

# Decoding undergraduates' English reading proficiency: EEG evidence and GPA-SES effects in a computerised English assessment

Thapanee Khemanuwong, Thanate Angsuwatanakul, Ekkapon Phairot, Shaik Abdul Malik Mohamed Ismail & Nagorn Bunyarit

**To cite this article:** Thapanee Khemanuwong, Thanate Angsuwatanakul, Ekkapon Phairot, Shaik Abdul Malik Mohamed Ismail & Nagorn Bunyarit (2026) Decoding undergraduates' English reading proficiency: EEG evidence and GPA-SES effects in a computerised English assessment, Cogent Education, 13:1, 2687213, DOI: [10.1080/2331186X.2026.2687213](https://doi.org/10.1080/2331186X.2026.2687213)

**To link to this article:** <https://doi.org/10.1080/2331186X.2026.2687213>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 19 Jun 2026.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

## Decoding undergraduates' English reading proficiency: EEG evidence and GPA–SES effects in a computerised English assessment

Thapanee Khemanuwong<sup>a</sup> , Thanate Angsuwatanakul<sup>b</sup> , Ekkapon Phairot<sup>c</sup> ,  
Shaik Abdul Malik Mohamed Ismail<sup>d</sup>  and Nagorn Bunyarit<sup>e</sup> 

<sup>a</sup>School of Liberal Arts, King's Mongkut Institute of Technology Ladkrabang, Bangkok, Thailand; <sup>b</sup>College of Biomedical Engineering, Rangsit University, Bangkok, Thailand; <sup>c</sup>Faculty of Education, Songkhla Rajabhat University, Mueang Songkhla, Thailand; <sup>d</sup>School of Educational Studies, Universiti Sains Malaysia, Penang, Malaysia; <sup>e</sup>College of Innovation and Industrial Management, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

### ABSTRACT

This study investigates undergraduate students' English reading proficiency through a sequential mixed-methods design that integrates neurocognitive evidence, behavioural performance, learner perceptions, and background characteristics within a computerised assessment context. In Phase 1, paper-based testing (PBT) and computer-based testing (CBT) were compared using electroencephalography (EEG) and qualitative inquiry. Five Thai EFL undergraduates completed parallel Thai Reading Evaluation and Decoding System (Thai-READS) tasks in both formats while EEG data were recorded using a single-channel MindWave Mobile 2 device. Wilcoxon signed-rank tests indicated no significant differences in Thai-READS scores between PBT and CBT ( $Z = -1.633$ ,  $p = 0.102$ ). EEG analyses revealed largely comparable cognitive processing across test modes; however, a significant difference in alpha frequency emerged for difficult items answered incorrectly ( $Z = -2.023$ ,  $p = 0.043$ ), with higher alpha frequency observed in the CBT condition, suggesting differential neural engagement under increased task difficulty. Semi-structured interviews further indicated that students generally perceived CBT as convenient, organised, and supportive of focus, while also expressing concerns related to technical reliability, digital navigation, and limited opportunities for annotation. Phase 2 analysed Thai-READS performance data from 3,274 undergraduates using a two-way analysis of variance to examine the effects of academic achievement (GPA) and socioeconomic status (SES). Results revealed significant main effects of both GPA and SES on reading proficiency ( $p < 0.05$ ), with GPA demonstrating a larger effect size ( $\eta^2 = 0.06$ ). The findings support the feasibility of CBT while highlighting the importance of considering cognitive processing and learner background to ensure valid and equitable assessment practices in Thai higher education.

### ARTICLE HISTORY

Received 31 January 2026  
Revised 3 June 2026  
Accepted 5 June 2026

### KEYWORDS



Computer-based testing (CBT); English reading comprehension; English reading proficiency; electroencephalography (EEG); neurocognitive evidence; academic achievement (GPA); socioeconomic status (SES)


### SUBJECTS

Higher Education; Education; Social Sciences, Assessment; Social Sciences, Teaching & Learning; Secondary Education; Social Sciences, Language Teaching & Learning; Language & Literature; Arts & Humanities, Research Methods in Education; Social Sciences, Biomedical Engineering; Engineering & Technology; Technology

## Introduction

Digital assessment has become a central feature of contemporary higher education, driven by rapid technological developments and the widespread shift to online learning during and after the COVID-19 pandemic. Among these developments, computer-based testing (CBT) has gained increased attention due to its practical advantages, including automated scoring, immediate feedback, enhanced test security, flexible administration, and reduced administrative workload (Nardi & Ranieri, 2019). These advantages are particularly evident in large university classes, where traditional paper-based assessment is often time-consuming and resource-intensive (Khemanuwong et al., 2024; Ockey, 2009; Wise, 2019). In response to these trends, many higher education institutions in Thailand, including King's Mongkut

**CONTACT** Ekkapon Phairot  [ekkapon.ph@skru.ac.th](mailto:ekkapon.ph@skru.ac.th)  Faculty of Education, Songkhla Rajabhat University, Songkhla 90000, Thailand

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/2331186X.2026.2687213>.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Institute of Technology Ladkrabang, have begun adopting technology-enhanced assessment systems such as the Thai Reading Evaluation and Decoding System (Thai-READS).

Despite the growing use of computer-based testing (CBT), questions remain about its fairness, accessibility, and comparability with traditional paper-based testing (PBT). These concerns are particularly relevant in the Thai higher education context, where students' exposure to digital learning tools and access to technological resources differ substantially across academic settings and socioeconomic backgrounds. Previous studies suggest that performance in CBT may be shaped by factors other than language proficiency, such as the cognitive demands of screen-based reading, interface navigation, and familiarity with digital testing environments (Chapelle & Voss, 2017; Dooley, 2008; Gkintoni & Halkiopoulou, 2025). In reading assessment, comprehension test scores are often used to make broader judgments about students' English reading ability. As computer-based testing (CBT) becomes more widely used in educational decision-making, it is important to examine not only whether test scores are consistent across paper-based and computer-based formats, but also whether CBT influences the cognitive processes involved in reading comprehension. These factors may affect both the validity of the test and the overall implications of using CBT in reading assessments.

Electroencephalography (EEG) offers a useful approach for addressing these concerns by providing direct evidence of the neurocognitive processes involved in assessment performance. Owing to its high temporal resolution, EEG makes it possible to examine moment-to-moment changes in attentional engagement, mental workload, and neural efficiency during task completion—dimensions that cannot be fully captured through behavioural data alone. Previous research has shown that EEG markers such as alpha and beta band activity, neural entropy, and attention-related indices are sensitive to variations in cognitive effort, even in cases where test scores or observable performance remain similar (Chaiyanan et al., 2025; Klimesch, 2012). In particular, little is known about whether computer-based testing (CBT) and paper-based testing (PBT) evoke different neurocognitive responses during reading assessment tasks.

To address these gaps, this study adopts a two-phase mixed-methods design. In the first phase, EEG is used to examine Thai university students' neurocognitive responses as they complete parallel Thai-READS reading comprehension tasks delivered in computer-based and paper-based formats. Particular attention is given to indicators of mental workload, attentional regulation, and neural dynamics during English text processing. The second phase broadens the scope of the investigation by analysing Thai-READS performance data from over 3,200 undergraduate students drawn from multiple faculties and spanning a range of GPA levels and socioeconomic status (SES). This large-scale dataset allows the neurocognitive findings to be interpreted in relation to broader patterns of English reading proficiency and learner background characteristics.

By integrating neurophysiological evidence with large-scale performance data, this study offers a multi-layered perspective on English reading assessment in technology-enhanced environments. The findings contribute to ongoing debates surrounding test mode effects, assessment validity, and equity in digital testing, and provide empirical evidence to inform the design and implementation of fair and effective computerised English assessments in Thai higher education.

### **Rationale for including GPA and SES in examining English reading proficiency**

In a study aimed at examining EFL undergraduates' English reading proficiency through neurocognitive and behavioural evidence derived from computerised assessment, the inclusion of grade point average (GPA) and socioeconomic status (SES) as learner variables is both theoretically justified and methodologically important. While EEG indicators offer insight into the cognitive processes underlying reading, observable performance outcomes, such as comprehension accuracy and test efficiency, are shaped not only by neural mechanisms but also by students' academic histories and sociocultural contexts (Lokare & Jadhav, 2026; Snow, 2002; Stanovich, 2009). Incorporating GPA and SES therefore allows for a more comprehensive account of individual differences in computerised reading assessment performance among Thai undergraduates.

GPA serves as a cumulative indicator of students' academic achievement, learning strategies, and sustained engagement with academic tasks. In EFL contexts, academic achievement has been shown to correlate strongly with reading comprehension ability, vocabulary knowledge, and metacognitive strategy

use (Grabe, 2009; Zimmerman, 2002). Students with higher GPAs typically demonstrate more efficient allocation of cognitive resources, stronger self-regulation, and greater familiarity with academic texts, all of which support successful performance in both paper-based and computer-based assessments. Examining GPA alongside EEG measures enables the investigation of whether neural indicators of cognitive effort and efficiency align with established markers of academic performance, and whether computerised testing environments impose differential cognitive demands across achievement levels (Bennett, 2003; Ramírez-Moreno et al., 2021).

Socioeconomic status (SES) represents broader structural and sociocultural conditions that influence learners' language development before and during higher education. A substantial body of research suggests that students from higher-SES backgrounds are more likely to benefit from print-rich home environments, greater access to educational resources, and wider opportunities for English language exposure beyond the classroom. These advantages may strengthen reading comprehension and, more broadly, contribute to the development of English reading proficiency (Bradley & Corwyn, 2002; Hart & Risley, 1995; Phairot et al., 2022). In contrast, students from lower-SES backgrounds often encounter persistent constraints related to limited English exposure and reduced access to digital technologies, factors that can increase cognitive effort and negatively affect performance in computerised assessments (Karim et al., 2024; OECD, 2019; Tan, 2024). These SES-related disparities are particularly evident in Thailand, where structural inequalities within the education system continue to influence learning outcomes. Recent evidence from PISA-Thailand demonstrates that resource disparities and systemic constraints, including uneven teacher distribution, are closely associated with achievement gaps among Thai students, reinforcing the role of socioeconomic conditions in shaping English language outcomes (Pholphirul et al., 2023).

By examining GPA and SES in combination, the present study accounts for the interaction between academic preparation and sociocultural background in shaping both neurocognitive efforts, as reflected in EEG patterns, and behavioural performance in computerised English reading assessment. This integrated perspective is essential for evaluating the validity and fairness of computer-based testing, as score equivalence alone may obscure meaningful differences in underlying cognitive processing across learner groups (American Educational Research Association (AERA) et al., 2014; Bennett, 2003; Langensee et al., 2024). Overall, the inclusion of GPA and SES strengthens the study's contribution to neuroeducation by situating EEG evidence within real-world learner characteristics and advancing a more nuanced understanding of English reading proficiency in technology-enhanced assessment contexts.

## Research novelty and contributions

This study provides a novel interdisciplinary contribution by integrating electroencephalography (EEG) with a computer-based reading comprehension assessment (Thai-READS) to investigate the neural and cognitive processes underlying English reading proficiency among Thai learners. While previous research has compared computer-based tests (CBT) and paper-based tests (PBT) through behavioural measures such as scores and response times, this study advances the field by incorporating real-time neural indicators of cognitive load, attention, and familiarity, offering a deeper understanding of how learners process reading tasks across modalities.

The first novelty of this study lies in demonstrating that EEG-derived features, including sample entropy, alpha and beta power, and frequency, can sensitively differentiate cognitive states during CBT and PBT reading tasks. The findings indicate that CBT is associated with higher entropy and longer response times, suggesting greater cognitive demand. In addition, the observed correlations between spectral features and test familiarity highlight the potential of EEG to index reading comprehension processes beyond what can be captured through traditional behavioural data alone.

Secondly, this study provides one of the first large-scale validations of the Thai-READS computerised assessment, drawing on data from more than 3,200 Thai undergraduates. By linking reading performance outcomes with key demographic variables, namely GPA and SES, the study contributes empirical evidence supporting the reliability and discriminative capacity of the Thai-READS system across diverse academic populations.

Finally, the integration of neural and behavioural data offers a novel methodological framework for evaluating reading comprehension in EFL contexts. The study demonstrates how EEG can complement

computerised assessment by uncovering cognitive mechanisms that are not fully accessible through test scores alone. This approach provides a foundation for future research in neuro-education, technology-enhanced reading assessment, and cognitive diagnostics. It also positions EEG-supported CBT as a promising tool for producing more nuanced, evidence-based evaluations of learners' broader English reading proficiency.

## Literature review

### *Computer-based and paper-based testing*

The rapid digitalisation of educational assessment has brought renewed attention to the comparability of computer-based testing (CBT) and paper-based testing (PBT). Chalhoub-Deville and Deville (1991) early projection that large-scale assessments would progressively migrate to digital platforms has now largely materialised, particularly in the aftermath of the COVID-19 pandemic, which compelled institutions worldwide to adopt online learning and assessment tools at unprecedented speed. Although PBT continues to be widely used—owing to familiarity, infrastructural constraints, and perceived reliability—CBT has become an essential component of contemporary assessment practice.

Research investigating CBT–PBT equivalence presents a complex picture, reflecting the multifaceted nature of test delivery. Studies comparing these modes (e.g. Anakwe, 2008; Barani, 2022; Bodmann & Robinson, 2004; Ebrahimi et al., 2019; Khoshshima et al., 2017; Noyes & Garland, 2008; Sawaki, 2001) highlight differences associated not only with medium, but also with navigation, interface design, information density, and examinees' digital experience. For instance, CBT allows for flexible sequencing, dynamic item presentation, and automated scoring, which can reduce administrative errors and provide immediate feedback—features shown to enhance test-taker engagement and acceptance (Alderson, 2000; Jamil et al., 2012; Karadeniz, 2009). These functions extend assessment beyond static formats and facilitate integration with adaptive testing algorithms, offering richer data for both instructional and diagnostic purposes.

CBT also provides administrative advantages (Deutsch et al., 2012). The capacity to deliver tests remotely has proven indispensable during health-related and environmental disruptions. Moreover, the cost reductions associated with eliminating printing and distribution make CBT particularly attractive for large-scale testing programmes. However, these benefits are context dependent; institutions with limited infrastructure or unstable internet connectivity may face challenges in implementing CBT equitably.

A key concern within the CBT–PBT comparability literature is the role of digital readiness. Students who are unfamiliar with screen-based reading may experience increased cognitive load due to scrolling, screen glare, or the novelty of the interface (Dooey, 2008). These factors can adversely affect performance unless mitigated. Evidence from Chapelle and Voss (2017) and Taylor et al. (1999) suggests that training tutorials can substantially reduce such mode effects, equalising performance between students with differing levels of digital experience. Moss (2020) further argues that cognitive load can be reduced when test-takers are provided with clear navigation guides and practice tasks, reinforcing the importance of scaffolding when transitioning to CBT.

Digital technologies are widely used in Thai universities; however, classroom-related challenges remain highly context-specific (Boontong et al., 2026). The Thai context, therefore, provides a useful example of the uneven distribution of digital literacy. Students in technology-intensive disciplines tend to possess greater familiarity with online platforms, whereas those in humanities, agriculture, or industrial education may encounter more challenges. As universities increasingly integrate digital assessments, it becomes essential to empirically determine whether CBT disproportionately disadvantages particular subgroups.

To address these concerns, the first phase of the present study compared matched CBT and PBT reading tasks administered to first-year English majors. The findings revealed no significant differences in mean performance between the modes, aligning with a broader body of literature reporting negligible score differences when tests are equivalently designed (Bodmann & Robinson, 2004; Shudong Wang et al., 2008). Even so, performance-based comparability does not fully address concerns about cognitive processing. Two test formats may produce similar scores while generating different cognitive demands, raising questions about the validity and construct equivalence of each mode. This gap points to the

need for cognitive and neurophysiological measures capable of capturing differences not visible through behavioural outputs alone.

### **Psychometric test mode effects and measurement invariance across paper-based and computer-based testing**

From a psychometric perspective, test mode effects and measurement invariance are essential for evaluating the equivalence of paper-based testing (PBT) and computer-based testing (CBT). Measurement invariance assumes that an assessment measures the same underlying construct across administration modes, enabling meaningful score comparisons (American Educational Research Association (AERA) et al., 2014). However, empirical research indicates that CBT can alter test-taking behaviour due to differences in presentation format, screen-based reading, navigation demands, and interaction requirements, potentially introducing construct-irrelevant variance (Bennett, 2003; Clariana & Wallace, 2002). These effects are particularly salient in reading assessments, where attentional control, working memory, and decoding processes are central, and where screen-based formats may increase extraneous cognitive load, especially for difficult items or under time pressure (Delgado et al., 2018; Keng et al., 2008).

Integrating EEG measures into test mode research extends traditional psychometric analyses by providing process-level evidence of cognitive and neural engagement during assessment. While PBT and CBT may appear equivalent based on behavioural outcomes alone, EEG indicators can reveal subtle differences in cognitive workload, attentional regulation, and neural efficiency associated with computerised reading tasks (Berka et al., 2007; Klimesch, 1999). Such neural variations, particularly during high-difficulty items or incorrect responses, suggest that CBT may influence underlying cognitive processing even when score differences are minimal. Consequently, combining psychometric theory with EEG evidence enables a more comprehensive evaluation of functional equivalence across testing modes and strengthens validity arguments for computerised English reading assessments in higher education (Bennett, 2003; Paek & Han, 2013).

### **EEG and cognitive processing in reading assessment**

Electroencephalography (EEG) has emerged as a valuable method for examining the cognitive processes underlying reading comprehension, offering high temporal resolution and sensitivity to subtle differences in mental workload. EEG measures the oscillatory patterns of electrical activity generated by the brain, which can be decomposed into frequency bands associated with specific cognitive states. Alpha activity (8–13 Hz) is inversely related to cognitive effort, with decreases reflecting heightened processing demands. Beta activity (14–30 Hz) is associated with sustained attention and engagement, whereas neural entropy reflects the complexity or variability of ongoing cognitive operations.

Reading comprehension is a cognitively demanding activity that requires simultaneous coordination of lower-level decoding processes and higher-level integrative skills such as inference-making, semantic integration, and monitoring of understanding. EEG provides an objective means of quantifying the cognitive effort involved in these processes, complementing behavioural indicators such as scores or completion time. This is particularly relevant in assessment contexts, where mode effects may alter cognitive processing without affecting observable performance.

Empirical research integrating EEG with reading assessment remains sparse, especially in English as a Foreign Language (EFL) context. Barani's (2022) ERP-based investigation of IELTS reading tasks represents one of the few studies examining neural responses to CBT versus PBT formats. The findings reveal that screen-based reading can elicit distinct neural patterns compared to paper-based reading, suggesting that delivery mode may influence how readers allocate attention or process textual information even when accuracy remains stable.

This line of inquiry is significant for test validity research. If CBT alters cognitive processing in ways that are not intended by test designers, the construct being measured may shift subtly from reading comprehension per se to the ability to navigate digital interfaces or manage screen-based cognitive load. EEG provides a mechanism for examining such potential shifts, enabling researchers to determine whether CBT and PBT measure the same underlying construct or introduce unintended construct-irrelevant variance.

In this regard, integrating EEG into comparability studies responds directly to calls for multi-method approaches to assessment validation. By comparing neural markers of cognitive load across CBT and PBT versions of the same reading tasks, the present study offers a novel contribution to understanding how test delivery modes affect cognitive processing in EFL reading assessment. This neurocognitive perspective complements traditional score-based analyses and supports more robust interpretations of mode equivalence within digital assessment environments.

### ***EEG brain activity of the test takers***

Electroencephalography (EEG) is a non-invasive technique used to measure and record the electrical activity of the brain. EEG is performed using a device called an EEG machine, which consists of a series of electrodes that are affixed to the scalp using a special adhesive (Angsuwatanakul et al., 2014; Pittau et al., 2014). These electrodes are used to measure the electrical activity of the brain, which is then recorded and displayed on the EEG machine as a series of waves or tracings. According to Angsuwatanakul et al. (2020), EEG has high temporal resolution, which makes it applicable for capturing neurocognitive processes and mental workload near-instantaneously. Moreover, EEG is generally considered to be a safe and painless procedure, and it does not involve the use of radiation or any other potentially harmful substances (Teplan, 2002). In literature, measuring electrical activity from the brain is useful establishing the potential of EEG to illuminate complex and transient creativity-related neural processes via electrical impulses. In addition to studying brain activity during testing, EEG can be used to study brain activity in test takers before, during, and after the testing process. This can provide insight into how test takers prepare for the test, how they react to the test environment, and how they process and react to the test results (Barani, 2022). This can be especially useful for identifying factors that may influence test performance, such as anxiety or fatigue.

In addition to studying brain activity during test-taking, researchers can also use EEG to study brain activity in response to different types of test stimuli. For example, EEG can be used to examine the brain activity of test takers while they are viewing different types of images or reading different types of text. As reported by Prisacari and Danielson (2017), computer-based testing can be implemented in educational settings without imposing additional cognitive load on students. One possible explanation for differences in scores between testing modes is that one mode imposes a greater extraneous cognitive load on the test taker than the other testing mode. If one testing mode produces high cognitive load, then working memory may be overloaded, which, in turn, may lead to poorer performance on that test. In the study, they explored the relationship between CBT and PBT testing mode and cognitive load in an undergraduate general chemistry class and concluded that there was no significant difference in the cognitive load imposed by CBT or PBT tests at the overall test level or by any question type. In addition to Crivelli et al. (2019), practices to improve cognitive function and neuroproductivity have enhanced the healthy subjects' attention, thereby, their alpha-to-beta ratio has increased. This can be useful for understanding how the brain processes and responds to different types of information, and for identifying factors that may influence test performance.

It is certainly feasible to use EEG to measure mental workload in both paper-based tests (PBT) and computer-based tests (CBT) for reading comprehension. EEG can be used to measure brain activity in real-time, which makes it an ideal tool for examining mental workload in test-taking situations. One approach to evaluating the feasibility of using EEG to measure mental workload in reading comprehension tests would be to compare the brain activity of test takers completing easy and difficult reading comprehension tests, both CBT and PBT. By comparing the brain activity of test takers completing easy and difficult tests, researchers can gain a better understanding of how mental workload varies with test difficulty and test format. In addition to evaluating mental workload, EEG can also be used to evaluate brain function during CBT and PBT reading comprehension exams. For example, researchers could use EEG to examine the brain activity of test takers as they read and comprehend the test materials, in order to better understand the cognitive processes involved in reading comprehension. This could provide valuable insights into how brain function is affected by different test formats and how to optimise test design to maximise performance.

A question that arises when one aims to put this knowledge into practical use is which variables should be measured to get the best mental workload assessment for a specific individual. This is not easy to answer based on the current literature because of several complications. Firstly, limited sets of variables are recorded and analysed in each study, excluding easy comparison of performance across variables. Secondly, variables are often investigated and reported at a group level rather than in individual subjects. Finally, some of these workload studies present challenging experimental methods in which variables being measured are confounded, for instance with body movements (potentially affecting heart rate and related variables) or visual information processing (potentially affecting eye- and EEG-based measures).

Overall, EEG has exhibited potential for achieving new insights on the neurocognitive basis of natural reading (Weiss & Mueller, 2003). Therefore, we aim to evaluate the feasibility of using EEG to measure mental workload in both PBT and CBT reading comprehension tests categorised as easy and difficult as well as use this methodology to evaluate brain function during CBT and traditional PBT reading comprehension exams.

## Research objectives

- RO1: To examine the neurocognitive differences between computer-based (CBT) and paper-based (PBT) English reading tasks of varying difficulty by analysing EEG-derived indicators including attention, response time, sample entropy, beta and alpha power, and beta and alpha frequency among undergraduate students.
- RO2: To evaluate large-scale performance patterns on the Thai-READS computer-based reading assessment across students with different GPA levels and socioeconomic status (SES) backgrounds at a Thai public university.
- RO3: To determine whether an interaction effect exists between GPA and SES in predicting students' performance on the Thai-READS computer-based reading assessment.

## Research questions

- RQ1: What neurocognitive differences arise between computer-based (CBT) and paper-based (PBT) English reading tasks of varying difficulty, as reflected in EEG-derived indicators such as attention, response time, sample entropy, beta and alpha power, and beta and alpha frequency among undergraduate students?
- RQ2: How does students' reading proficiency on the Thai-READS computer-based assessment differ across GPA and socioeconomic status (SES) groups at a Thai public university?
- RQ3: Is there an interaction between GPA and SES in predicting students' performance on the Thai-READS computer-based reading assessment?

## Research hypotheses

- H<sub>0</sub>1: There is no significant main effect of GPA on students' Thai-READS performance.
- H<sub>0</sub>2: There is no significant main effect of SES on students' Thai-READS performance.
- H<sub>0</sub>3: There is no significant interaction effect between GPA and SES on students' Thai-READS performance.

## Methodology

### Research design

This study employed a two-phase sequential mixed-methods design to investigate cognitive processing during reading comprehension assessment using electroencephalography (EEG) in both computer-based testing (CBT) and paper-based testing (PBT) formats. EEG provides high temporal resolution, enabling

the capture of subtle variations in cognitive effort, attentional engagement, and neural complexity across test modalities (Angsuwatanakul et al., 2021; Pittau et al., 2014).

Phase 1 consisted of a pilot study aimed at evaluating the feasibility of EEG measurement during reading comprehension tasks and identifying preliminary trends in cognitive load associated with CBT and PBT. This sequential design allowed initial quantitative EEG data to inform subsequent analysis and interpretation of cognitive processes in reading assessment. Phase 2 extended the study to a large-scale assessment of reading proficiency among undergraduate students, examining the influence of demographic variables, academic background, and high-school GPA on test performance.

### **Participants**

**Phase 1:** Five English as a Foreign Language (EFL) learners at CEFR B1 level (2 males, 3 females) were recruited. Inclusion criteria required participants who are:

1. healthy and free from neurological or memory disorders;
2. not taking medications affecting cognitive performance;
3. having normal or corrected-to-normal vision.

All participants provided written informed consent and were informed of their right to withdraw at any stage. The small, homogeneous sample ensured feasibility and controlled for potential confounds such as anxiety, fatigue, or prior experience with digital interfaces.

**Phase 2:** This study ensured the adherence to ethical principles. Participants were provided with comprehensive information regarding the research's objectives, procedures, potential risks, and benefits, accompanied by a clear explanation of their unequivocal right to withdraw from the study at any point without any adverse consequences. Robust measures were implemented to safeguard the confidentiality of personal information, with all data anonymised and securely stored. Throughout the study, participants were treated with respect, and their privacy and dignity were upheld, free from any form of discrimination. The data collected were exclusively utilised for research purposes and reported in an aggregated, anonymised fashion. The subsequent section presents an overview of the participants' demographic profile, encompassing their gender, regions, SES, high-school GPA, and schools. As depicted in Table 1, the participant composition included 1,901 females (58.10%), 1,350 males (41.20%), and 23 individuals (0.70%) who chose not to disclose their information.

The participants were originally from across six distinct hometown regions of Thailand. Predominantly, the central region was home to the largest number of participants, with 1,967 individuals, representing 60.10% of the total. Following this, the northeastern region had 358 participants, constituting 10.90% of the sample, while the eastern region accounted for 356 participants, representing 10.90%. In addition, the southern region contributed 246 participants, making up 7.50% of the cohort, followed by 224 northern region participants, and 123 western region participants (3.80%), constituting 6.80% and 3.80% respectively.

Table 1 provides insight into the participants' socioeconomic backgrounds, as determined by their family's monthly income. A significant portion of participants fell into three distinct categories, with 1,157 (35.30%), 1,063 (32.50%), 1,054 (32.20%), indicating their SES based on their family's monthly earnings of above 40,000 Baht, below 20,000 Baht, and between 20,000 and 40,000 Baht, respectively.

Additionally, Table 1 presents the high school GPA distribution of participants. The largest group consisted of 1,025 participants (31.30%) with GPAs of above 3.49. Following closely, there were 980 participants (29.90%) who attained GPAs in the range between 3.00 and 3.49, while 765 participants (23.40%) achieved GPAs between 2.50 and 2.99. Additionally, 504 participants held GPAs below 2.50, constituting the smallest portion of participants (15.40%).

In Table 1, a breakdown of the participants' respective schools during the data collection phase is presented. Notably, the largest representation came from the School of Science, comprising 935 participants, accounting for 28.60% of the total. Following closely, the School of Engineering was represented by 701 participants, constituting 21.40% of the sample. Moreover, the School of Agricultural Technology contributed 592 participants, making up 18.10% of the overall cohort. Additionally, there were 407 participants from the School of Architecture, Art, and Design representing 12.40%, while the School of

**Table 1.** Background of participants (N = 3,274).

Demographic information		Frequency	Percent
Gender	Male	1,350	41.20
	Female	1,901	58.10
	Prefer not to specify	23	0.70
Region	Central	1,967	60.10
	Eastern	356	10.90
	Western	123	3.80
	Northern	224	6.80
	Southern	246	7.50
	North Eastern	358	10.90
	SES	Below 20,000 Baht	1,063
	20,000–40,000 Baht	1,054	32.20
	Above 40,000 Baht	1,157	35.30
GPA	Below 2.50	504	15.40
	2.50–2.99	765	23.40
	3.00–3.49	980	29.90
	Above 3.49	1,025	31.30
	School	School of Science	935
	School of Information and Technology	141	4.30
	School of Industrial Education and Technology	249	7.60
	School of Engineering	701	21.40
	School of Architecture, Art, and Design	407	12.40
	School of Agro-Industry	139	4.20
	School of Agricultural Technology	592	18.10
	School of Integrated Innovative Technology	110	2.30
Total		3,274	100

Industrial Education and Technology had 249 participants, accounting for 7.60%. Besides that, approximately 400 participants were from three different schools, that is, 141 participants (4.30%) from the School of Information and Technology, 139 participants (4.20%) from the School of Agro-Technology, and 110 participants (2.30%) the School of Integrated Innovative Technology, respectively.

## Ethical considerations

Ethical approval for the study (EC-KMITL\_69\_043) was granted by the institutional research ethics committee prior to data collection. All research procedures were conducted in accordance with recognised ethical guidelines for research involving human participants. Informed consent was obtained from all participants, confidentiality was ensured, and participants were informed of their right to withdraw from the study at any time without penalty.

## Research instruments

### EEG measurement

EEG data were recorded using the MindWave Mobile 2 (NeuroSky Inc., San Jose, USA), a portable device that captures prefrontal cortical activity through electrodes attached to the forehead and earlobes (Alnemari, 2017). The use of a single-channel EEG device was justified by the applied nature of the study, which focused on computer-based reading assessment. The MindWave Mobile 2 was selected because it is portable, affordable, easy to administer, and minimally intrusive for test takers. These features made it suitable for collecting EEG data in an authentic assessment setting while allowing participants to complete the reading tasks under conditions closely resembling a normal computer-based test.

In this study, the device enabled the collection of general neurocognitive indicators related to mental workload, attentional engagement, and response behaviour without disrupting the testing process. Its lightweight design also reduced the likelihood that the EEG equipment would distract participants, interfere with screen-based reading, or increase test-related nervousness. This was particularly important because the study aimed to examine reading performance and cognitive engagement during computer-based and paper-based testing, rather than to localise specific neural sources. Therefore, the single-channel EEG approach was considered appropriate for exploratory educational research seeking to complement behavioural test data with accessible neurophysiological indicators. It allowed the study to

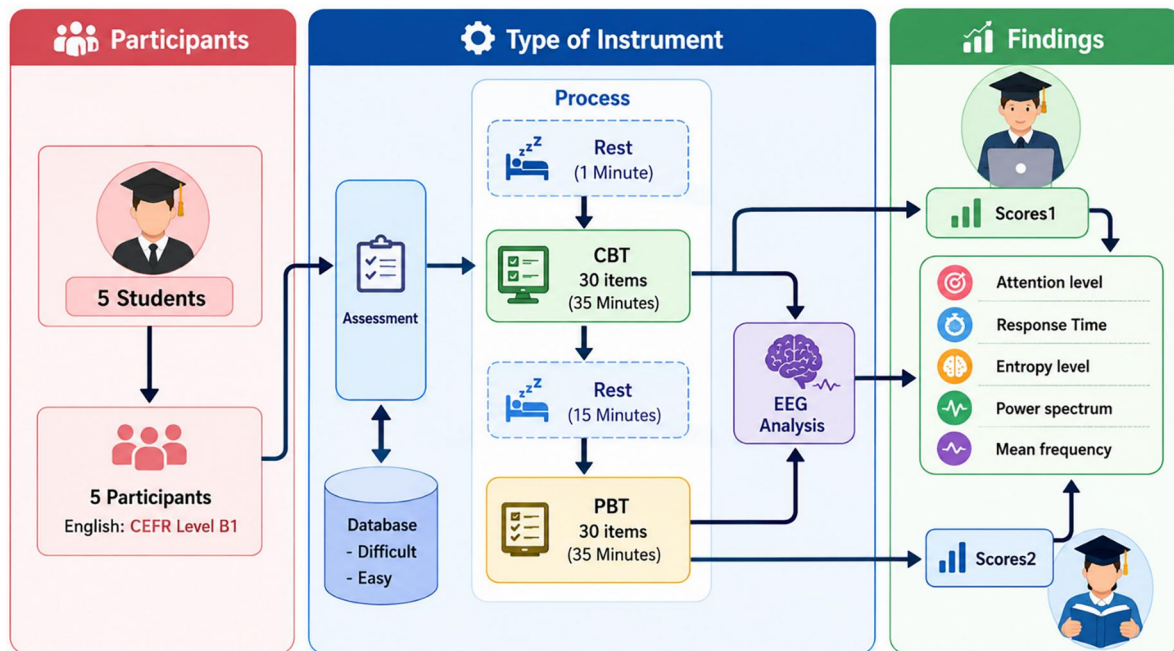


Figure 1. Experimental process of CBT and PBT.

balance methodological feasibility, participant comfort, and ecological validity while providing useful evidence of patterns of cognitive demand and attentional engagement during reading comprehension tasks.

Key metrics included: (1) Alpha power (8–13 Hz): inversely related to cognitive effort; (2) Beta power (14–30 Hz): associated with sustained attention and engagement; (3) Sample entropy: reflecting the complexity of ongoing cognitive operations; and (4) Attention level index: proprietary algorithm from MindWave Mobile.

EEG signals were amplified, filtered, and synchronised with each test item via button-press triggers, allowing precise temporal alignment of neural activity with task performance.

### Phase 1

Phase 1 was designed as a pilot feasibility study intended to explore the potential of EEG-based indicators for examining cognitive processing during CBT and PBT. The participants were English language learners with a CEFR level of B1 ( $n = 5$ , 2 males and 3 females) and did not have any history of memory disorders or any other known medical conditions that could affect their cognitive abilities. They also were not taking any medications that could potentially affect their cognition. The inclusion criteria for the study required that participants be healthy and free from any evidence of neurological disorders or potential organic causes for memory disorders. The main research instrument used for data collection was the Thai-READS administered in both CBT and PBT formats. All participants provided written informed consent and were able to withdraw from the study at any time. The experimental process was sequential, as shown in Figure 1.

In line with a study by Straker et al. (2008), the participants were also given a tutorial on how to answer computerised questions and were seated in a comfortable chair approximately 50 cm away from a computer screen for the CBT test. They were asked to do 30 items in 30 minutes, pressing a button to move on to the next item once they completed it. After the CBT test, the participants were given a 15-minute rest before starting the PBT test, which was done on paper. They were asked to press a button to proceed to the next page, and then did 30 items in 30 minutes. A button press was used to send a trigger signal that synchronised the EEG recordings with each exam question.

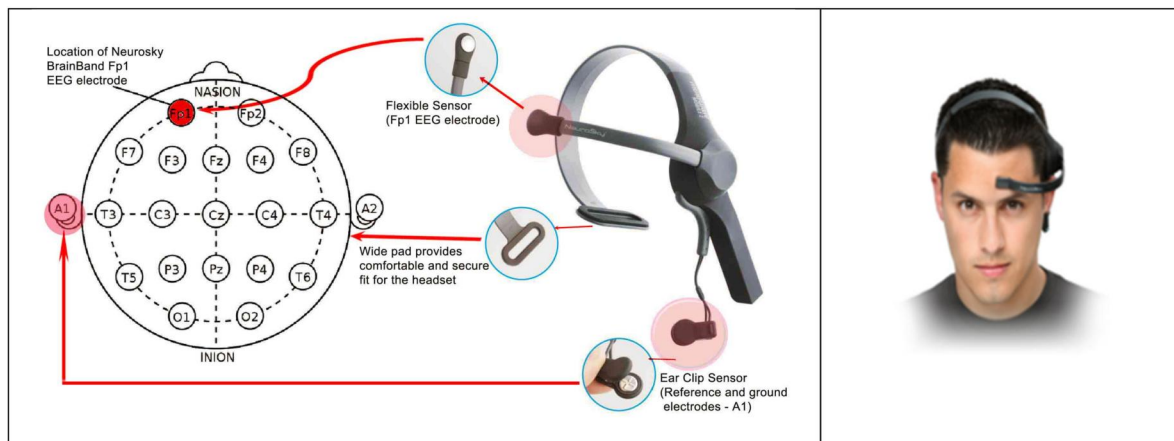


Figure 2. Neurosky mindwave mobile (Alnemari 2017).

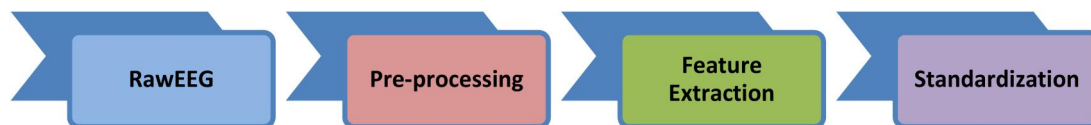


Figure 3. Procedure for analysing EEG signals.

### *Electrophysiology and data analysis*

EEG data was recorded using MindWave Mobile 2 (NeuroSky Inc.; San Jose, USA), a portable EEG device that is designed to be easy to use and relatively low-cost. It works by amplifying and filtering the electrical signals detected at the microvolt level from the scalp using electrodes attached to the earlobes and forehead. This allows it to measure and record the brain activity of test takers as they complete cognitive tasks such as reading comprehension. This obtains brain signals primarily from the prefrontal cortex, as shown in Figure 2.

EEG data collected during the computer-based testing (CBT) and paper-based testing (PBT) sessions were processed using Unity software (Unity Software Inc., San Francisco, USA) and subsequently exported to MATLAB for further analysis. Extracted features included indices of neural activity such as mean alpha and beta power and frequency, an attention-level index, and sample entropy. Response time for each test item was also recorded. EEG features were analysed across four task conditions: easy items answered correctly (E1), easy items answered incorrectly (E0), difficult items answered correctly (D1), and difficult items answered incorrectly (D0). Owing to the exploratory nature of the pilot phase and the limited sample size, multiple EEG indicators were examined to identify potential patterns rather than to test confirmatory hypotheses. Consequently, statistically significant results are interpreted with caution and are regarded as preliminary trends requiring validation in larger-scale studies. To ensure comparability across measures, all variables were standardised using z-score normalisation, such that each feature had a mean of zero and a variance of one, as illustrated in Figure 3.

### *Reading assessment*

The Thai Reading Evaluation and Decoding System (Thai-READS) is a standardised reading assessment administered in English and designed to evaluate EFL reading comprehension among Thai learners (Hui et al., 2020; Karim et al., 2024; Khemanuwong et al., 2018). The instrument has been widely implemented within Thai higher education and is appropriate for examining English reading proficiency in technology-enhanced assessment contexts.

The Thai-READS reading comprehension test was administered in both CBT and PBT formats. Each participant completed 30 items per modality within 30 minutes, separated by a 15-minute rest period to reduce fatigue. CBT items were navigated on a computer screen, while PBT items were answered on paper with EEG recording synchronised via button press.

## Methods of data analysis

Data analysis was conducted separately for Phase 1 (i.e. EEG pilot study) and Phase 2 (i.e. large-scale assessment) to address the study's research objectives. Prior to analysis, all datasets were screened for completeness, outliers, and normality. Descriptive statistics, including medians, standard deviations, and frequency distributions, were computed to summarise participant characteristics, reading proficiency, and EEG indicators.

### Phase 1: EEG and test mode analysis (CBT vs. PBT)

EEG data obtained during the pilot phase were preprocessed and analysed to examine differences in cognitive processing between computer-based testing (CBT) and paper-based testing (PBT). Raw EEG signals were filtered and synchronised with item-level responses using button-press triggers. EEG features including alpha power, beta power, alpha frequency, beta frequency, attention index, and sample entropy were extracted for each test item and categorised into four task conditions: easy items answered correctly (E1), easy items answered incorrectly (E0), difficult items answered correctly (D1), and difficult items answered incorrectly (D0). To ensure comparability across features, all EEG measures were standardised using z-score normalisation.

Given the small sample size and the non-normal distribution of several EEG indicators, Wilcoxon signed-rank tests were employed to compare CBT and PBT conditions across the four task categories. Effect sizes ( $r$ ) were calculated for statistically significant results to assess the magnitude of observed differences. Response time data were analysed alongside EEG indicators to support interpretation of cognitive workload and attentional engagement across testing modes.

### Phase 2: two-way ANOVA of reading performance by GPA and SES

In Phase 2, data from 3,274 undergraduate participants were analysed to examine the effects of academic achievement (GPA) and socioeconomic status (SES) on English reading proficiency measured through a computerised assessment. A two-way analysis of variance (two-way ANOVA) was conducted to investigate the main effects of GPA and SES, as well as their interaction effect, on reading comprehension scores.

GPA was categorised into four levels (i.e. below 2.50, 2.50–2.99, 3.00–3.49, and above 3.49), while SES was classified into three groups based on monthly household income (i.e. below 20,000 Baht, 20,000–40,000 Baht, and above 40,000 Baht). Prior to conducting the two-way ANOVA, assumptions of independence, normality, and homogeneity of variances were examined. Given the large sample size, the ANOVA procedure was considered robust to minor deviations from normality. Levene's test was used to assess the homogeneity of variances assumption.

When significant main or interaction effects were identified, post hoc comparisons were performed using appropriate multiple-comparison procedures to locate specific group differences. Effect sizes were reported using partial eta squared ( $\eta^2$ ) to indicate the magnitude of observed effects. The two-way ANOVA results were interpreted to determine whether reading performance varied systematically across GPA and SES groups and whether the influence of GPA on reading proficiency differed by socioeconomic background.

All statistical analyses were conducted using standard statistical software, with the level of statistical significance set at  $p < .05$ . The integration of large-scale behavioural data from Phase 2 with EEG-based findings from Phase 1 enabled a comprehensive examination of English reading proficiency from both neurocognitive and educational perspectives.

## Results

### Phase 1: comparison of Thai-READS performance and EEG analysis between PBT and CBT

A Wilcoxon signed rank test, a non-parametric test, was employed to compare the five participants' score on Thai-READS between PBT and CBT. As can be seen in [Table 2](#), the findings demonstrate that

**Table 2.** Results of Wilcoxon signed rank test for Thai-READS performance.

Test formats	Test statistics		
	Median	Z	Asymp. Sig. (2-tailed)
PBT	20	-1.633	0.102
CBT	19		

**Table 3.** Results of Wilcoxon signed rank test for EEG analysis.

Test items and responses	EEG analysis	Median		Z	Asymp. Sig. (2-tailed)
		PBT	CBT		
1. Difficult items with incorrect answers	Attention	0.51	0.48	-0.271	0.786
	Response Time	0.38	0.40	-0.813	0.416
	Sample Entropy	0.53	0.45	-1.219	0.223
	Beta Power	0.36	0.36	-0.677	0.498
	Alpha Power	0.49	0.45	-1.841	0.066
	Beta Frequency	0.44	0.51	-1.483	0.138
	Alpha Frequency	0.50	0.59	-2.023	0.043*
2. Difficult items with correct answers	Attention	0.51	0.50	-0.813	0.416
	Response Time	0.33	0.39	-0.552	0.581
	Sample Entropy	0.47	0.54	-0.674	0.500
	Beta Power	0.35	0.38	-0.365	0.715
	Alpha Power	0.32	0.37	-1.214	0.225
	Beta Frequency	0.50	0.44	-0.944	0.345
	Alpha Frequency	0.53	0.51	-0.405	0.686
3. Easy items with incorrect answers	Attention	0.50	0.45	-0.674	0.500
	Response Time	0.46	0.42	-0.135	0.892
	Sample Entropy	0.48	0.47	-0.674	0.500
	Beta Power	0.37	0.39	-0.677	0.498
	Alpha Power	0.43	0.39	-0.135	0.893
	Beta Frequency	0.47	0.56	-0.674	0.500
	Alpha Frequency	0.49	0.42	-0.944	0.345
4. Easy items with correct answers	Attention	0.47	0.41	-1.214	0.225
	Response Time	0.38	0.30	-1.214	0.225
	Sample Entropy	0.40	0.51	-0.677	0.498
	Beta Power	0.39	0.26	-0.944	0.345
	Alpha Power	0.35	0.32	-0.944	0.345
	Beta Frequency	0.58	0.54	-0.135	0.893
	Alpha Frequency	0.39	0.45	-0.730	0.465

\*The difference was at 0.05 of significant level.

there was no statistically significant difference in the participants' score between PBT (Md = 20) and CBT (Md = 19),  $Z = -1.633$ ,  $p = 0.102$ . The test results from PBT and CBT for each participant are provided in Appendix A.

To provide a deeper analysis of EEG analysis in different aspects which aligns with RO1, aiming to examine the neurocognitive differences between computer-based (CBT) and paper-based (PBT) English reading tasks of varying difficulty by analysing EEG-derived indicators (Appendix B), a Wilcoxon signed rank test was used to compare EEG indicators (i.e. attention, response time, sample entropy, beta power, alpha power, and beta frequency) between PBT and CBT across two levels of test item difficulty (i.e. easy and difficult) and the participants' responded answers (i.e. correct and incorrect). The results in Table 3 reveal that overall results of EEG analysis in difficult test items with correct answer, and easy test items with correct and incorrect answers showed no statistically significant differences between PBT and CBT,  $p > 0.05$ .

However, there was a statistically significant difference in difficult test items with incorrect answers based on alpha frequency analysis,  $Z = -2.023$ ,  $p < 0.05$ , with a large effect size ( $r = 0.90$ ), following Cohen's (1988) criterion. The results further indicate that the participants' score (Md = 0.59) on CBT was significantly greater than score on PBT (Md = 0.50). The significant increase in alpha frequency during CBT for difficult items answered incorrectly may indicate differential cognitive processing or neural engagement under higher task difficulty when students interacted with a computer-based testing environment

### Phase 1: student perceptions of CBT and PBT

To complement the quantitative EEG analysis, which indicated lower mental workload and more stable attentional engagement during the computer-based test (CBT), a qualitative component was conducted

**Table 4.** Themes and representative student excerpts from phase 1 interviews.

Theme	Subtheme	Excerpts (students S1–S5)
Perceived Advantages of CBT	1.1 Convenience and streamlined navigation	S1: 'On the computer I can scroll straight down ... with paper I flip back and forth and sometimes lose the page.' S3: 'CBT feels more organised. Everything stays in one place.' S5: 'I can jump to questions faster. It's cleaner than turning pages.'
	1.2 Immediate scoring and faster feedback	S2: 'If the computer shows the score right away, it saves me the stress of waiting.' S4: 'Digital tests can calculate instantly. I think it's fairer and faster.'
	1.3 Reduced paper use and institutional savings	S5: 'Using computers is better for the environment. No printing needed.' S3: 'It feels modern. The university doesn't have to prepare so many booklets.'
	1.4 Enhanced focus and reduced visual clutter	S1: 'Seeing one section at a time helps me focus.' S2: 'On the screen I read more smoothly and feel less distracted.'
Perceived Limitations of CBT	2.1 Concerns about technical reliability	S4: 'If the system freezes or the internet is slow, it might affect the test.' S1: 'In real exams, what if the computer stops working? Paper feels safer.'
	2.2 Test anxiety related to digital navigation	S3: 'Sometimes when I scroll, I'm scared I'll miss a sentence.' S5: 'I'm not used to reading long passages on the screen, so I feel a bit tense.'
	2.3 Preference for physical marking and note-taking	S2: 'On paper I can circle words or make quick notes.' S4: 'If digital marking is easier, maybe I'd switch, but right now paper feels more natural.'

to explore students' subjective experiences with CBT and paper-based testing (PBT). Semi-structured interviews with five participants (S1–S5) generated rich accounts of how students perceived both formats. Inductive thematic analysis identified two overarching themes: (1) perceived advantages of CBT, and (2) perceived limitations of CBT. These themes and their representative excerpts are summarised in Table 4, which serves as the primary reference for interpreting the qualitative findings.

### **Theme 1: perceived advantages of CBT**

Students expressed predominantly positive perceptions of the CBT environment, identifying four advantages: (1) convenience and streamlined navigation, (2) potential for faster scoring, (3) reduced paper use, and (4) improved focus. Each subtheme is closely illustrated by participant excerpts in Table 4.

#### **Convenience and streamlined navigation**

Students consistently described the CBT format as more convenient and efficient than PBT. As shown in Table 4, S1 highlighted the advantage of continuous scrolling, noting that it prevented the disruption caused by flipping paper pages. Similarly, S3 emphasised that CBT 'feels more organised', and S5 reported quicker access to questions. These statements suggest that CBT reduces physical interruptions, supporting a smoother and more continuous engagement with the text.

#### **Immediate scoring and faster feedback**

Although the testing system used in this study did not provide instant scores, students commonly referred to immediate scoring as an expected feature of digital assessment. As illustrated in Table 4, S2 associated faster feedback with reduced test-related anxiety, while S4 viewed automated scoring as both efficient and fair. These responses indicate that students value CBT not only for the testing experience itself, but also for its perceived potential to improve how results are delivered.

#### **Reduced paper use and institutional savings**

Participants also acknowledged the environmental and administrative benefits of CBT. As shown in Table 4, S5 linked digital delivery to reduced paper waste, while S3 commented on the lower institutional workload involved in preparing printed materials. Although these factors are unlikely to affect test performance

directly, they contributed to students' generally favourable perceptions of CBT as a modern and sustainable assessment mode.

### ***Enhanced focus and reduced visual clutter***

Several participants reported that CBT helped them maintain focus by limiting visual distractions. As indicated in Table 4, S1 appreciated being able to view 'one section at a time', and S2 described screen-based reading as smoother and less distracting than reading from paper. These perceptions are consistent with the EEG findings, which showed lower mental workload and more stable patterns of attentional engagement in the CBT condition.

## ***Theme 2: perceived limitations of CBT***

Despite these positive perceptions, students also identified several limitations of CBT. Their comments highlighted concerns related to technical reliability, anxiety associated with digital navigation, and difficulties using on-screen annotation tools. These issues are supported by the participant excerpts presented in Table 4.

### ***Concerns about technical reliability***

Students expressed apprehension about potential system failures, even though no issues occurred during the study. As reflected in Table 4, S4 worried about freezing screens or poor internet performance, while S1 perceived paper as more dependable. These concerns highlight structural and infrastructural challenges that institutions must address before large-scale implementation.

### ***Anxiety related to digital navigation***

Participants unfamiliar with long-form screen reading reported heightened anxiety. Table 4 shows that S3 feared missing information when scrolling, and S5 felt tense when reading lengthy passages on screen. This suggests that digital reading skills and familiarity may influence students' comfort levels with CBT.

### ***Preference for physical marking and note-taking***

Students also noted limitations in digital annotation. As seen in Table 4, S2 preferred paper because it allowed him to circle words and create quick notes, while S4 found digital marking less intuitive. These comments indicate that students value the flexibility of handwritten strategies, and that improvements in digital annotation tools may enhance user acceptance of CBT.

The qualitative findings, supported by representative excerpts in Table 4, indicate that students generally view CBT as a feasible, efficient, and acceptable medium for reading comprehension assessment. Students particularly valued its navigational ease, organisational clarity, potential for rapid scoring, and reduced environmental impact. These perceptions strongly align with the EEG findings, which showed more stable attention in the CBT condition. At the same time, students' concerns such as potential technical failures, discomfort with digital navigation, and limited annotation capabilities were less frequently expressed and appear manageable through targeted training, user-friendly interface design, and adequate test administration infrastructure. Overall, the evidence suggests that CBT offers a feasible and generally effective approach for administering the adapted Thai-READS in Thai higher education.

## ***Phase 2: main effects of GPA and SES on Thai-READS performance***

As for RO2, two-way ANOVA was utilised to examine a statistically significant main effect of high-school GPA levels and socioeconomic status (SES) backgrounds on Thai-READS computer-based reading assessment. The results in Table 5 show that there was a statistically significant main effect of high-school GPA on Thai-READS performance among the participants,  $F(3, 3262) = 65.45, p < 0.05$ , with a large effect size ( $\eta^2 = 0.06$ ) as determined by Cohen's (1988) criterion.

The results of Post-Hoc comparisons using the Tukey HSD test in Appendix C indicate that Thai-READS performance of the participants with GPA above 3.49 ( $M = 32.93, SD = 13.42$ ) was significantly

**Table 5.** Results for tests of between-subjects effects.

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared
Corrected Model	47379.104 <sup>a</sup>	11	4307.191	29.631	0.000	0.091
Intercept	2221987.054	1	2221987.054	15286.020	0.000	0.824
SES	7911.201	2	3955.601	27.212	0.000	0.016
GPA	28539.860	3	9513.287	65.446	0.000	0.057
SES * GPA	922.360	6	153.727	1.058	0.386	0.002
Error	474166.712	3262	145.361			
Total	3147272.000	3274				
Corrected Total	521545.817	3273				

<sup>a</sup>R Squared = 0.091 (Adjusted R Squared = 0.088).

greater than the participants with GPA between 3.00 and 3.49 ( $M = 27.81$ ,  $SD = 12.27$ ) and between 2.50 and 2.99 ( $M = 25.48$ ,  $SD = 11.03$ ), and below 2.50 ( $M = 24.23$ ,  $SD = 10.92$ ). The participants with GPA between 3.00 and 3.49 ( $M = 27.81$ ,  $SD = 12.27$ ) outperformed on Thai-READS than the participants with GPA between 2.50 and 2.99 ( $M = 25.48$ ,  $SD = 11.03$ ), and below 2.50 ( $M = 24.23$ ,  $SD = 10.92$ ). However, the participants with GPA between 2.50 and 2.99 ( $M = 25.48$ ,  $SD = 11.03$ ) showed no significant difference on Thai-READS as compared to the participants with GPA below 2.50 ( $M = 24.23$ ,  $SD = 10.92$ ).

The results in Table 5 also point out that there was a statistically significant main effect of SES on Thai-READS performance among the participants,  $F(2, 3262) = 27.21$ ,  $p < 0.05$ , with a small effect size ( $\eta^2 = 0.02$ ) as determined by Cohen's (1988) criterion. Hence,  $H_{01}$  and  $H_{02}$  were rejected.

The results of Post-Hoc comparisons using the Tukey HSD test in Appendix C reveal that the participants with better SES outperformed on Thai-READS than the lower SES. For example, the Thai-READS performance of the participants with monthly income of above 40,000 Baht ( $M = 31.27$ ,  $SD = 13.21$ ) was significantly greater than the participants with monthly income between 20,000 and 40,000 Baht ( $M = 27.36$ ,  $SD = 12.22$ ) and below 20,000 Baht ( $M = 26.06$ ,  $SD = 11.74$ ). It was also found that Thai-READS average score of the participants with monthly income between 20,000 and 40,000 Baht ( $M = 27.36$ ,  $SD = 12.22$ ) was significantly greater than the participants with monthly income of below 20,000 Baht ( $M = 26.06$ ,  $SD = 11.74$ ).

## Phase 2: interaction effect of GPA and SES on Thai-READS performance

A two-way analysis of variance (ANOVA) was conducted to examine the interaction effect between high-school GPA and socioeconomic status (SES) on students' performance in the computer-based Thai-READS English reading assessment. The analysis included 3,274 undergraduate participants and tested whether the effect of GPA on reading performance varied across different SES backgrounds.

The results indicated that the interaction effect between GPA and SES was not statistically significant,  $F(6, 3262) = 1.06$ ,  $p = 0.386$ , with a very small effect size (partial  $\eta^2 = 0.002$ ). Therefore,  $H_{03}$  failed to be rejected. This finding suggests that the influence of high-school GPA on Thai-READS performance did not differ meaningfully across SES groups. In other words, the relationship between academic achievement and English reading proficiency was consistent regardless of students' socioeconomic background.

Although the interaction effect was not significant, the overall model was statistically significant,  $F(11, 3262) = 29.63$ ,  $p < 0.001$ , explaining approximately 9.1% of the variance in Thai-READS performance (adjusted  $R^2 = 0.088$ ). As reported previously, significant main effects were observed for both GPA and SES, indicating that each variable independently contributed to variation in reading performance. However, the absence of a significant interaction effect implies that these contributions operated additively rather than synergistically.

The plot in Appendix D illustrates a consistent increase in Thai-READS scores as GPA levels rise across all SES groups. Students from higher socioeconomic backgrounds demonstrate higher estimated Thai-READS scores at each GPA level, with the greatest divergence observed in the highest GPA group, indicating a potential interaction between academic achievement and socioeconomic status.

Taken together, these results indicate that while both academic achievement and socioeconomic status are important predictors of English reading proficiency in a computerised assessment context, SES does not moderate the effect of GPA on Thai-READS performance. This finding supports the robustness

of GPA as an indicator of reading proficiency across diverse socioeconomic groups and suggests that the computerised Thai-READS assessment functions similarly across combinations of GPA and SES levels.

## Discussion and conclusion

Building on the absence of a significant interaction effect between GPA and socioeconomic status (SES), the findings suggest that the computerised Thai-READS assessment functions consistently across combinations of academic achievement and socioeconomic background, supporting its robustness as a large-scale assessment tool. Against this backdrop, the present study examined undergraduates' English reading proficiency by integrating neurocognitive evidence from EEG, behavioural performance in a computerised assessment, and key learner background variables, namely GPA and SES. By combining experimental, large-scale quantitative, and qualitative data, the findings provide a comprehensive perspective on the cognitive, psychometric, and equity-related implications of computer-based English reading assessment in Thai higher education.

At the behavioural level, the absence of significant differences in Thai-READS scores between paper-based testing (PBT) and computer-based testing (CBT) supports the comparability of performance outcomes across test modes. This finding aligns with previous research suggesting that well-designed CBT systems can yield results equivalent to traditional paper formats. However, EEG analyses revealed a more nuanced pattern. While most neurocognitive indicators showed no significant differences between CBT and PBT, a significant increase in alpha frequency was observed for difficult items answered incorrectly in the CBT condition. This suggests that under heightened task difficulty, CBT may elicit subtle differences in cognitive regulation or attentional processing that are not captured by test scores alone. These findings reinforce psychometric arguments that score equivalence does not necessarily imply equivalence in underlying cognitive processes and highlight the value of EEG in examining test mode effects beyond behavioural outcomes.

The qualitative findings further contextualised these neurocognitive results. Students generally perceived CBT as efficient, organised, and supportive of focus, which is consistent with EEG evidence indicating stable attentional engagement during CBT. At the same time, concerns related to technical reliability, screen-based navigation, and limited opportunities for annotation underscore potential sources of anxiety or cognitive friction, particularly for learners less accustomed to digital reading environments. Together, these findings suggest that CBT is largely feasible and acceptable, but its cognitive demands may vary depending on task difficulty and user familiarity.

The findings from the large-scale analysis in Phase 2 revealed significant main effects of both grade point average (GPA) and socioeconomic status (SES) on Thai-READS performance, with GPA emerging as the stronger predictor of English reading proficiency. This pattern highlights the central role of sustained academic preparation and effective learning strategies in supporting reading achievement, while also reaffirming the contribution of sociocultural and economic resources to language development. Importantly, no significant interaction was observed between GPA and SES, indicating that the relationship between academic achievement and reading proficiency is broadly consistent across socioeconomic groups. This result is consistent with Tan's (2024) synthesis, which emphasises the strong influence of academic and instructional factors alongside, rather than subsumed by, socioeconomic conditions in shaping student learning outcomes. Taken together, these findings provide evidence for the fairness of the computerised Thai-READS assessment, suggesting that it does not systematically advantage or disadvantage students on the basis of SES when differences in academic preparation are taken into account.

The global relevance of these findings extends beyond Thailand's national policy context and deserves explicit consideration. The significant main effects of GPA and SES on reading proficiency, together with the absence of a moderation effect between these variables, provide internationally relevant evidence for the design of equitable computerised assessments. PISA 2018 data consistently show that SES is one of the strongest predictors of reading performance across participating countries, including emerging economies in Southeast Asia (OECD, 2019). However, the findings of the present study suggest that a well-designed CBT system can demonstrate construct-relevant sensitivity across diverse learner backgrounds without intensifying existing socioeconomic inequalities. This outcome is particularly important in the current global context, where concerns about digital inequality in assessment have increased following the widespread adoption of online and computer-based testing during and

after the COVID-19 pandemic (Litchfield et al., 2021; Miras et al., 2023). The empirical evidence that CBT did not systematically disadvantage lower-SES students in the Thai context therefore contributes to the growing international evidence base supporting the responsible scaling of digital assessment in developing and middle-income countries, including ASEAN member states, where uneven digital infrastructure continues to create unequal conditions for CBT implementation.

At a broader level, the study's two-phase mixed-methods framework, which integrates neural, behavioural, and large-scale psychometric evidence, reflects the comprehensive approach to assessment validation recommended in internationally recognised testing standards (American Educational Research Association (AERA) et al., 2014). The World Bank and UNESCO have consistently highlighted the inadequacy of reading proficiency data in low- and middle-income countries, while the 2024 SDG 4 Extended Report indicated that only 58% of students globally achieved minimum reading proficiency at the end of primary schooling in 2019. Against this backdrop, validated, scalable, and cognitively sensitive assessment instruments such as the computerised Thai-READS system offer a potentially replicable model for national contexts seeking to develop technically sound digital assessment tools. The methodological comprehensiveness of the present study strengthens its potential contribution as a reference model for EFL assessment practice across Asia, Africa, and Latin America, where national language assessments are increasingly being transferred to digital platforms without sufficient empirical examination of their construct validity or cognitive equivalence across diverse learner groups (Delgado et al., 2018; Keng et al., 2008).

In conclusion, this study demonstrates that computer-based English reading assessment can function as a valid and equitable alternative to paper-based testing in higher education. By integrating EEG evidence with performance data and learner characteristics, the study advances understanding of how assessment modality, cognitive processing, and contextual factors jointly shape reading proficiency. The findings offer practical implications for assessment design, institutional policy, and the implementation of digital testing systems in EFL contexts. Future research should build on this work by employing longitudinal designs and additional cognitive measures to further explore learners' engagement with technology-enhanced assessments in increasingly digital learning environments.

The implications of this study align closely with Sustainable Development Goal 4 (SDG 4), which emphasises inclusive, equitable, and quality education, as well as with Thailand's national higher-education and digital transformation policies. Evidence of comparable performance outcomes between computer-based and paper-based English reading assessments, with no observed disadvantage for students from lower socioeconomic backgrounds, is consistent with the aims of SDG 4.1 and 4.3 in promoting equity in higher education assessment (Jie, 2022). At the national level, these findings are consistent with Thailand's 20-Year National Strategy (2018–2037), which prioritises human capital development, educational quality, and the integration of digital technologies to enhance competitiveness and reduce inequality (Office of the National Economic & Social Development Council [NESDC], 2018). Moreover, the results directly support policy directions of the Ministry of Higher Education, Science, Research and Innovation (MHESI), which advocates for digital transformation, technology-enhanced assessment, and data-informed quality assurance in higher education. By showing that computerised assessment can be both cognitively manageable and fair across GPA and SES groups while also highlighting the need for careful interface design and learner support, this study provides empirical evidence to guide the responsible implementation of CBT in line with national policy goals. Importantly, the conservative use of EEG as a complementary tool for understanding cognitive engagement reinforces policy efforts to promote innovation grounded in pedagogical validity rather than technological adoption alone.

### **Limitations and future research directions**

The limitations of the present study should be acknowledged. First, the participant sample was drawn from a single university context, which may limit the generalisability of the findings to other institutional settings. Differences in instructional practices, levels of technological familiarity, and patterns of English language exposure across universities and disciplines may influence both reading performance and cognitive engagement during computerised assessment.

Second, EEG data were collected using a single-channel device (MindWave Mobile 2: NeuroSky Inc.; San Jose, USA), selected for its practicality and suitability for assessment-based research contexts. Although this system does not provide the spatial resolution associated with multi-channel research-grade EEG equipment, it allows for the examination of general patterns of mental workload and attentional engagement during reading tasks. As with other EEG-based approaches, measurements may be influenced by individual variability, testing conditions, and transient affective or attentional states. In addition, the analysis was limited to a small set of EEG indicators, which offers only a partial representation of the neurocognitive processes involved in reading comprehension.

Finally, the cross-sectional design of the study constrains causal interpretations of the relationships observed among EEG responses, English reading proficiency, grade point average (GPA), and socioeconomic status (SES). The SES variables included were based on broad background information and may not fully capture the complexity of learners' educational and technological environments.

Future research could extend the present work by including larger and more diverse samples across multiple institutions in order to enhance generalisability. Longitudinal or experimental designs that directly compare paper-based and computer-based testing conditions would provide stronger evidence regarding the causal mechanisms underlying test mode effects on cognitive and neural processing. Further methodological refinement may also be achieved through the use of multi-channel EEG systems in combination with complementary measures such as eye-tracking or functional neuroimaging. Finally, the inclusion of more detailed indicators of learners' socioeconomic and technological backgrounds would support a more comprehensive account of how individual and contextual factors interact to shape English reading proficiency in computerised assessment contexts.

## Implications

The findings of this study have practical implications for teaching, assessment, and research in higher education. The use of EEG highlights its potential as a complementary tool for examining learners' cognitive effort and attentional engagement during reading tasks. While EEG is not intended for routine classroom use, insights derived from neurocognitive data can inform the design of reading assessments and instructional activities that better align task demands with students' cognitive capacities.

Differences observed between computer-based testing (CBT) and paper-based testing (PBT) suggest that assessment mode should be treated as a substantive design consideration rather than a neutral delivery choice. Computer-based formats may introduce additional cognitive demands related to screen-based reading and interface interaction, which can influence performance independently of language proficiency. Institutions implementing digital assessments should therefore attend to interface design, task length, and students' prior experience with computerised testing to support fairness and score interpretability.

The effects associated with grade point average (GPA) and socioeconomic status (SES) further indicate that learner background plays a meaningful role in shaping both reading performance and cognitive effort. These findings reinforce the need for differentiated instructional and assessment support, particularly for students from lower-achievement or lower-SES backgrounds who may experience increased cognitive load in technology-enhanced testing environments. From a policy perspective, the results point to the importance of considering student diversity when scaling up computerised assessments in higher education.

Future research should extend this work by examining how instructional interventions informed by cognitive load and attentional patterns influence reading development over time. Longitudinal and multi-institutional studies would help establish the robustness of EEG-informed insights and clarify their value for educational decision-making. Such work would contribute to a more evidence-based integration of digital assessment tools in higher education contexts.

In sum, this study contributes to applied educational research by linking neurocognitive evidence with assessment performance and learner characteristics in an EFL context. By situating neural measures within real classroom and institutional settings, the findings offer actionable guidance for educators, assessment designers, and policymakers seeking to improve the validity, equity, and effectiveness of English reading assessment in diverse student populations.

## Acknowledgments

The authors sincerely thank the students, faculty members, and employers who participated in this research and generously shared their insights. We also gratefully acknowledge the School of Liberal Arts, King Mongkut's Institute of Technology Ladkrabang, for providing funding support, and the College of Biomedical Engineering, Rangsit University, for providing research resources. The authors further extend their appreciation to the anonymous reviewers for their valuable feedback and constructive suggestions, which significantly helped improve the quality of the manuscript.

## Author contributions

CRedit: **Thapanee Khemanuwong**: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing; **Thanate Angsuwatanakul**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **Ekkapon Phairot**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **Shaik Abdul Malik Mohamed Ismail**: Conceptualization, Methodology, Writing – original draft, Writing – review & editing; **Nagorn Bunyarit**: Conceptualization, Data curation, Methodology, Visualization, Writing – original draft, Writing – review & editing.

## Informed consent statement

Informed consent was obtained from all subjects involved in the study.

## Review board statement

Full board ethical approval (EC-KMITL\_69\_043) was granted from the KMITL Research Ethics Committee at King's Mongkut Institute of Technology Ladkrabang, Thailand.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This research was funded by the School of Liberal Arts, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand (2563–0219006).

## About the authors

**Thapanee Khemanuwong**, PhD, earned her Bachelor's degree in English Language Education (Second Class Honors) from Chulalongkorn University, Thailand. She completed both her master's and doctoral degrees in English Language Education/TESOL at Universiti Sains Malaysia. She is currently a lecturer at the School of Liberal Arts, King Mongkut's Institute of Technology Ladkrabang. Her research interests include instructional coaching, English language assessment, reading comprehension, English Language Teaching (ELT), English for Specific Purposes (ESP), TESOL education, artificial intelligence in language learning, EEG-based comparisons of computer-based and paper-based testing, and neurocognitive approaches to language assessment.

**Thanate Angsuwatanakul**, PhD, received a Bachelor of Engineering degree (B.E.) in Electrical Engineering in 2001 and a Master of Engineering degree (M.E.) in Biomedical Electronics Engineering in 2007 from King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. In March 2020, he received a Doctor of Philosophy degree (Ph.D.) in Biomedical Engineering from Kyushu University, Fukuoka, Japan. From 2008 to present, he has been an instructor of College of Biomedical Engineering, Rangsit University, Bangkok, Thailand. His research interests include biomedical electronics, brain function analysis, neuroinformatics, medical signal and image processing, and EEG-based comparisons of computer-based and paper-based testing.

**Ekkapon Phairot**, PhD, has currently worked as a lecturer at the Faculty of Education, Songkhla Rajabhat University. He earned his Ph.D. from Universiti Sains Malaysia. His academic work spans several interconnected areas; including quantitative research in education, English language teaching, educational neurotechnology, and artificial intelligence integration in EFL classrooms. His research in English language teaching focuses particularly on Thai learners'

communication, English for Specific Purposes (ESP), teaching strategies in ESL and EFL contexts, and classroom interaction. He is also interested in EEG-based comparisons of computer-based and paper-based testing, as well as the pedagogical applications of AI in language learning environments.

**Shaik Abdul Malik Mohamed Ismail**, PhD, is an Honorary Associate Professor at the School of Educational Studies, Universiti Sains Malaysia (USM). He also serves as a Senior Research Fellow at the Industry and Community Network Division of USM. Dr. Ismail earned his PhD in Curriculum and Instruction from the University of Wisconsin-Madison, USA. His areas of expertise include curriculum and instruction, English language studies, and educational administration. His research primarily focuses on reading benchmarking, reading assessment, curriculum development, educational management and leadership, and EEG-based comparisons of computer-based and paper-based testing.

**Nagorn Bunyarit**, earned his Bachelor's and Master's degrees in English Language and Linguistics. He is currently a lecturer at the College of Industrial and Innovation Management, King Mongkut's Institute of Technology Ladkrabang. His research interests include curriculum design, English language assessment, teaching strategies in EFL contexts, EEG-based comparisons of computer-based and paper-based testing, and CEFR-based education.

## ORCID

Thapanee Khemanuwong  <http://orcid.org/0000-0002-6655-0676>

Thanate Angsuwatanakul  <http://orcid.org/0000-0001-6155-6529>

Ekkapon Phairot  <http://orcid.org/0009-0004-5218-7222>

Shaik Abdul Malik Mohamed Ismail  <http://orcid.org/0000-0002-5314-0774>

Nagorn Bunyarit  <http://orcid.org/0009-0005-2571-6413>

## Data availability statement

The datasets generated and/or analysed during the current study are not publicly available but are available from the first, the second, and the corresponding authors upon reasonable request.

## References

- Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press. DOI: [10.1017/CBO9780511732935](https://doi.org/10.1017/CBO9780511732935)
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Alnemari, M. (2017). *Integration of a low-cost EEG headset with the internet of thing framework* [Doctoral dissertation]. UC Irvine.
- Anakwe, B. (2008). Comparison of student performance in paper-based versus computer-based testing. *Journal of Education for Business*, 84(1), 13–17. <https://doi.org/10.3200/JOEB.84.1.13-17>
- Angsuwatanakul, T., Inthaphuk, N., O'Reilly, J. A., Ismail, S. A. M. M., Chantragatrawi, C., & Khemanuwong, T. (2021). Brain function analysis of computer-based and paper-based tests using EEG: A pilot study. In *Globalization revisited: Building organization resilience with digital transformation: Proceedings of the 4th PIM international conference*, 3 March 2021.
- Angsuwatanakul, T., Iramina, K., & Kaewkamerdpong, B. (2014). Multi-scale sample entropy as a feature for working memory study. In *The 7th 2014 biomedical engineering international conference*, November (pp. 1–5). IEEE.
- Angsuwatanakul, T., O'Reilly, J., Ounjai, K., Kaewkamerdpong, B., & Iramina, K. (2020). Multiscale entropy as a new feature for EEG and fNIRS analysis. *Entropy*, 22(2), 189. <https://doi.org/10.3390/e22020189>
- Barani, R. (2022). A comparison of paper-based tests and computer-based tests in the emergence of COVID-19 era: Evidence by event-related potentials. *Journal of Research in Curriculum Instruction and Educational Technology*, 8(3), 15–41. <https://doi.org/10.21608/jrciet.2022.251885>
- Bennett, R. E. (2003). Online assessment and the comparability of score meaning. *Educational Measurement: Issues and Practice*, 22(1), 5–12. <https://doi.org/10.1111/j.1745-3992.2003.tb00101.x>
- Berka, C., et al. (2007). EEG correlates of task engagement and mental workload. *Biological Psychology*, 75(2), 137–152. <https://doi.org/10.1016/j.biopsycho.2007.01.006>
- Bodmann, S. M., & Robinson, D. H. (2004). Speed and performance differences among computer-based and paper-pencil tests. *Journal of Educational Computing Research*, 31(1), 51–60. <https://doi.org/10.2190/GRQQ-YT0F-7LKB-F033>
- Boontong, N., Pakladapongsin, A., Metee, J., Japang, P., & Khemanuwong, T. (2026). Enhancing conversational competence with short simulation-based clips in a Thai university first-year EFL course. *International Journal of Learning, Teaching and Educational Research*, 25(2), 217–242. <https://doi.org/10.26803/ijlter.25.2.11>
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53(1), 371–399. <https://doi.org/10.1146/annurev.psych.53.100901.135233>

- Chalhoub-Deville, M., & Deville, C. (1991). Computer-Adaptive testing in second language contexts. *Annual Review of Applied Linguistics*, 19, 273–299.
- Chapelle, C. A., & Voss, E. (2017). Utilizing technology in language assessment. In *Language testing and assessment* (pp. 149–161). Springer.
- Chaiyanan, C., Angsuwatanakul, T., Iramina, K., & Kaewkamnerdpong, B. (2025). Simultaneous EEG-fNIRS data on learning capability via implicit learning induced by cognitive tasks. *Data*, 10(8), 131. <https://doi.org/10.3390/data10080131>
- Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: Key factors associated with the test mode effect. *British Journal of Educational Technology*, 33(5), 593–602. <https://doi.org/10.1111/1467-8535.00294>
- Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Crivelli, D., Fronda, G., Venturella, I., & Balconi, M. (2019). Supporting mindfulness practices with brain-sensing devices. Cognitive and electrophysiological evidences. *Mindfulness*, 10(2), 301–311. <https://doi.org/10.1007/s12671-018-0975-3>
- Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on comprehension. *Educational Research Review*, 25, 23–38. <https://doi.org/10.1016/j.edurev.2018.09.003>
- Deutsch, T., Herrmann, K., Frese, T., & Sandholzer, H. (2012). Implementing computer-based assessment—A web-based mock examination changes attitudes. *Computers & Education*, 58(4), 1068–1075. <https://doi.org/10.1016/j.compedu.2011.11.013>
- Dooley, P. (2008). Language testing and technology: Problems of transition to a new era. *ReCALL*, 20(1), 21–34. <https://doi.org/10.1017/S0958344008000311>
- Ebrahimi, M. R., Toroujeni, S. M. H., & Shahbazi, V. (2019). Score equivalence, gender difference, and testing mode preference in a comparative study between computer-based testing and paper-based testing. *International Journal of Emerging Technologies in Learning (IJET)*, 14(07), 128. <https://doi.org/10.3991/ijet.v14i07.10175>
- Gkintoni, E., & Halkiopoulos, C. (2025). Mapping EEG metrics to human affective and cognitive models: An interdisciplinary scoping review from a cognitive neuroscience perspective. *Biomimetics*, 10(11), 730. <https://doi.org/10.3390/biomimetics10110730>
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Brookes.
- Hui, K. S., Saeed, K. M., & Khemanuwong, T. (2020). Reading comprehension ability of future engineers in Thailand. *MEXTESOL Journal*, 44(4), 1–21.
- Jamil, M., Tariq, R. H., & Shami, P. A. (2012). Computer-based vs paper-based examinations: Perceptions of university teachers. *Turkish Online Journal of Educational Technology-TOJET*, 11(4), 371–381.
- Jie, H. U. (2022). *Assessing students' digital reading performance: An educational data mining approach*. Routledge.
- Karim, A., Khemanuwong, T., Ferdous, T., Humaira, T., Sultana, S., & Ismail, S. A. M. M. (2024). Do gender, socio-economic status and regional difference predict the reading comprehension of English as a foreign language (EFL) learners? *MEXTESOL Journal*, 48(4), 1–27.
- Keng, L., McClarty, K. L., & Davis, L. L. (2008). Item-level analysis of online versus paper administrations. *Applied Measurement in Education*, 21(3), 207–226. <https://doi.org/10.1080/08957340802161779>
- Khemanuwong, T., Angsuwatanakul, T., Phairot, E., Iramina, K., Puttasakul, T., & Bunyarit, N. (2024). Brain function analysis using EEG evidence: New insights into English paper-based versus computer-based tests. In *2024 16th biomedical engineering international conference (BMEiCON)*, Chon Buri, Thailand (pp. 1–5). IEEE. <https://doi.org/10.1109/BMEiCON64021.2024.10896313>
- Khemanuwong, T., Mohamed, A. R., & Ismail, S. A. M. M. (2018). Developing a Thai READS encoder to gauge EFL reading proficiency of Thai undergraduate students. *Journal of Teaching and Learning English in Multicultural Contexts (TLEMC)*, 2(1), 23–34.
- Khoshshima, H., Hosseini, M., & Toroujeni, S. M. H. (2017). Cross-mode comparability of computer-based testing (CBT) versus paper-pencil based testing (PPT): An investigation of testing administration mode among Iranian intermediate EFL learners. *English Language Teaching*, 10(2), 23–32. <https://doi.org/10.5539/elt.v10n2p23>
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance. *Brain Research. Brain Research Reviews*, 29(2-3), 169–195. [https://doi.org/10.1016/S0165-0173\(98\)00056-3](https://doi.org/10.1016/S0165-0173(98)00056-3)
- Klimesch, W. (2012).  $\alpha$ -band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617. <https://doi.org/10.1016/j.tics.2012.10.007>
- Langensee, L., Rumetshofer, T., & Mårtensson, J. (2024). Interplay of socioeconomic status, cognition, and school performance in the ABCD sample. *NPJ Science of Learning*, 9(1), 17. <https://doi.org/10.1038/s41539-024-00233-x>
- Lokare, V. T., & Jadhav, P. M. (2026). Improving online learner comprehension via EEG signal analysis. *Biomedical Engineering: Applications, Basis and Communications*, 38(04), 2550044. <https://doi.org/10.4015/S1016237225500449>
- Litchfield, I., Shukla, D., & Greenfield, S. (2021). Impact of COVID-19 on the digital divide: A rapid review. *BMJ Open*, 11(10), e053440. <https://doi.org/10.1136/bmjopen-2021-053440>
- Miras, S., Ruiz-Bañuls, M., Gómez-Trigueros, I. M., & Mateo-Guillen, C. (2023). Implications of the digital divide: A systematic review of its impact in the educational field. *Journal of Technology and Science Education*, 13(3), 936–950. <https://doi.org/10.3926/jotse.2249>

- Moss, P. G. (2020). Training students for online exams reduces cognitive over-load. Retrieved from <https://paulg-moss.com/2020/05/22/training-students-for-online-examsreduces-cognitive-overload/> (on 12 June 2026)
- Nardi, A., & Ranieri, M. (2019). Comparing paper-based and electronic multiple-choice examinations with personal devices: Impact on students' performance, self-efficacy and satisfaction. *British Journal of Educational Technology*, 50(3), 1495–1506. <https://doi.org/10.1111/bjet.12644>
- Noyes, J. M., & Garland, K. J. (2008). Computer-vs. paper-based tasks: Are they equivalent? *Ergonomics*, 51(9), 1352–1375. <https://doi.org/10.1080/00140130802170387>
- Ockey, G. J. (2009). Developments and challenges in the use of computer-based testing for assessing second language ability. *The Modern Language Journal*, 93(s1), 836–847. <https://doi.org/10.1111/j.1540-4781.2009.00976.x>
- OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed*. OECD Publishing.
- Office of the National Economic and Social Development Council [NESDC]. (2018). *National strategy (2018–2037)*. National Strategy Secretariat Office. [in Thai]
- Paek, I., & Han, K. T. (2013). IRTPRO and invariance logic in assessment. *Applied Psychological Measurement*, 37(3), 242–252.
- Phairot, E., Khemanuwong, T., Boonthong, N., Angsuwatanakul, T., & Uampittaya, P. (2022). Examining gender and socioeconomic status predictability in English reading comprehension among Thai EFL undergraduates. *Journal of Faculty of Education Pibulsongkram Rajabhat University.*, 9(1), 96–111.
- Pholphirul, P., Rukumnuaykit, P., & Teimtd, S. (2023). Teacher shortages and educational outcomes in developing countries: Empirical evidence from PISA-Thailand. *Cogent Education*, 10(2), 1–26. <https://doi.org/10.1080/2331186X.2023.2243126>
- Pittau, F., Grouiller, F., Spinelli, L., Seeck, M., Michel, C. M., & Vulliemoz, S. (2014). The role of functional neuroimaging in pre-surgical epilepsy evaluation. *Frontiers in Neurology*, 5(1), 31–45. <https://doi.org/10.3389/fneur.2014.00031>
- Prisacari, A. A., & Danielson, J. (2017). Computer-based versus paper-based testing: Investigating testing mode with cognitive load and scratch paper use. *Computers in Human Behavior*, 77, 1–10. <https://doi.org/10.1016/j.chb.2017.07.044>
- Ramírez-Moreno, M. A., Díaz-Padilla, M., Valenzuela-Gómez, K. D., Vargas-Martínez, A., Tudón-Martínez, J. C., Morales-Menendez, R., Ramírez-Mendoza, R. A., Pérez-Henríquez, B. L., & Lozoya-Santos, J. J. (2021). EEG-based tool for prediction of university students' cognitive performance in the classroom. *Brain Sciences*, 11(6), 698. <https://doi.org/10.3390/brainsci11060698>
- Sawaki, Y. (2001). Comparability of conventional and computerized tests of reading in a second language. *Language Learning & Technology*, 5(2), 38–59. <https://doi.org/10.64152/10125/25127>
- Shudong Wang, Hong Jiao, Young, M. J., Brooks, T., & Olson, J. (2008). Comparability of Computer-Based and Paper-and-Pencil Testing in K–12 Reading Assessments. *Educational and Psychological Measurement*, 68(1), 5–24. <https://doi.org/10.1177/0013164407305592>
- Snow, C. E. (2002). *Reading for understanding*. RAND.
- Stanovich, K. E. (2009). Matthew effects in reading. *Journal of Learning Disabilities*, 42(5), 387–409.
- Straker, L., Pollock, C., Burgess-Limerick, R., Skoss, R., & Coleman, J. (2008). The impact of computer display height and desk design on muscle activity during information technology work by young adults. *Journal of Electromyography and Kinesiology: official Journal of the International Society of Electrophysiological Kinesiology*, 18(4), 606–617. <https://doi.org/10.1016/j.jelekin.2006.09.015>
- Tan, C. Y. (2024). Socioeconomic status and student learning: Insights from an umbrella review. *Educational Psychology Review*, 36(4), 1–63. <https://doi.org/10.1007/s10648-024-09929-3>
- Taylor, C., Kirsch, I., Eignor, D., & Jamieson, J. (1999). Examining the relationship between computer familiarity and performance on computer-based language tasks. *Language Learning*, 49(2), 219–274. <https://doi.org/10.1111/0023-8333.00088>
- Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, 2(2), 1–11.
- Weiss, S., & Mueller, H. M. (2003). The contribution of EEG coherence to the investigation of language. *Brain and Language*, 85(2), 325–343. [https://doi.org/10.1016/s0093-934x\(03\)00067-1](https://doi.org/10.1016/s0093-934x(03)00067-1)
- Wise, S. L. (2019). Controlling construct-irrelevant factors through computer-based testing: Disengagement, anxiety, & cheating. *Education Inquiry*, 10(1), 21–33. <https://doi.org/10.1080/20004508.2018.1490127>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner. *Theory Into Practice*, 41(2), 64–70. [https://doi.org/10.1207/s15430421tip4102\\_2](https://doi.org/10.1207/s15430421tip4102_2)