) conditions displayed as grand-average wave (line) with subject-wise standard error (shaded region). A statistically significant ($p < 0.05$) difference between the distributions of average amplitude in the N200 (violet) and P300 (green) time windows across subjects between conditions was marked by a black star. On the immediate right, we display topographical plot of the main ERP peaks in both conditions (colder-to-warmer as lower-to-higher).

icant differences between hits and misses at the group level, single-trial decoding performance remained modest overall. This result is not unexpected given the limited size of the present dataset and, more specifically, the relatively small number of miss trials available for training and evaluation. Because misses are by definition less frequent than hits, the *Hit vs Miss* problem suffers from a marked class imbalance, which required repeated undersampling and reduced the effective number of training samples. Under these conditions, moderate classification performance should not be interpreted as a lack of relevance of the proposed features, but rather as a consequence of the difficulty of the problem and the preliminary nature of the dataset.

Nevertheless, some trends are worth noting. For *Hit vs Miss*, ERP features yielded better performance than band-power features in both the subject-specific and generalized models, in line with the statistical results showing greater sensitivity of the evoked markers. Combining ERP and band features further improved performance in the generalized model, suggesting that these feature families may provide partially complementary information. Overall, these findings support the feasibility of the approach, while also highlighting that substantially larger datasets will be required to properly assess its utility for robust single-trial decoding.

Several limitations should be acknowledged. First, the present study is based on a small sample of participants, and the number of miss events per participant remains limited. Second, the classification framework relied on a simple LDA model and a reduced feature set, which was appropriate for a preliminary feasibility study but may not fully exploit the information contained in the

EEG signals. Future work should therefore include additional participants, longer or repeated sessions to increase the number of lapse events, and more advanced feature extraction and decoding strategies. These developments will make it possible to determine more reliably whether the proposed aperiodic StAR paradigm can support real-time vigilance monitoring.

Despite these limitations, the present results provide encouraging evidence that an aperiodic RVS paradigm using semi-transparent textured flickers can probe fluctuations in attention in a sustained monitoring task. More broadly, they reinforce the emerging view that almost invisible flicker-based approaches can offer a useful compromise between signal sensitivity and user comfort. This makes them particularly relevant for ecological passive BCI applications, especially in operational contexts such as aviation, where continuous and unobtrusive monitoring of vigilance could help detect lapses early and support adaptive assistance systems.

CONCLUSION

In this preliminary study, we showed that aperiodic semi-transparent StAR flickers can provide an unobtrusive neural probe of vigilance fluctuations during prolonged monitoring. Flicker-evoked ERP responses, especially the N200 and P300, were sensitive to upcoming attentional lapses at the group level, whereas conventional alpha- and theta-band power features were not. Although single-trial classification performance remained limited, likely because of the small sample size and low number of misses, the present results support the feasibility of the approach. Future work will include collecting more data, integrat-

ing subjective assessments, and extending the analysis to additional neurophysiological markers such as ERSP and inter-trial coherence (ITC). These findings motivate further investigation of aperiodic flicker-based passive BCI strategies for continuous and user-centered vigilance monitoring.

References

- [1] Hamann A et al. The state of the art in assessing mental fatigue in the cockpit using head-worn sensing technology. *Frontiers in Neuroergonomics*. 2026;Volume 6 - 2025.
- [2] Aricò P et al. Adaptive automation triggered by eeg-based mental workload index: A passive brain-computer interface application in realistic air traffic control environment. *Frontiers in Human Neuroscience*. 2016;Volume 10 - 2016.
- [3] Sebastiani M, Di Flumeri G, Aricò P, Sciaraffa N, Babiloni F, Borghini G. Neurophysiological vigilance characterisation and assessment: Laboratory and realistic validations involving professional air traffic controllers. *Brain Sciences*. 2020;10(1):48.
- [4] Dehais F et al. Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fnirs-eeg passive bci. In: 2018 IEEE international conference on systems, man, and cybernetics (SMC). 2018, 544–549.
- [5] Borghini G, Astolfi L, Vecchiato G, Mattia D, Babiloni F. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*. 2014;44:58–75.
- [6] Diez PF, Correa AG, Orosco L, Laciari E, Mut V. Attention-level transitory response: A novel hybrid bci approach. *Journal of Neural Engineering*. 2015;12(5):056007.
- [7] Roy RN, Bonnet S, Charbonnier S, Campagne A. Mental fatigue and working memory load estimation: Interaction and implications for eeg-based passive bci. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2013, 6607–6610.
- [8] Yagi Y, Coburn KL, Estes KM, Arruda JE. Effects of aerobic exercise and gender on visual and auditory p300, reaction time, and accuracy. *European Journal of Applied Physiology and Occupational Physiology*. 1999;80(5):402–408.
- [9] Gramann K, Gwin JT, Bigdely-Shamlo N, Ferris DP, Makeig S. Visual evoked responses during standing and walking. *Frontiers in Human Neuroscience*. 2010;4:202.
- [10] Stróżak P, Jakubowska A, Piotrowski K, Krejtz I, Duchowski AT, Krejtz K. Vection reduces the amplitude of visual event-related potentials. *Frontiers in Psychology*. 2016;7:1952.
- [11] Baldwin CL, Coyne JT. Relative validity of the auditory and visual secondary task probes for assessing workload. *Human Factors*. 2005;47(1):157–169.
- [12] Roy RN, Charbonnier S, Campagne A, Bonnet S. Efficient mental workload estimation using task-independent eeg features. *Journal of Neural Engineering*. 2016;13(2):026019.
- [13] Ladouce S, Darmet L, Torre Tresols JJ, Velut S, Ferraro G, Dehais F. Improving user experience of ssvep bci through low amplitude depth and high frequency stimuli design. *Sci. Rep.*. 2022;12(1):8865.
- [14] Ladouce S, Tresols JT, Le Goff K, Dehais F. Eeg-based assessment of long-term vigilance and lapses of attention using a user-centered frequency-tagging approach. *Journal of Neural Engineering*. 2025;22(3):036018.
- [15] Dehais F, Cabrera Castillos K, Ladouce S, Clisson P. Leveraging textured flickers: A leap toward practical, visually comfortable, and high-performance dry eeg code-vep bci. *Journal of Neural Engineering*. 2024;21(6):066023.
- [16] Blanc C, Dehais F, Scannella S. Neuromodulation with transparent textured flicker preserves apha-band entrainment and improves visual comfort: A flanker paradigm. *International Journal of Psychophysiology*. 2025:113307.
- [17] Gomel J, Tresols JT, Cimarosto P, Castillos KC, Dehais F. Pre-decision feedback in code-modulated visual evoked potentials brain-computer interface for an 11-class keypad typing task. In: 2025 IEEE International Conference on Systems, Man, and Cybernetics (SMC). 2025.
- [18] Castillos KC, Ladouce S, Darmet L, Dehais F. Burst c-vep based bci: Optimizing stimulus design for enhanced classification with minimal calibration data and improved user experience. *NeuroImage*. 2023;284:120446.
- [19] Cimarosto P, Velut S, Cabrera Castillos K, Torre-Tresols JJ, Roy RN, Dehais F. Near-invisible c-vep-based passive bci for mental workload monitoring. *Journal of Neural Engineering*. 2026.
- [20] Martel A, Dane S, Blankertz B. EEG predictors of covert vigilant attention. *Journal of Neural Engineering*. 2014.
- [21] Shou G, Dasari D, Ding L. Pre-stimulus alpha and post-stimulus N2 foreshadow imminent errors in a single task. *Neuropsychologia*. 2015;77:346–358.
- [22] O'Connell RG, Dockree PM, Robertson IH, Bellgrove MA, Foxe JJ, Kelly SP. Uncovering the neural signature of lapsing attention: Electrophysiological signals predict errors up to 20 s before they occur. *Journal of Neuroscience*. 2009.