Task Resolution Time Estimation through Cognitive Load: An EEG Study of Chess Players

Milton O. Candela-Leal (milton.candela@exatec.tec.mx)
Mauricio A. Ramírez-Moreno (mauricio.ramirezm@tec.mx)
Jorge De-J. Lozoya-Santos (jorge.lozoya@tec.mx)
Department of Mechatronics, School of Engineering and Sciences,
Tecnologico de Monterrey, Av. Eugenio Garza Sada 2501 Sur, Monterrey 64849, Mexico

Abstract

Assessing attention is essential for optimizing performance. This study identifies a single-channel EEG biomarker based on cognitive load to estimate Task Resolution Time (TRT). Thirty-seven chess players were recorded with an 8-channel EEG headset while solving chess problems under two conditions: distracting noise (65 dB) and ambient noise (40 dB). Participants were grouped by chess expertise (ELO rating), and cognitive load was measured via theta (4-8 Hz) power on the C4 channel. EEG signals underwent preprocessing with a bandpass filter, Artifact Subspace Reconstruction (ASR), and Independent Component Analysis (ICA). Power estimation (Welch) was normalized to a resting 30-second Eyes Open (EO) period. TRT analysis indicated shorter engagement times and slightly lower performance in novices under noise, while experts remained relatively stable, possibly due to better cognitive resilience. This biomarker could be further integrated into portable EEG systems for real-time neurofeedback in educational and workplace settings.

Keywords: biomarker; brain modeling; chess; cognitive load; electroencephalography; problem-solving; theta; white noise

Introduction

Cognitive load denotes information-processing (attentional or working-memory) demands (Block et al., 2010). Monitoring and understanding human mental performance has been crucial since both low and high cognitive load affect efficiency and decision making (Wascher et al., 2014; Watt & Blanchard, 1994). Low cognitive load suggests boredom and lack of engagement (Watt & Blanchard, 1994), while high cognitive load suggests mental fatigue and errors (Wascher et al., 2014). By having a tool that measures cognitive load in real-time (Chen & Wang, 2017), it is possible to design reactive interfaces that maintain cognitive load at an optimal level in which the subject is committed to the task without reaching high mental fatigue. In Chen and Wang (2017), authors implemented the Attention Monitoring and Alarm Method (AMAM), which continuously monitored students' attention during e-learning sessions by analyzing their neural activity.

Cognitive Load Theory (CLT) is defined as the complexity of the performed task and the expertise of the learner; it is divided into intrinsic (task complexity and learner expertise), extraneous (superfluous processes not directly contributing to learning), and germane (learning processes dealing with intrinsic cognitive load) (Van Merrianboer & Sweller, 2010). It has been widely studied as it has a direct impact on the working memory (Young et al., 2014), and thus instructional methods have been developed to improve learning by decreasing extraneous cognitive load, linking cognitive processes to biological evolution and real-life tasks for complex learning (van Merriënboer & Sweller, 2005). In this study, cognitive load was induced through two distinct methods. Firstly, extraneous cognitive load, as some participants were subjected to continuous distracting white noise, similar to the static sound often heard on radios. Secondly, intrinsic cognitive load, via manipulating the complexity of chess problems.

The distracting white noise setup was inspired by Soviet and Russian chess grandmaster Mikhail Botvinnik and his odd training regime. He used to blast the radio while he was playing training games as well as having nicotine-puffing opponents blow smoke in his eyes during practice games in order to acclimatize himself for a tournament setting (Timman, 2006). Constant exposure to distracting noise has been shown to exert a psychological effect of annoyance, from a low transitive noise of air conditioning (Soeta et al., 2023) to footsteps at a slow pace in a building at noon (Frescura et al., 2022).

On the other hand, increasing chess problem complexity would increase cognitive load. Given that the task is more complex, more mental resources would be dedicated to solving the task, increasing cognitive load. A reliable method to assess a participant's cognitive state is using an Electroencephalography (EEG) to measure brainwaves (Blanco-Ríos et al., 2024; Olivas et al., 2021; Ramírez-Moreno et al., 2021). Various studies found correlations between EEG and task complexity, such as during flight (Hankins & Wilson, 1998), at chess problems (Fuentes-García et al., 2019), and in cognitive tests (Charles & Nixon, 2019). EEG has also been proven beneficial to detect subjective cognitive states such as STEM interest (Olivas et al., 2021), mental fatigue (Ramírez-Moreno et al., 2021), and emotion (Blanco-Ríos et al., 2024), in addition to having shown high correlations with subjective reports of workload (Berka et al., 2007), and thus able to discriminate it (Raufi & Longo, 2022).

Spectral analysis at EEG traces significant similarities to cognitive state identification (Lozoya-Santos et al., 2022); spectral content is divided into spectral bands: Delta (δ: 1-4 Hz) with deep sleep, theta (θ: 4-7 Hz) with interruptions of consciousness in the parietal and temporal regions, alpha (α: 8-12 Hz) with relaxation and general attention in parietal and occipital regions, beta (β: 13-29 Hz) with active thinking in the frontal region, and gamma (γ: 30-50 Hz) with perception and high-order cognition (Kropotov, 2009).

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

EEG literature agrees that cognitive load can be either detected in theta (Morales-Menendez et al., 2021) or alpha (Aguilar-Herrera et al., 2021) bands, as low-frequency bands reflect relaxation, drowsiness and mental fatigue states (Krishnan & Yaacob, 2019). However, due to alpha being so broad and complex that it has recently been found to correlate with creativity (Beaty et al., 2019; Stevens & Zabelina, 2019), only theta will be used to measure cognitive load. Theta activity mainly occurs in the frontal cortical area (Dan & Reiner, 2017) and has shown to be positively correlated with prolonged concentration while executing a task (Gevins & Smith, 2003), high-demand periods of time (Fairclough et al., 2005), and high-load (Gevins, 2000).

Due to the variability of spectral bands and cognitive states, some authors even combine frequency bands in the form of EEG indices. The most common EEG indices to track time-on-task effect and workload manipulation during the performance of a given task are engagement index ($\frac{\beta}{\theta+\alpha}$) (Freeman et al., 1999) and Task Load Index (TLI) ($\frac{\theta}{\alpha}$) (Gevins & Smith, 2003). Although these indices are effective, they are often computationally expensive and hardware-costly to calculate, as they rely on calculating beta at frontal sites, alpha at parietal sites, and theta at frontal or midline sites. Previous implementations included separate montages: Frontal (F3, F4, F7, F8), temporal, and parietal pooled over four sites (Cz, P3, Pz, P4) (Hockey et al., 2009). In an attempt to reduce this computational and hardware cost, the current paper estimates Task Resolution Time (TRT) only using ”effort”, calculated based on the theta bandpower at the fronto-central region (Gevins, 1997), which indicates increased demand and focused attention (Yamada, 1998).

Similarly, previous EEG chess literature found an increase in theta at faster games (Villafaina et al., 2019) and even identified C4 as a statistically significant channel positively correlated with problem difficulty (Fuentes-García et al., 2019). Hence, θ_{C4} bandpower was used to track the participant's effort while solving the chess problem. Additionally, a spike detection algorithm was applied in the smoothed signal in order to identify when the participant stopped making an effort, thus estimating the TRT.

Regarding the study design, participants were instructed to remain still and silent, which led to the non-recording of true labels of resolution time. They were also not allowed to skip to the next chess problem upon resolution but to wait until the time for each problem elapsed. To ensure the estimated time accurately reflects the participants' TRT, a group comparison was conducted on the calculated TRT for quiet and noise participants, stratified by chess level (novice, expert).

It was then hypothesized that novice individuals would be more affected by white distracting noise than expert individuals. Specifically, novice participants exposed to noise would perform worse and exhibit shorter TRT than those in quiet conditions. In contrast, expert participants would remain unaffected by noise, with the noisy group exhibiting TRTs similar to the quiet group.

Methods

Participants

A total of 37 chess players (35 men, 2 women) with age ranging from 18 to 23 years (μ = 20.88, σ = 1.36) participated in the study. All participants signed an informed consent form in accordance with the declaration of Helsinki (1964). They were informed about the study, their rights, risks, and benefits, and remained in close contact with the research team. For the inclusion criteria, they were required to be familiarized with solving chess problems, possess at least 800 ELO, and not manifested being under medical treatment or having a psychiatric diagnosis.

Apparatus

EEG For brainwave measurement, an 8-electrode Unicorn Hybrid Black EEG system (g.tec medical engineering GmbH, Schiedlberg, Austria) was used; it recorded brain activity signals in microvolts (μV) at a sampling frequency of 250 Hz. This non-invasive device consists of a portable cap with active gel electrodes arranged according to the international 10-20 placement system: One at the frontal region (Fz), three at the central region (C3, Cz, C4), one at the parietal region (Pz), two at the parieto-occipital region (PO7, PO8), and one at the occipital region (Oz). The electrode distribution is displayed in Figure 1.

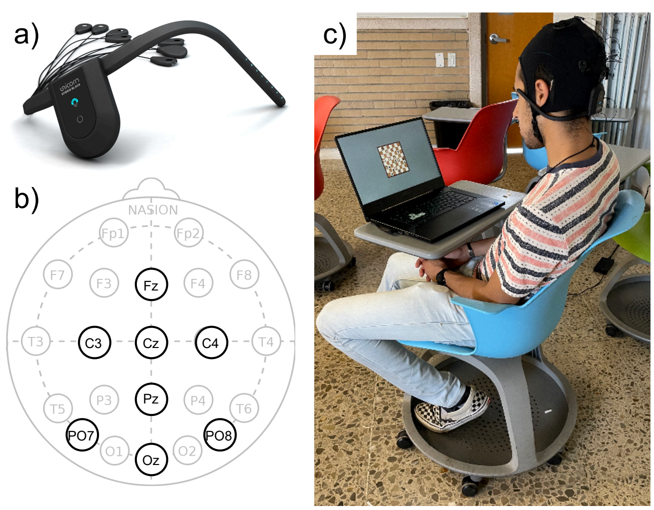

Figure 1: a) Unicorn Hybrid Black EEG wireless system without the cap. b) 10-20 international placement channel distribution: Fz, C3, Cz, C4, Pz, PO7, Oz, PO8. c) Experimental setup: A participant wearing the EEG cap solves the chess problem in his mind while staring at the screen.

Unicorn Suite Hybrid Black (g.tec medical engineering GmbH, Schiedlberg, Austria) was used to check impedance levels on all electrodes and ensure their levels were maintained low. The software indicated by color coding the impedance level of each electrode; thus, each electrode started in red color (high impedance) or yellow color (low

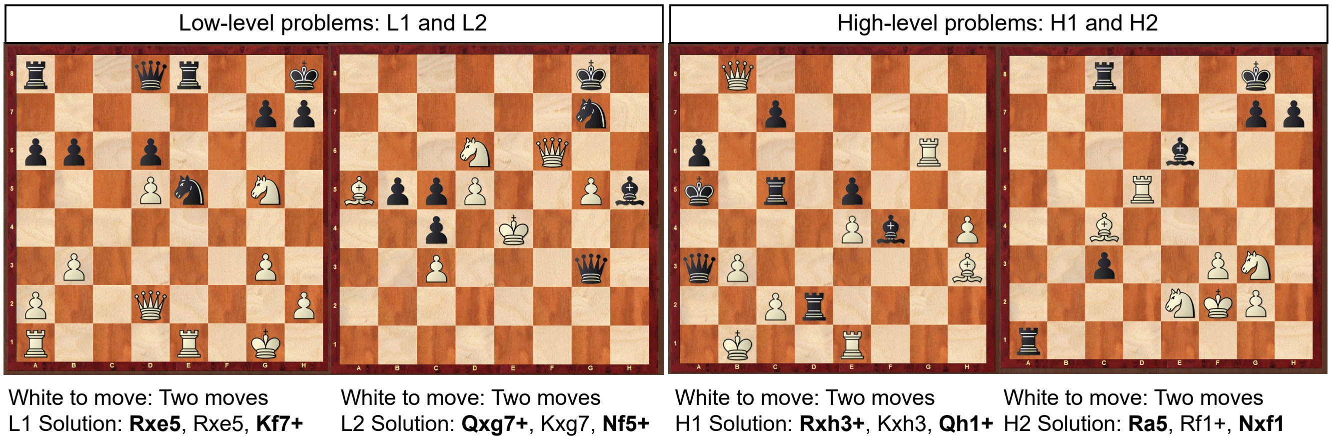

Figure 2: Two moves chess problem-solving tasks for each difficulty level from (Fuentes-García et al., 2019; Pereira et al., 2020): The two low-level problems are on the left, and the two high-level problems are on the right. Each problem was treated as a different trial for analysis purposes, for which an equal average was calculated.

impedance), and the gel was applied gradually to each electrode until all electrodes were marked in green color (low impedance) by the software. EEG signals were transmitted via Bluetooth protocol to a laptop and recorded using Unicorn Suite Hybrid Black (g.tec medical engineering GmbH, Austria).

Noise The noise level was measured using the Apple Watch Series 6 (Apple Inc., Cupertino, California, USA) Noise application, which has shown to be as accurate as a class 1 sound level meter (Muhonen et al., 2023). The average noise level for the quiet group was 40 dB. For the noisy group, a white noise signal was generated by Audacity (https://www.audacityteam.org/), further reproduced by the computer that showed the chess problems, which increased the noise level to 65 dB.

Experimental Protocol

Participants were randomly assigned to one of two between-subjects conditions: A quiet environment (40 dB) or a noisy environment (65 dB white noise). The experiment consisted of 4 chess tactical problems: 2 easy and 2 complex, adapted from (Fuentes-García et al., 2019; Pereira et al., 2020), displayed in Figure 2. These are problems where a definitive advantage is obtained in a clearly defined combination of 2 moves, thus being +3 points ahead by capturing a minor piece, such as bishop/knight, or making a quality trade (minor piece for a rook).

The chess problems are presented via PsychoPy® (Peirce et al., 2019) (https://www.psychopy.org/), an open-source platform that enables precise and automated design of psychology experiments. It offers a range of tools to display images, play sounds or videos, and time scenes, ensuring a standardized experimental process. Through this platform, an experiment was designed based on the following protocol: 30 seconds of Eyes Open (EO), 30 seconds of Eyes Closed (EC), and 150 seconds per problem (600 seconds for all 4 problems). The objective of having a 30-second EO and EC period was to normalize data, thus reducing intra-subject variability and allowing inter-subject comparisons.

During the test, the participants were instructed to solve the problem in their mind and wait for the remaining time of said problem to run out even if he completes it satisfactorily, as it has been common practice (Fuentes-García et al., 2019; Pereira et al., 2020). This approach was employed to reduce signal noise caused by physical movements and to ensure an equal sample size, regardless of whether participants solved the exercise quickly or slowly.

Unfortunately, due to device malfunction and running out of battery, data from 6 subjects were removed. Additionally, after executing the pre-processing steps, data from another 5 subjects were removed due to bad signal quality, resulting in a total of 26 subjects having usable data (24 men, 2 women).

Chess level effect size Participants' chess skill was quantified using their reported ELO rating at enrollment. The ELO rating system, developed by Arpad Elo (Elo & Sloan, 2008) and used by the Fédération Internationale des Échecs (FIDE), is a standardized scale used to estimate relative skill levels of players in competitor-versus-competitor games; it ranges from 0 to 3000, with higher values indicating greater skill (Di Fatta et al., 2009).

To reduce the effect of experience level on autonomic stress responses, groups were balanced using a stratified sampling method based on these ELO scores. Hence, the average ELO of the quiet group (N = 13) was 1302 $\pm$ 370.90, while the average ELO of the noisy group (N = 13) was 1351 $\pm$ 467.44. For analysis purposes, groups according to ELO were created; this was set to a threshold of 1200 ELO, close to all the subjects' average ELO of 1326 $\pm$ 414.17. The average ELO of the novice group (N = 13, ELO $<$ 1200) was 982 $\pm$ 106.74, while average ELO of the the expert group (N = 13, ELO $\geq$

1200) was 1670 ± 299.20. Thus, the quiet group (N = 13) comprised 6 novice and 7 expert participants, while the noisy group (N = 13) comprised 7 novice and 6 expert participants.

Signal pre-processing

EEG signal data pre-processing was conducted as described in Candela-Leal et al. (2025) and Ramírez-Moreno et al. (2023), beginning with the standardized pre-processing (PREP) pipeline (Bigdely-Shamlo et al., 2015) from the EEGLAB package (https://sccn.ucsd.edu/eeglab), which detrends using high-pass of 1 Hz, removes line noise by using multiples of 60 Hz, and also provides a robust average reference, which iteratively identifies bad channels to interpolate them. Furthermore, a Finite Impulse Response (FIR) Butterworth 4th order bandpass filter 0.1-50 Hz was applied.

Data was further cleaned using the Artifact Subspace Reconstruction (ASR) algorithm (Mullen et al., 2013) ($\kappa = 20$), which corrected periods with short-time high-amplitude artifacts. Additionally, a linear filter FIR transition band [0.25-0.75 Hz] removed channel drift, and finally, the ICLabel plugin was used to perform Independent Component Analysis (ICA) and remove artifacts from muscle, heart, eye, and noise, retaining only components with over 70% probability of being brain signals.

Signal processing

EEG pre-processed signal was processed using Scipy (1.8.1) in Python (3.9.0). Data was decomposed in one frequency band: theta (θ), between 4-8 Hz. Power Spectral Density (PSD) was calculated via Fast Fourier Transform (FFT) using the *welch* method and a 1-second window size, further calculating θ spectral power using *simps* in its frequency ranges.

θ_{C4} bandpower was normalized as in Equation 1, based on (Ramírez-Moreno et al., 2021). In which θ normalized (θ_N) is calculated by subtracting the average θ at EO ($\bar{\theta}_{EO}$) from the θ at the chess problem (θ_P), further divided by the average θ at EO ($\bar{\theta}_{EO}$).

$$\theta_N = \frac{\theta_P - \bar{\theta}_{EO}}{\bar{\theta}_{EO}} \tag{1}$$

The normalized signal (θ_N) was smoothed using a cubic smoothing spline via the *smooth.spline* function built in R. This transformation exhibited a more evident pattern within the signal during the task, in addition to improving the peak detection algorithm's performance.

Task Resolution Time Estimation

Task Resolution Time (TRT) is defined as the time required for a participant to complete the given task. However, given that a fixed time for each chess problem to be solved was set (150 seconds), it was not possible to know whether a participant solved a task faster or slower than other participants.

In this sense, a novel method for task resolution time estimation is proposed, based on (Fuentes-García et al., 2019) findings, which revealed a statistically significant increase in the theta band at C4 channel θ_{C4} as participants solved chess problems of increased difficulty, thus suggesting it to be helpful in controlling load during cognitive tasks.

The peak estimation algorithm takes into account the first highest θ_{C4} power for each problem and then calculates the time at which the peak occurred; this was performed for each problem, resulting in a per-problem TRT estimation, further averaged among chess problems to yield an unique TRT for each participant.

Statistical analysis

A two-sample t-test was applied on the average TRT for each subject for group comparisons. Due to the bandpower being normalized, and the average of the estimated TRTs for each chess problem was calculated, a t-test was then chosen in order to determine whether the difference in TRTs for each group was statistically significant.

When conducting group comparisons using TRT, a two-tailed t-test was employed. The objective was to determine whether the TRT distributions are different among quiet and noisy groups across chess level; exhibiting different behaviors when facing a problem-solving task.

Results

Chess level & performance

As expected, ELO ratings strongly correlated with performance (Di Fatta et al., 2009), explaining 82% of the variance in problems correctly solved ($p < 0.001$). In order to assess the influence of noise in this relationship, participants were separated by condition in Figure 3. Both quiet and noisy groups showed similar regression slopes, with $R^2 = 0.81$ and $R^2 = 0.84$, respectively, indicating that ELO predicted performance consistently across environments.

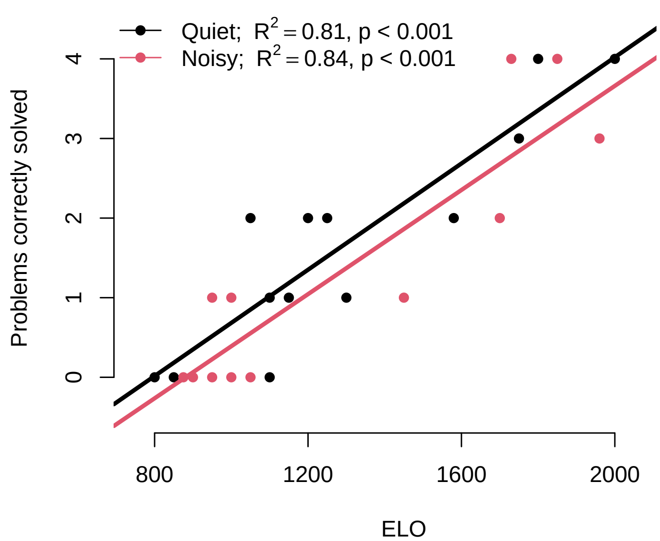

Figure 3: Performance on chess problems as a function of ELO rating, colored by noise condition. Both groups show similar regression trends and high R^2 values, indicating that ELO is a strong predictor of performance regardless of noise.

Influence of noise in performance

Distracting noise has a psychological effect of annoyance (Soeta et al., 2023), leading to an examination of its effect on participant performance. Although between-group differences were not statistically significant (novices: t = 1.0000, p = 0.1735; experts: t = -0.6386, p = 0.2686), descriptive trends emerged: novices performed slightly better in quiet environments, while experts performed slightly better in noisy ones.

To capture individual-level effects, the number of participants who outperformed the mean of the opposite condition within their group was compared in Table 1. Among experts, 4 of 6 participants in the noisy condition exceeded the quiet group's mean, while only 2 of 7 novices did so; hence, suggesting an expertise-dependent modulation of noise effects.

Table 1: Mean performance and participant counts by group and condition. M_o indicates the mean of the opposite condition within the same expertise group. $n > M_o$ reflects how many participants outperformed that value.

Group	Condition	Mean (M)	SD	n	$n > M_o$
Novices	Quiet	0.67	0.31	6	-
	Noisy	0.29	0.18	7	2
Experts	Quiet	2.57	0.43	7	-
	Noisy	3.00	0.48	6	4

TRT estimation

According to the previously described methodology, TRT was estimated using the smoothed θ_{C4} bandpower. In Figure 4, the θ_{C4} signals of two subjects from the quiet group are shown; the left-side corresponds to the subject #1 (ELO = 1000, performance = 1), while the right-side corresponds to the subject #2 (ELO = 1750, performance = 3). For visualization purposes, the calibration period (EO, EC) was excluded, as it was used in order to normalize the smoothed signal; hence, the figure only shows the signal θ_{C4} during the problem-solving task (600 seconds, 150 seconds / problem).

The figure also shows whether the chess problem was a low-level problem (L1, L2) or a high-level problem (H1, H2), in addition to the estimated peak per problem, calculated based on the first highest θ_{C4} bandpower for each problem. From the figure, at least one clear peak for each problem is evident, which shows the accumulated cognitive load when solving the chess problems, further dropping after the exercise is solved. TRT then corresponds to the duration from the start of a problem to the time at which the θ_{C4} peak occurred.

TRT chess level group comparison

The experimental protocol outlined that the effect size of chess expertise was controlled through a stratified sampling technique, which accounted for ELO ratings. This approach minimized the impact of chess expertise, as a great difference in problems correctly solved can be seen at Figure 3, despite controlling the problems' difficulty level.

The boxplot in Figure 5 shows the estimated TRT for the quiet and noisy group across the chess level (novice, expert). The TRT between novice participants in quiet and noisy groups was significantly different (t = 4.5717, p = 0.0011), which was also the case to a lesser extent between expert participants in quiet and noisy groups (t = 1.9911, p = 0.0361). Overall, the novice participants at the quiet group (NQ) experimented the highest TRT among all the subgroups: Novice noisy (NN), expert quiet (EQ), expert noisy (EN) ($\bar{x}_{NQ} = 87.00$, $\bar{x}_{NN} = 72.57$, $\bar{x}_{EQ} = 78.28$, $\bar{x}_{EN} = 63.83$), in addition to being significantly different from expert participants at the noisy setting (t = 4.2766, p = 0.0053).

Discussion

The TRT calculation shown in Figure 4, has limitations due to the nature of the task. These include double peaks caused by participants second-guessing their answers (subject #2, H1), an increase in θ_{C4} related to anticipation of the next problem (subject #1, L2), and movement-related artifacts at the end of the test (subject #1, H2; subject #2, L1). To reduce the unconscious brain's response to programmed stimuli, future protocols should incorporate randomized timing to prevent anticipation of upcoming stimuli. A semi-randomized approach with a mix of constant and variable time intervals could address this issue.

Overall, the same trend is displayed in the TRT group comparison using chess level and performance: Lower-level, novice chess players exhibited higher TRT in the quiet group setting compared to other groups. This is reasonable, as participants that are not proficient in the task tend to take more time to solve it, with a decrease in TRT in the quiet group from novice to expert (t = 1.6374, p = 0.0723); thus showing expertise in a task decreases the time required to complete it correctly. While the difference in correctly solved tasks between novices and experts is minor in both noise and quiet environments, participants showed longer TRT in Figure 5. This suggests that noise led to quicker decisions, especially in novices, possibly due to cognitive overload (Gevins, 2000), which also explains the decreased performance in novices in Table 1. The previous findings are consistent with reported findings of human behavior experiments (Chen & Wang, 2017; Fairclough et al., 2005; van Merriënboer & Sweller, 2005; Van Merrianboer & Sweller, 2010).

Distracting white noise affected participants' TRT and performance, but its impact depended on their chess expertise. Overall, it slightly reduced both their performance and TRT, but more striking results are exhibited when taking into account their chess level. While no significant accuracy difference was observed, novices in the noisy group showed shorter TRT and slightly reduced performance, suggesting noise may have disrupted their cognitive strategies. Thus, the noise affected their decision-making by lowering their tolerance for cognitive load. As a result, they rushed through the task,

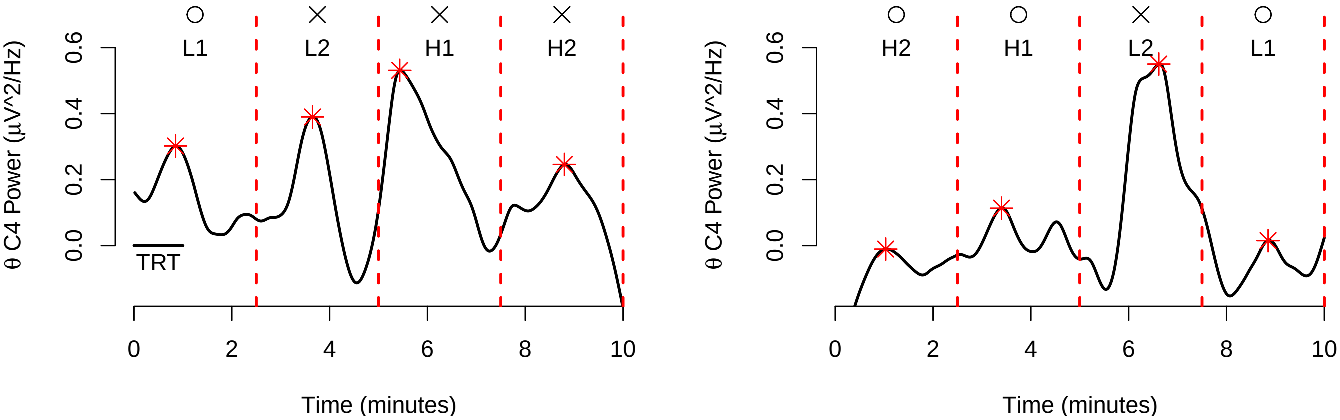

Figure 4: Theta bandpower at the C4 channel during the problem-solving task. Left: Subject #1 (Novice, ELO = 1000), right: Subject #2 (Expert, ELO = 1750). Problem labeling is based on Figure 2, problems correctly solved were marked with a circle ○, problems incorrectly solved were marked with a multiplication ×, and peaks were marked with a red asterisk ∗.

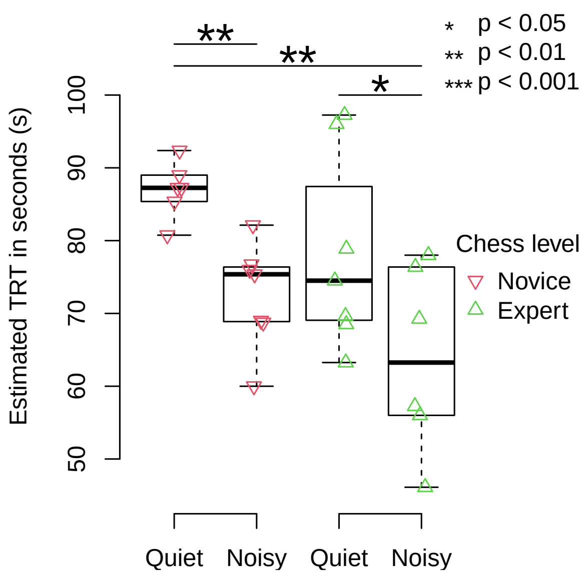

Figure 5: Boxplot of estimated TRT for each experimental condition (quiet, noisy) across chess level, measured by ELO.

spending less time solving the chess problems compared to the quiet group, which led to reduced performance.

On the other hand, the noisy environment slightly reduced expert participants' TRT and improved their performance. As the novice participants, noise also reduced their tolerance for cognitive load, reducing their task time. However, their performance remained unaffected if not improved by a distracting noisy environment; this suggests a better autonomic modulation by the expert participants when compared to the novice participants, found in previous EEG chess studies (Villafaina et al., 2021) with a decrease in alpha brainwaves, this then leading to an increase in performance as they were accustomed to problem-solving in these types of environments. Experts may be more accustomed to noise due to prior tournament experience, aiding performance.

Although response timestamps were not recorded, the TRT calculation results align with expected behavior patterns. Future studies will record TRT using non-intrusive methods, such as key presses, to preserve EEG signal integrity and minimize participant pressure. Despite the small sample size (37 participants), similar studies have used comparable or fewer subjects, such as 1 (Fuentes et al., 2018), 13 (Fuentes-García et al., 2019), 14 (Villafaina et al., 2019), and 28 (Villafaina et al., 2021); yet expanding the sample size and diversifying tasks beyond chess problems could enhance generalizability. While previous EEG and chess studies focus on brainwave activity, they have not used these measures to calculate TRT. This study is novel in doing so, as it uses a language-independent test like chess, avoiding the influence of language on cognitive load. This allows TRT to complement accuracy measures by capturing effort-related engagement, even in the absence of explicit task success. Future efforts should focus on refining TRT measurement using keystroke data and validate the biomarker across EEG systems and task types to ensure robust and generalizable models.

The presented single-channel attention biomarker (θ_{C4}), derived from the cognitive load metric "effort" (Gevins, 1997), was sensitive to cognitive load and thus task load, being able to estimate TRT. Although TRT is not a direct measure of task success, it captures participants' cognitive engagement duration, offering complementary insights into task load and attention. Next steps involve a real-time implementation to monitor cognitive load and develop adaptive interfaces that sustain focus without causing mental fatigue. Its portability and cost-effectiveness make it suitable for integration into Brain-Computer Interfaces (BCI), with potential applications in education (Aguilar-Herrera et al., 2021) and workplaces (Ramírez-Moreno et al., 2021). These findings emphasize the value of integrating neurophysiological markers like TRT with behavioral outcomes to better understand task engagement under cognitive stressors.

Acknowledgments

The authors acknowledge the technical support of the Writing Lab, Institute for the Future of Education, Tecnologico de Monterrey, in the production of this work; in addition to Mostla Tecnologías Educativas Emergentes for the Unicorn Hybrid Black headset, electrodes and gel.

They also thank undergraduate students Damian Mateo Zúñiga Cardenas, Ana Mariel Puga Torrez, and Victor Alberto López Lavoignet for their help with recruitment and data collection.

References

Aguilar-Herrera, A. J., Delgado-Jimenez, E. A., Candela-Leal, M. O., Olivas-Martinez, G., Alvarez-Espinosa, G. J., Ramirez-Moreno, M. A., de Jesus Lozoya-Santos, J., & Ramirez-Mendoza, R. A. (2021). Advanced learner assistance system's (ALAS) recent results. *2021 Machine Learning-Driven Digital Technologies for Educational Innovation Workshop*, 26–33.

Beaty, R. E., Seli, P., & Schacter, D. L. (2019). Network neuroscience of creative cognition: Mapping cognitive mechanisms and individual differences in the creative brain [Creativity]. *Current Opinion in Behavioral Sciences*, *27*, 22–30.

Berka, C., Levendowski, D., Lumicao, M., Yau, A., Davis, G., Zivkovic, T., Olmstead, R., Tremoulet, P., & Craven, P. (2007). Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, *78*, B231–44.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Front. Neuroinform.*, *9*, 16.

Blanco-Ríos, M. A., Candela-Leal, M. O., Orozco-Romo, C., Remis-Serna, P., Vélez-Saboyá, C. S., Lozoya-Santos, J. d. J., Cebral-Loureda, M., & Ramírez-Moreno, M. A. (2024). Real-time eeg-based emotion recognition for neurohumanities: Perspectives from principal component analysis and tree-based algorithms. *Frontiers in Human Neuroscience*, *18*.

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica*, *134*(3), 330–343.

Candela-Leal, M. O., Alanis-Espinosa, M., Murrieta-González, J., Lozoya-Santos, J. J., & Ramírez-Moreno, M. A. (2025). Neural signatures of stem learning and interest in youth. *Acta Psychologica*, *255*, 104949.

Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. *Applied Ergonomics*, *74*, 221–232.

Chen, C.-M., & Wang, J.-Y. (2017). Effects of online synchronous instruction with an attention monitoring and alarm mechanism on sustained attention and learning performance. *Interactive Learning Environments*, *26*(4), 427–443.

Dan, A., & Reiner, M. (2017). Real time eeg based measurements of cognitive load indicates mental states during learning. *Journal of Educational Data Mining*, *9*(2), 31–44.

Di Fatta, G., Haworth, G. M., & Regan, K. W. (2009). Skill rating by bayesian inference. *2009 IEEE Symposium on Computational Intelligence and Data Mining*.

Elo, A. E., & Sloan, S. (2008). *The rating of chessplayers : Past and present*. Ishi Press International.

Fairclough, S. H., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology*, *56*(2), 171–184.

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three eeg indices with a visual tracking task. *Biological Psychology*, *50*(1), 61–76.

Frescura, A., Lee, P. J., Soeta, Y., & Ariki, A. (2022). Effects of spatial characteristics of footsteps sounds and non-acoustic factors on annoyance in lightweight timber buildings. *Building and Environment*, *222*, 109405.

Fuentes, J. P., Villafaina, S., Collado-Mateo, D., de la Vega, R., Gusi, N., & Clemente-Suárez, V. J. (2018). Use of biotechnological devices in the quantification of psychophysiological workload of professional chess players. *Journal of Medical Systems*, *42*(3).

Fuentes-García, J. P., Pereira, T., Castro, M. A., Carvalho Santos, A., & Villafaina, S. (2019). Psychophysiological stress response of adolescent chess players during problem-solving tasks. *Physiology & Behavior*, *209*, 112609.

Gevins, A. (1997). High-resolution eeg mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385.

Gevins, A. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, *10*(9), 829–839.

Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, *4*(1–2), 113–131.

Hankins, T., & Wilson, G. (1998). A comparison of heart rate, eye activity, eeg and subjective measures of pilot mental workload during flight. *Aviation, space, and environmental medicine*, *69 4*, 360–7.

Hockey, G. R. J., Nickel, P., Roberts, A. C., & Roberts, M. H. (2009). Sensitivity of candidate markers of psychophysiological strain to cyclical changes in manual control load during simulated process control. *Applied Ergonomics*, *40*(6), 1011–1018.

Krishnan, P., & Yaacob, S. (2019). Drowsiness detection using band power and log energy entropy features based on eeg signals. *International Journal of Innovative Technology and Exploring Engineering*, *8*(10), 830–836.

Kropotov, J. D. (2009). *Quantitative EEG, event-related potentials and neurotherapy*. Elsevier.

Lozoya-Santos, J. d. J., Ramírez-Moreno, M. A., Diaz-Armas, G. G., Acosta-Soto, L. F., Leal, M. O. C., Abrego-Ramos, R., & Ramirez-Mendoza, R. A. (2022, May). Current and future biometrics: Technology and applications. In *Biometry* (pp. 1–30). CRC Press.

Morales-Menendez, R. A., Ramírez-Moreno, M. A., & de J Lozoya-Santos, J. (2021). Real-time biofeedback system for interactive learning using wearables and iot. *2021 International Conference on Industrial Engineering and Operations Management (IEOM)*, 2959–2970.

Muhonen, E. G., Abouzari, M., Yang, Y., Zeng, F.-G., & Djalilian, H. R. (2023). Audiometric validation of a smart watch decibel meter. *JAMA Otolaryngology–Head & Neck Surgery*, *149*(5), 466.

Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., Cauwenberghs, G., & Jung, T.-P. (2013). Real-time modeling and 3d visualization of source dynamics and connectivity using wearable EEG. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2184–2187.

Olivas, G., Candela, M., Ocampo, J., Acosta, L., Aguilar, A., Delgado, E., Alanis, M., Ramirez, M., Lozoya, J., Murrieta, J., & Ramírez, R. (2021). Detecting change in engineering interest in children through machine learning using biometric signals. *2021 Machine Learning-Driven Digital Technologies for Educational Innovation Workshop*, 33–40.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Pereira, T., Castro, M. A., Villafaina, S., Carvalho Santos, A., & Fuentes-García, J. P. (2020). Dynamics of the prefrontal cortex during chess-based problem-solving tasks in competition-experienced chess players: An fnir study. *Sensors*, *20*(14), 3917.

Ramírez-Moreno, M. A., Carrillo-Tijerina, P., Candela-Leal, M. O., Alanis-Espinosa, M., Tudón-Martínez, J. C., Roman-Flores, A., Ramírez-Mendoza, R. A., & de J. Lozoya-Santos, J. (2021). Evaluation of a fast test based on biometric signals to assess mental fatigue at the workplace—a pilot study. *International Journal of Environmental Research and Public Health*, *18*(22), 11891.

Ramírez-Moreno, M. A., Cruz-Garza, J. G., Acharya, A., Chatufale, G., Witt, W., Gelok, D., Reza, G., & Contreras-Vidal, J. L. (2023). Brain-to-brain communication during musical improvisation: A performance case study. *F1000Research*, *11*, 989.

Raufi, B., & Longo, L. (2022). An evaluation of the eeg alpha-to-theta and theta-to-alpha band ratios as indexes of mental workload. *Frontiers in Neuroinformatics*, *16*.

Soeta, Y., Kambara, A., & Onogawa, E. (2023). Psychological responses to low-level transient sounds in background sounds generated by an air conditioner. *Applied Acoustics*, *202*, 109147.

Stevens, C. E., & Zabelina, D. L. (2019). Creativity comes in waves: An EEG-focused exploration of the creative brain. *Current Opinion in Behavioral Sciences*, *27*, 154–162.

Timman, J. (2006, May). *Botvinnik's secret games*. Hardinge Simpole Publishing.

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177.

Van Merrianboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies: Cognitive load theory. *Medical Education*, *44*(1), 85–93.

Villafaina, S., Castro, M. A., Pereira, T., Carvalho Santos, A., & Fuentes-García, J. P. (2021). Neurophysiological and autonomic responses of high and low level chess players during difficult and easy chess endgames – a quantitative eeg and hrv study. *Physiology & Behavior*, *237*, 113454.

Villafaina, S., Collado-Mateo, D., Cano-Plasencia, R., Gusi, N., & Fuentes, J. P. (2019). Electroencephalographic response of chess players in decision-making processes under time pressure. *Physiology & Behavior*, *198*, 140–143.

Wascher, E., Rasch, B., Sänger, J., Hoffmann, S., Schneider, D., Rinkenauer, G., Heuer, H., & Gutberlet, I. (2014). Frontal theta activity reflects distinct aspects of mental fatigue. *Biological Psychology*, *96*, 57–65.

Watt, J. D., & Blanchard, M. J. (1994). Boredom proneness and the need for cognition. *Journal of Research in Personality*, *28*(1), 44–51.

Yamada, F. (1998). Frontal midline theta rhythm and eyeblinking activity during a vdt task and a video game: Useful tools for psychophysiology in ergonomics. *Ergonomics*, *41*(5), 678–688.

Young, J. Q., Van Merrienboer, J., Durning, S., & Ten Cate, O. (2014). Cognitive load theory: Implications for medical education: Amee guide no. 86. *Medical Teacher*, *36*(5), 371–384.