The Review on Eye-Tracking Studies in L2 Assessment

Caoxi[1] (iD)
Ma Zhenni[2] (iD)

Contributor roles: Caoxi: Writing- Original draft preparation. Ma Zhenni: Writing- Reviewing and Editing

Abstract

The use of eye-tracking methodology in L2 assessments has become increasingly common, focusing on examining the cognitive validity of tests and the processing patterns of test takers. This paper reviews both empirical and theoretical studies on eye-tracking in L2 assessments and its applications in language education research. Through the lens of cognitive psychology, the paper examines the fundamental principles and primary applications of eye-tracking in L2 assessment. It also discusses potential theoretical and pedagogical implications. Finally, it highlights the limitations of current studies in this field and offers suggestions for future research to broaden the scope of eye- tracking studies in language testing and assessment.

Keywords: Language testing; eye tracking; cognitive validity; language performance

---

1   Tangshan University. caoxi815@shisu.edu.cn

2   Xinjiang Normal University

Resumen de la investigación de seguimiento del movimiento ocular en la evaluación de la segunda lengua

Resumen

La aplicación del método de seguimiento del Movimiento ocular en la evaluación de la segunda lengua es cada vez más común, centrándose en comprobar la efectividad cognitiva de las pruebas y el modo de procesamiento de los candidatos. Este trabajo resume la investigación empírica y teórica del rastreo ocular en la evaluación de la segunda lengua y su aplicación en la investigación de la educación lingüística. Se exploran los principios básicos y los principales usos del rastreo ocular en la evaluación de la segunda lengua desde la perspectiva de la psicología cognitiva. Se discuten también posibles implicaciones teóricas y didácticas. Finalmente, este trabajo destaca las limitaciones de la investigación actual en este campo y ofrece recomendaciones para futuras investigaciones para ampliar el alcance de la investigación de seguimiento ocular en pruebas y evaluaciones de lenguaje.

Palabras clave: prueba de lenguaje; seguimiento del movimiento ocular; eficacia cognitiva; expresión lingüística

Introduction

Eye tracking, the concurrent registration of eye movements, is widely employed in psychological research, where it has been said to provide the 'gold criterion' in psycholinguistic research (Rayner, 2009; Godfroid *et al*., 2020). Eye tracking technology has been utilized in language processing research due to its controlled nature, offering insights into attention and various cognitive behaviors of learners. From 2013 to 2022, there has been a rise in studies exploring the use of eye-tracking technology in L2 assessments, including cognitive validity, correlation between viewing patterns and language performance, text readability, attention to visual cues in listening tests, and reading behaviors of students with varying comprehension levels. These studies, situated at the intersection of psycholinguistics, second language acquisition, and cognitive sciences, demonstrate the diverse and rich landscape of contemporary eye-tracking research in language testing and assessment, engaging with related disciplines.

Researchers employed eye-tracking technology to explore the depths of processing patterns of test takers and the cognitive validity of testing because eye-tracking offers millisecond-precise information on what a participant is visually attending to and most likely processing at any given time (Godfroid *et al*., 2020). In doing so, eye-movement data can advance researchers' understandings of the performance of test takers, as well as how different types of test items—such as multiple-choice questions and the argumentative writing—can influence the cognitive behaviors of test takers. Eye-movement data can assist researchers in comprehending how test takers interact with testing materials, and how this interaction may vary based on learners' backgrounds, task instructions, item types, and test delivery methods. Findings from eye-tracking studies can provide valuable insights for teachers and test developers in enhancing the design and planning of testing materials to optimize learning outcomes.

In this paper, we review studies on the use of eye-tracking technologies in L2 assessments, as well as their applications in test design and pedagogical practices. Eye-tracking researchers have expanded the scope of language testing by incorporating other data-collection methods such as recall interviews, keystroke logging, and speak-aloud prototypes for data triangulation. We hope that this review will offer valuable insights for future research in this area.

Eye-tracking Technology in Second Language Studies

Eye-tracking technology employs concealed near-infrared illuminators to create corneal reflection patterns, facilitating eye movement measurement. High-speed image sensors within eye-tracking hardware capture data, which is later processed to generate a 3D model of the reader's eye movements. This model determines the pupil's precise location and analyzes illuminator reflections and their coordinates (Al-Edaily *et al*., 2013; Baazeem *et al*., 2021).

Eye-tracking technology provides data on oculomotor events, including total dwell time, fixation rate, fixation counts, and regression counts. Most eye-tracking software can track participants' focus within user-defined areas of interest (AOIs) in the visual field. For complex AOIs, where human perception surpasses software capabilities, manual scan-path analysis may be required (Holmqvist *et al*., 2011). In psychology, the interpretation of eye-tracking data reveals cognitive behaviors, demonstrating a strong link between eye movements and mental processes in language comprehension (e.g., Batty, 2021; Bax a Chan, 2019; Godfroid *et al*., 2020).

Furthermore, eye-tracking has become a valuable alternative to traditional assessment methods, emphasizing cognitive validity and its influence on examinees' performance in reading, listening, and writing assessments, including placement and diagnostic tests. A frequently analyzed variable in L2 assessment studies is the total time spent focusing on specific AOIs (Godfroid, 2019).

Researchers often integrate quantitative data from eye-tracking devices with qualitative insights from verbal reports, such as recall interviews, self-report checklists, and think-aloud protocols. This mixed-methods approach enables a more comprehensive triangulation and interpretation of participants' viewing behaviors.

Eye-Tracking Measures Adopted in L2 Assessment

Eye-tracking is considered a valuable tool for investigating cognitive behaviors during language processing and has been applied to auditory, visual, and multimodal language studies. Eye-tracking systems can provide the location, sequence, and duration of eye movements within areas of interest (AOIs), as well as real-time data on pupil size and blink rates (Holmqvist *et al.*, 2011). Eye-tracking measures can be conceptualized in multiple ways. Lai *et al.* (2013) categorized eye-movement measures based on the scale of measurement and the type of eye movement. Temporal measures quantify gaze behaviors over time and are believed to offer insights into when and for how long cognitive processing occurs, as well as the associated processing load. The spatial scale represents eye movements in spatial terms, focusing on locations, distances, directions, or sequences. Spatial measures offer information on where and how cognitive processing occurs. The count scale quantifies eye movements by number, proportion, rate, and/or frequency. Similarly, in SLA and bilingualism research, Godfroid (2019) classified eye-tracking measures into three main categories: fixation, regression, and eye movement dynamics. According to this classification, there are four subtypes of fixation-based measures: (1) fixation counts, probabilities, and proportions; (2) fixation duration; (3) fixation latency (e.g., the time it takes for a participant to fixate on a specific area of interest); and (4) fixation location. In addition to fixations and saccadic eye movements, measures like pupil size and blink rate have emerged in recent studies to explore visual information processing and investigate cognitive mechanisms (Eckstein *et al.*, 2017).

The results clearly indicate an uneven distribution of the types of eye-tracking measures used in L2 research. While a significant proportion of studies in the sample employed fixation measures, other types of measures, such as regressions, saccades, blinks, and pupil measures, were used much less frequently. The broad application of fixations across various subfields of assessment studies can be attributed to their ability to reveal insights into cognitive operations and online language processing, such as attention, reading processes, and listening processes.

Moreover, this study has found that fixation temporal measures have been predominantly used by researchers, followed by fixation count measures. Specifically, the two research designs most commonly adopted—reading and the visual world paradigm—primarily focus on analyzing fixation temporal and count measures, respectively. In addition, it has been found that the distribution of measure types in this field varies depending on the topics being investigated. Therefore, the selection of eye-tracking measures should be guided by research questions, research designs, and theoretical frameworks.

Eye-tracking Technology in Reading Tests

Eye-tracking technology in studies of reading tests has primarily focused on test validation, text readability, processing of scientific terms in technology-related articles, and the effects of item formats on the interpretation of scores. Compared to listening and writing assessments, reading tests have received considerably more attention. Concerning test types, high-stake tests, such as TOFEL and IELTS gained more attention compared to local-level tests. The survey participants mainly consist of undergraduate and graduate students from universities. Various research studies have utilized eye movement data to analyze test behaviors of secondary school students, including differences in approaches to discourse synthesis. These findings shed light on key issues in L2 assessment research, such as test design, score interpretation validity, and pedagogical practices.

Assessment-related research has also highlighted the potential of using eye-tracking technology to investigate the cognitive validity of reading tests. Specifically, eye-tracking technology can identify the key elements of a text that distinguish cognitive processing between successful and unsuccessful test takers, a crucial aspect for test designers (Yaneva *et al.*, 2021). In the case of evaluation of validity and test takers' proficiency, studies indicated that eye-tracking data provide evidence in score interpretation based on strong cognitive models about the process of test takers. Moreover, Yaneva *et al.* (2021) examined how eye-tracking data can illuminate the comprehensive evaluation of the validity of tests. In their research, they studied how the presence of options impacted the way medical students responded to questions by collecting data on eye-tracking features (dwell time, fixation counts and mean fixation duration). Results showed that the presence of options significantly influenced the approach test takers used in responding, but little evidence was found that examinees constrained the problems by reading the options.

Assuming that eye-movement data accurately represent the mental processes the test is designed to measure, researchers explored Chinese ESL learners' reading behaviors in the iBT TOEFL reading section using eye-tracking data (gaze plots, fixation count, and angles between fixations). Results indicated that vocabulary questions seemingly failed to elicit test takers' inferencing ability that test designers intended to measure, and expeditious reading rarely occurred in the processing of reading. Bax and Chan (2019) investigated the cognitive validity of two level-specific English Proficiency Reading Tests. In their study, investigators recorded twenty-four students' eye-movement data (AOI, fixation duration, fixation count, and visit count) while they were doing cloze and multiple-choice on a computer. Findings indicated the range of cognitive processes elicited by different reading item types at the two levels as well as different reading patterns between stronger and weaker test takers.

Eye-tracking studies have explored test takers' processing of science-related test items, with a particular emphasis on comparing attentional shifts between successful and unsuccessful problem solvers (Lindner *et al.*, 2014). These studies examine information processing across different testing stages and reveal that reading strategies vary depending on the type of problem (Langenfeld *et al.*, 2016). For example, research tracking the eye movements of high- and low-proficiency students while solving interactive problems shows that students employ different processing strategies based on the problem type.

Eye-movement data can also serve as indicators for measuring reading strategies. Wilkinson and Payne (2006) found that during time-limited reading tasks, readers typically focus on the beginning of paragraphs or the opening sections of discourse, then quickly skim through the remainder of the text to locate relevant information. Dolgunsz (2016) demonstrated a positive correlation between attention and reading acquisition in second-language learners. More efficient readers tend to apply vocabulary inference strategies, and their language proficiency and vocabulary knowledge significantly influence the strategies they use in second-language reading.

Recent studies have expanded the application of eye-tracking beyond language learners, examining, for example, how computer programmers read code. Peitek *et al.* (2020) investigated programmers' reading habits by recording the eye movements of 21 programmers (12 novices and 9 experts), revealing that experience influenced readers' behavior and the order in which they processed the text. This finding appears to confirm the conclusions of other researchers, including Wang *et al.* (2018), who also conducted a study on the e-book reading process of 6 university students. Their study showed that participants with prior knowledge spent more time on the target areas of the e-book, engaged in longer reading times, and found the images in these areas more engaging.

Bax and Chan (2019) analyzed the cognitive validity of two level-specific English Proficiency Reading Tests by examining eye-movement data from 24 students completing cloze and multiple-choice tasks on a computer. Their findings highlighted the range of cognitive processes triggered by different reading item types at the two levels and the variations in reading patterns between stronger and weaker test takers.

Baazeem *et al.* (2021) conducted a pilot study to investigate the use of real-time processing data for assessing text readability in Arabic. By analyzing eye-movement data from 41 participants while reading six classical Arabic poems, the researchers identified a strong correlation between eye-movement patterns and readability predictions. The findings demonstrated that eye-tracking data, due to its real-time and precise nature, surpassed linguistic features in effectively capturing the readability level of the text.

Eye-tracking Technology in Reading-While-Listening Tests

In the realm of second language (L2) acquisition, the integration of reading and listening activities has long been recognized as a valuable tool for enhancing language proficiency. The advent of eye-tracking technology has provided researchers with unprecedented insights into the cognitive processes underlying these activities. Reading-While-Listening (RWL) tasks involve presenting written text alongside auditory input, requiring learners to process information from both modalities simultaneously. This dual-modality presentation has been shown to enhance vocabulary acquisition and reading comprehension in L2 learners (Chang, 2009; Webb & Chang, 2014). Eye-tracking studies have further elucidated the mechanisms behind these benefits, revealing that RWL tasks can lead to more efficient word recognition and deeper semantic processing compared to reading-only tasks (Conklin *et al.*, 2020).

However, the cognitive load associated with RWL tasks is a critical factor that influences their effectiveness. Research indicates that the simultaneous processing of written and auditory input can increase cognitive load, potentially leading to divided attention and reduced comprehension (Field, 2015). The alignment of eye movements with auditory input is particularly important; misalignment can disrupt the integration of visual and auditory information, hindering comprehension (Aryadoust, 2019). Therefore, understanding how learners allocate their attention between written text and auditory input is crucial for optimizing RWL tasks.

Reading comprehension in RWL tasks is influenced by the interplay between visual and auditory processing. Eye-tracking studies have shown that learners spend more time reading text and less time processing images in RWL conditions compared to reading-only conditions (Pellicer- Sánchez *et al*., 2018). This suggests that the presence of auditory input can enhance text comprehension by providing additional context and reducing the cognitive load associated with reading (Chang & Millet, 2015). However, the increased cognitive load in RWL tasks can also lead to divided attention, particularly among less proficient learners (Field, 2015). Aryadoust (2019) found that during the first hearing of a listening passage, learners tend to skim through the text, focusing more on the auditory input. In contrast, during the second hearing, learners allocate more attention to the written text, indicating a shift in cognitive strategy. This dynamic allocation of attention suggests that RWL tasks require learners to balance between listening and reading, which can be cognitively demanding.

The use of eye-tracking in RWL research presents several methodological challenges. Ensuring accurate synchronization between auditory and visual stimuli is crucial for valid results (Conklin & Alotaibi, 2023). Additionally, the selection of appropriate materials and the control of variables such as word frequency, sentence length, and syntactic complexity are essential to minimize confounding factors (Clifton *et al*., 2007). Furthermore, the analysis of eye-tracking data requires sophisticated statistical methods to account for the complex interactions between different variables (Godfroid, 2020).

The findings from eye-tracking studies have significant implications for language teaching and learning. RWL tasks can be particularly beneficial for L2 learners, as they offer a multisensory learning experience that enhances vocabulary acquisition and reading comprehension (Chang, 2009). However, the effectiveness of RWL tasks may depend on learners' proficiency levels and the complexity of the materials used. For beginners, RWL tasks should be designed with simpler texts and clear auditory support to avoid overwhelming cognitive load (Lightbown, 1992). Advanced learners, on the other hand, may benefit from more challenging materials that require deeper integration of visual and auditory information (Pellicer-Sánchez *et al*., 2018).

Moreover, the use of RWL tasks in classroom settings should be balanced with other reading activities to ensure a comprehensive learning experience. Teachers can incorporate RWL tasks into their lesson plans to provide learners with varied exposure to the target language, while also encouraging independent reading to develop fluency and autonomy (Day & Robb, 2015). Additionally, the findings from eye-tracking studies can inform the design of language assessments, ensuring that test formats align with the cognitive processes involved in RWL tasks (Aryadoust, 2019).

Eye-tracking Technology in Listening Tests

Eye-tracking technology has also been applied to uncover the cognitive processes, including viewing behaviors, levels of attention in real-time, the effect of the spatial location of options on the processing order, and the correlation between gaze bias and domain knowledge of test takers. In these studies, researchers can infer from gaze behaviors the allocation of attention of examinees when they are completing test items, such as multiple-choice questions and short-video prompts. It has been concluded that listeners tended to focus their attention on objects that are anticipated to appear in the video-based listening tests. (Altmann, 2011) even when they are told to ignore the auditory stream and instead look elsewhere (Salverda & Altmann, 2011).

Eye-tracking has also been used to capture psycholinguistic processing in listening comprehension tests, such as in the studies of the effects of test methods on listening performance and cognitive load and attention to visual cues in L2 listening tests (Batty, 2021). For instance, recent studies have focused on the influence of test methods

on cognitive load and listening performance by examining test takers' brain activity patterns (measured by near-infrared spectroscopy) and gaze behaviors (measured by eye- tracking) in while-listening performance (WLP) and post-listening performance (PLP) formats. As indicated by data on fixation rate, fixation duration, and neuroimaging, the research found that, consistent with both the cognitive load theory (Sweller, 2011) and the study by Aryadoust *et al.* (2020), WLP tests imposed a lighter cognitive load upon test takers than notetaking-while-listening in the PLA-Audio. Additionally, listening-while-answering test items during the WLP tests require less bottom-up and top-down processing than answering test items in the PLP-Question phase. Being the first study that integrated the data of eye-tracking and neuroimaging, this study was able to explain the observed test method's effect on gaze behavioral measures and brain activity patterns.

Research of visual effects on the interpretation of test takers' performance on video-based L2 listening tests has been based on test scores and self-reported verbal data (Wagner, 2010; Ockey, 2007). However, few studies have employed eye-tracking technology to investigate examinees' gaze patterns. To address this gap, Suvorov (2015) conducted a study on eye-movement measures—including fixation rate, dwell rate, and total dwell time—for context and content videos. The data revealed significant differences in fixation rates and total dwell time values, although no significant correlation was found between these eye-tracking measures and test scores. The findings nonetheless demonstrated that test takers utilized both visual and auditory information while listening. With the increasing shift towards online education, there has been an exploration into whether video instruction effectiveness can be evaluated remotely using standard web cameras. By analyzing the synchronized eye movements of 1000 participants through standard web cameras, the study found that students' eye-movement patterns in instructional videos were similar. Furthermore, the study established a correlational relationship between eye movements and test performance, suggesting that stronger students were able to both comprehend the video content better and perform better on tests.

Response formats have been shown to be an influential factor in test takers' performance in listening assessments (Brunfaut, 2016), yet few studies have examined the impact of spatial effects on multiple-choice spatial relations. To address this gap, an eye-tracking study was conducted to investigate how the response order in multiple-choice questions influenced item difficulty. In the first part of the study, it was found that participants focused significantly longer on responses placed higher on the screen, marking the first time such results have been observed. The second part of the study tested whether items are more difficult when the key is placed in later positions in the Aptis Listening Test items. Consistent with previous research (Hohensinn & Baghaei, 2017), the results indicated a direct effect of key position on item difficulty in a sample of 200 live Aptis items. The findings further suggested that the spatial location of the key in multiple-choice listening tests influenced the amount of processing it receives and the item's difficulty. Given the widespread use of multiple-choice tasks in language assessments, these findings are crucial, particularly for tests that randomize response order, and for candidates who, by chance, encounter many keys in later positions.

Eye-tracking Technology in Writing Tests

It has also been shown that studies on L2 reading and listening have steadily increased in recent years. In contrast, eye-tracking has not seen widespread adoption in the study of L2 writing. This may be due to ongoing exploration by writing researchers regarding the applicability of eye-tracking in studying the writing process, with recent studies beginning to demonstrate its affordance in this area (e.g., Ranalli *et al.*, 2019). The analysis of participants' eye movements during composition often combines digital screen recordings with visualizations of eye movements (Rvsz *et al.*, 2019). Under these conditions, the resulting video streams must be manually annotated by creating moment-by- moment segments. This is a time-consuming procedure, and such methodological complexity can hinder the development of L2 eye-tracking writing research. One potential solution is to fix the screen areas where participants can type their texts, but this may diminish the authenticity of experiments. This is particularly true in source-based academic writing—where authors frequently alternate between texts or scroll up and down—engaging in various cognitive processes such as reading, viewing, skimming, scanning, confusion, confirmation, rebuttal, and more, which would be extremely difficult, if not impossible, to disentangle in gaze behavior data. In such cases, writing researchers need to employ multiple methods to investigate these processes; for example, Rvsz *et al.* (2019) combined keystroke logs, eye-tracking, and stimulated recall comments, demonstrating the methodological complexity required in this field of study

Research on second-language (L2) writing encounters challenges as feedback from instructors often focuses on the final written products rather than the various processes involved in their creation, including planning, formulation, and evaluation. Utilizing eye-tracking technology enables researchers to observe writing processes in a naturalistic setting, avoiding additional cognitive burden on participants compared to methods like concurrent think-aloud. By examining students' source use, valuable insights can be gained to explain differences in integrated writing performance, as indicated in previous studies. Therefore, the application of eye-tracking technology has also been used to uncover cognitive processes in writing tasks, revealing underlying motivational issues that may have limited participants' learning.

Zhu *et al.* (2021) investigated the correlation between discourse synthesis skills and overall integrated writing performance in both Chinese (students' L1) and English (their L2). The results of integrated writing tests, eye-movement data, and stimulated interviews revealed that discourse synthesis skills are important indicators of integrated writing performance, supporting the validity of the two writing tasks that assessed underlying discourse synthesis skills. Researchers in the field of writing processes have utilized keystroke logging (KL) and eye-tracking (ET) to analyze and visualize process engagement, aiming to enhance the assessment of L2 writing. In their study, two Chinese L1 students enrolled at a U.S. university were selected as case studies. They completed argumentative writing tasks while a KL-ET system tracked their processes and generated visualizations for personalized tutoring. The findings indicated that eye-tracking technology enabled a deeper understanding of the participants' writing development and identified motivational issues that hindered their learning.

Limitations and Directions for Future Research

This review aimed to investigate research on eye movement in L2 assessments, including listening, reading, and writing tests. By synthesizing this research, we hope that it will provide implications for further research. However, several limitations should be acknowledged. First, this research investigated research on eye-tracking technology in L2 assessments by synthesizing conclusions, methods as well as implications of relevant findings, however, this review only examined L2 assessments with eye-tracking technology, future investigations are suggested to provide an in-depth review on how eye-tracking technology, combined with other neuro-related data, such as neuroimaging data to give a full picture of the cognitive processes of test takers.

Despite significant advancements, several limitations remain prevalent in current eye-tracking research within second language (L2) assessment, highlighting the necessity for further exploration and methodological refinement.

First, the predominance of fixation-based measures in existing studies (e.g., fixation counts and durations) indicates an imbalance, with other potentially insightful metrics, such as saccades, regressions, pupil dilation, and blink rates, underutilized. Such measures can offer deeper insights into cognitive load, emotional arousal, and attentional shifts. Future research should diversify eye-tracking measures, exploring their unique contributions to understanding cognitive processes in L2 assessments.

Second, a substantial proportion of studies have concentrated on reading assessments, particularly high-stakes tests like TOEFL and IELTS. However, eye-tracking studies in listening, writing, and speaking assessments remain limited. Particularly, eye-tracking research in oral proficiency tests and interpreting and translation tasks is virtually non-existent. Future studies should address these gaps, examining whether eye-tracking can effectively capture cognitive processes in these modalities, potentially enriching construct validity evidence.

Third, methodological challenges persist regarding the integration and synchronization of eye- tracking with multimodal data. Accurate synchronization between auditory and visual stimuli in Reading-While-Listening (RWL) tasks, for example, remains challenging yet crucial. Future research should employ advanced synchronization techniques and software to minimize timing discrepancies and enhance the validity of multimodal studies. Moreover, current eye-tracking research predominantly involves controlled laboratory settings, which may limit ecological validity. There is a need for increased research in authentic, real-world contexts, including online and classroom-based assessments, to better reflect naturalistic L2 processing and behavior. The adaptation of remote eye- tracking

technology, which leverages standard web cameras, offers promising opportunities for increased ecological validity and large-scale data collection. In addition, there is a growing—yet still insufficient—integration of eye-tracking data with other neurocognitive methodologies, such as neuroimaging techniques (*f*MRI, EEG) or physiological measures (heart rate variability, galvanic skin response). Future studies should pursue multimethod triangulation to create comprehensive cognitive profiles of L2 learners, thereby enhancing the robustness of cognitive validity evidence.

Integration of eye-tracking technology with artificial intelligence (AI) presents promising avenues for future research. AI-driven analytical methods, such as machine learning algorithms, can be employed to analyze complex and extensive eye-tracking datasets more efficiently, identifying subtle cognitive patterns and behaviors that traditional analyses might overlook. AI technology could also enhance adaptive testing systems, dynamically adjusting test difficulty based on real-time analysis of learners' visual engagement and cognitive load. Furthermore, AI integration facilitates automated analysis of eye-tracking data for immediate feedback, providing learners and educators with actionable insights to tailor instructional strategies effectively.

Finally, participant diversity remains relatively limited in eye-tracking studies, with research primarily involving university students from specific language backgrounds. Future studies should broaden participant demographics to include varied proficiency levels, age groups, linguistic backgrounds, and learning contexts. This diversity would help generalize findings and provide deeper insights into differential cognitive processes across populations.

To sum up, while the integration of eye-tracking technology has profoundly enriched L2 assessment research, addressing these outlined limitations can significantly advance our understanding of cognitive processes underlying language assessment tasks. Future research efforts are encouraged to diversify methodologies, expand assessment modalities, increase ecological validity, integrate complementary neurocognitive measures, and enhance participant diversity, contributing to the robust evolution of eye-tracking studies in applied linguistics.

Conclusion

The integration of eye-tracking technology within applied linguistics has significantly reshaped research paradigms, opening pathways to more nuanced understanding of language processing and learner cognition. Eye-tracking, which records ocular movements, fixations, and saccades, has emerged as an invaluable methodological tool in applied linguistics research, particularly due to its capacity to yield direct, fine-grained insights into cognitive processes that underlie language comprehension and production (Godfroid, 2019; Conklin *et al.*, 2018; Rayner & Clifton, 2009).

Eye-tracking technology serves as a valuable bridge between theoretical constructs in applied linguistics and empirical cognitive evidence. Recent research emphasizes its effectiveness in capturing subconscious cognitive processes, including attentional allocation, lexical access, and syntactic parsing, thus offering concrete validation to theoretical frameworks in second language acquisition (SLA) and language pedagogy (Godfroid *et al.*, 2020; Michel *et al.*, 2021). For instance, eye-tracking provides empirical substantiation to Schmidt's (1990) Noticing Hypothesis, elucidating how learners consciously or subconsciously notice linguistic forms during exposure (Godfroid & Uggen, 2013). By measuring fixation durations and gaze sequences, researchers gain critical insights into learner engagement and cognitive load, bridging qualitative interpretations of learner behavior with quantitative cognitive metrics (Winke *et al.*, 2013).

Crucially, the relationship between eye-tracking and applied linguistics is not merely methodological but epistemological, reshaping the kinds of questions researchers can ask. Applied linguistics research has increasingly leveraged eye-tracking to explore nuanced aspects of language processing such as metaphor comprehension (Holmqvist & Andersson, 2017), pragmatic inference- making (Fukuta & Yamashita, 2021), and multimodal literacy development (Pellicer-Sánchez, 2020). Eye-tracking's capacity to precisely locate where and how long learners visually engage with textual or visual stimuli provides invaluable qualitative insights. This aligns closely with the editorial orientation of the *Colombian Applied Linguistics Journal* (*CALJ*), which traditionally prioritizes non-experimental, qualitative approaches that foreground learners' subjective experiences and interpretative processes.

Despite its quantitative roots, eye-tracking data can indeed be effectively integrated into qualitative research designs. Gaze plot analyses, for example, have been qualitatively interpreted to illustrate learners' strategic reading behaviors, decision-making processes, and metacognitive strategies (Conklin *et al.*, 2018). Complementing qualitative methods such as stimulated recall or think- aloud protocols, eye-tracking data enrich the interpretative depth, ensuring triangulation of data sources and enhancing credibility in qualitative studies (Fukuta & Yamashita, 2021; Pellicer-Sánchez, 2020). Thus, rather than contradicting *CALJ*'s qualitative methodological preferences, eye-tracking serves to extend and deepen qualitative insights, providing explicit, real-time evidence of cognitive phenomena previously accessible only through retrospective self-report.

Recent studies employing eye-tracking have successfully illustrated qualitative processes in vocabulary acquisition and reading comprehension. Godfroid *et al.* (2020) qualitatively analyzed gaze sequences to explore second language vocabulary retention, highlighting nuanced individual differences in learning trajectories and strategies. Similarly, Michel *et al.* (2021) interpreted fixation patterns to elucidate the cognitive demands of syntactic complexity in second language reading, offering insights that purely qualitative methodologies might overlook.

Another significant implication is methodological transparency and reproducibility. Eye-tracking enhances methodological rigor in qualitative applied linguistics research by providing explicit visual representations of learner cognition, significantly reducing the interpretative ambiguity typically associated with qualitative data analysis. This technology allows applied linguists to substantiate qualitative claims with robust visual evidence, strengthening interpretative arguments (Conklin *et al.*, 2018).

Furthermore, eye-tracking's inherent multimodality aligns well with contemporary qualitative approaches within applied linguistics, which emphasize learners' interactions across different modalities (Pellicer-Sánchez, 2020). By simultaneously tracking visual engagement with textual, graphical, and multimodal input, researchers can offer richly contextualized interpretations of learner behaviors, thereby resonating strongly with the qualitative ethos favored by *CALJ*.

From a pedagogical standpoint, eye-tracking research generates qualitative insights with direct implications for instructional practices. For example, qualitative analyses of gaze data have informed classroom practices by elucidating effective reading strategies, guiding the design of learner-friendly texts, and identifying specific instructional interventions for learners experiencing cognitive overload or attentional difficulties (Winke *et al.*, 2013; Godfroid *et al.*, 2020).

Given the substantial qualitative interpretative potential of eye-tracking data, applied linguists adopting eye-tracking methodologies are advised to articulate explicitly how these data align with and enrich qualitative analytic frameworks. Such explicit articulation would facilitate acceptance of eye- tracking studies in *CALJ*, given the journal's preference for qualitative and interpretative methodologies.

In conclusion, eye-tracking technology makes a significant contribution to applied linguistics research by providing nuanced, real-time insights into cognitive processes underlying language acquisition and processing. Although traditionally quantitative, eye-tracking data are readily amenable to qualitative interpretation and integration into qualitative research frameworks. By bridging empirical cognitive insights and qualitative interpretative analyses, eye-tracking enables applied linguistics researchers to engage deeply with learner cognition and interaction, advancing the field's theoretical, methodological, and pedagogical frontiers. Consequently, eye-tracking research, when explicitly positioned within qualitative methodological paradigms, aligns seamlessly with the scholarly ethos and methodological preferences of *CALJ*.

References

Al-Edaily, A., Al-Wabil, A., & Al-Ohali, Y. (2013). Interactive screening for learning difficulties: Analyzing visual patterns of reading Arabic scripts with eye tracking. In *HCI International 2013—Posters' Extended Abstracts* (Vol. 374, pp. 3–7). Springer. https://doi.org/10.1007/978-3-642-39476-8_1

Altmann, G. T. M. (2011). Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta Psychologica, 137*(2), 190–200. https://doi.org/10.1016/j.actpsy.2010.09.009

Altmann, G. T. M. (2011). The mediation of eye movements by spoken language. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 979–1004). Oxford University Press.

Aryadoust, V. (2019). Eye tracking in language assessment: A review of eye-tracking studies in language testing and learning. *International Journal of Listening, 33*(1), 1–25. https://doi.org/10.1080/10904018.2018.1455641

Aryadoust, V., Foo, S., & Ng, L. Y. (2020). Integrating eye-tracking and neuroimaging in language assessment: Insights into test methods and cognitive load. *Language Testing, 37*(3), 417–442. https://doi.org/10.1177/0265532219898388

Aryadoust, V., Foo, S., & Ng, L. Y. (2022). What can gaze behaviors, neuroimaging data, and test scores tell us about test method effects and cognitive load in listening assessments? *Language Testing, 39*(1), 56–89. https://doi.org/10.1177/02655322211026876

Baazeem, I., Al-Khalifa, H., & Al-Salman, A. (2021). Cognitively driven Arabic text readability assessment using eye-tracking. *Applied Sciences, 11*(18), 8607. https://doi.org/10.3390/app11188607

Batty, A. O. (2021). An eye-tracking study of attention to visual cues in L2 listening tests. *Language Testing, 38*(4), 511–535. https://doi.org/10.1177/0265532220982896

Bax, S., & Chan, S. (2019). Using eye-tracking research to investigate language test validity and design. *System, 83*, 64–78. https://doi.org/10.1016/j.system.2019.07.005

Brunfaut, T. (2016). Assessing listening. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 97–112). Mouton de Gruyter. https://doi.org/10.1515/9781614516873-007

Chang, A. C.-S. (2009). Gains to L2 listeners from reading while listening versus listening only in extensive listening. *System, 37*(4), 554–565. https://doi.org/10.1016/j.system.2009.09.009

Chang, A. C.-S., & Millet, S. (2015). Developing L2 listening fluency through extended listening-focused activities. TESOL Quarterly, 49(4), 1075–1097. https://doi.org/10.1002/tesq.198

Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge University Press. https://doi.org/10.1017/9781108233279

Conklin, K., Pellicer-Sánchez, A., & Vilkaitė-Lozdienė, L. (2020). *Eye-tracking in second language reading: A research synthesis and methodological guide.* Routledge. https://doi.org/10.4324/9780429203112

Conklin, K., & Alotaibi, N. (2023). Methodological issues in eye-tracking studies of reading while listening. *Studies in Second Language Acquisition, 45*(3), 783–805. https://doi.org/10.1017/S0272263122000305

Day, R., & Robb, T. (2015). Extensive reading. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed., pp. 294–306). National Geographic Learning/Cengage Learning.

Dolgunsöz, G. (2016). The effect of L2 reading proficiency on the use of reading strategies and metacognitive awareness. *Journal of Language and Linguistic Studies, 12*(2), 42–55.

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience, 25*, 69–91. https://doi.org/10.1016/j.dcn.2016.11.001

Field, J. (2015). *Listening in the language classroom*. Cambridge University Press.

Fukuta, J., & Yamashita, J. (2021). Eye-tracking the development of second language pragmatic inference. *Journal of Pragmatics, 171*, 39–53. https://doi.org/10.1016/j.pragma.2020.10.019

Godfroid, A. (2019). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. Routledge. https://doi.org/10.4324/9781315177022

Godfroid, A., Lin, C. H., & Ryu, C. (2020). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Studies in Second Language Acquisition, 42*(4), 841–875. https://doi.org/10.1017/S0272263119000700

Godfroid, A., & Uggen, T. (2013). Attention to irregular verbs by beginning learners of German: An eye-movement study. *Studies in Second Language Acquisition, 35*(2), 291–322. https://doi.org/10.1017/S0272263112000906

Godfroid, A., Winke, P., & Conklin, K. (2020). Exploring the depths of second language processing with eye tracking: An introduction. *Second Language Research, 36*(3), 243–255. https://doi.org/10.1177/0267658320930157

Hohensinn, C., & Baghaei, P. (2017). The effect of the position of the correct option in multiple-choice tests on item difficulty and test-takers' eye movements. *Educational Assessment, 22*(4), 291–304. https://doi.org/10.1080/10627197.2017.1381554

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697083.001.0001

Holmqvist, K., & Andersson, R. (2017). Eye-tracking in metaphor research: Methodological considerations and empirical evidence. *Metaphor and Symbol, 32*(1), 1–16. https://doi.org/10.1080/10926488.2017.1297625

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., Lee, M. H., Chiou, G. L., Liang, J. C., & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review, 10*, 90–115. https://doi.org/10.1016/j.edurev.2013.10.001

Langenfeld, T., Sauro, S., & Jarodzka, H. (2016). Reading strategies in L2 multiple-choice testing: An eye-tracking study. *System, 59*, 1–15. https://doi.org/10.1016/j.system.2016.04.004

Lightbown, P. M. (1992). Getting quality input in the second/foreign language classroom. *TESL Canada Journal, 9*(2), 9–26. https://doi.org/10.18806/tesl.v9i2.602

Lindner, M. A., Eitel, A., Thoma, G.-B., Dalehefte, I. M., Ihme, J. M., & Köller, O. (2014). Tracking the decision-making process in multiple-choice assessment: Evidence from eye movements. *Journal of Educational Measurement, 51*(3), 241–255. https://doi.org/10.1111/jedm.12056

Li, W., Chen, Y., & Zhang, J. (2020). Exploring remote eye-tracking for online education: The relationship between students' gaze patterns in instructional videos and learning outcomes. *Computers & Education, 156*, 103960. https://doi.org/10.1016/j.compedu.2020.103960

Michel, M., Révész, A., Lu, X., & Lee, M. (2021). Investigating second language writing processes across tasks using keystroke logging and eye-tracking. *Second Language Research, 37*(4), 623–648. https://doi.org/10.1177/0267658320927760

Ockey, G. J. (2007). Construct implications of including still images or video in computer-based listening tests. *Language Testing, 24*(4), 517–537. https://doi.org/10.1177/0265532207080986

Peitek, N., Siegmund, J., Apel, S., Hofmeister, J., Kästner, C., Parnin, C., Bethmann, A., Leich, T., Saake, G., & Brechmann, A. (2020). A look into programmers' heads: A study of program comprehension processes. *IEEE Transactions on Software Engineering, 46*(4), 442–462. https://doi.org/10.1109/TSE.2018.2876736

Pellicer-Sánchez, A., Tragant, E., Conesa, I., & Serrano, R. (2018). Reading while listening to stories in a foreign language: The effect of input modality on vocabulary learning. Studies in Second Language Acquisition, 40(3), 554–578. https://doi.org/10.1017/S0272263117000344

Pellicer-Sánchez, A. (2020). Eye-tracking and reading in applied linguistics: Multimodal literacy development in L2 contexts. *Language Teaching, 53*(4), 517–531. https://doi.org/10.1017/S0261444820000238

Ranalli, J., Feng, H.H., & Chukharev-Hudilainen, E. (2019). Exploring the potential of process-tracing technologies to support assessment for learning of L2 writing. *Assessing Writing, 36*, 77–89. https://doi.org/10.1016/j.asw.2018.05.004

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506. https://doi.org/10.1080/17470210902866788

Rayner, K., & Clifton, C., Jr. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology, 80*(1), 4–9. https://doi.org/10.1016/j.biopsycho.2008.05.002

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writing processes from a psycholinguistic perspective: The role of task complexity. *Studies in Second Language Acquisition, 41*(4), 867–891. https://doi.org/10.1017/S0272263119000022

Salverda, A. P., & Altmann, G. T. M. (2011). Attentional and cognitive control in visual world eye-tracking. *Journal of Memory and Language, 65*(3), 209–231. https://doi.org/10.1016/j.jml.2011.03.004

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Suvorov, R. (2015). The use of eye tracking in second language listening research: A case of studying the visual input in L2 listening tests. *Language Testing, 32*(4), 463–483. https://doi.org/10.1177/0265532214562099

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation* (Vol. 55, pp. 37–76). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Wagner, E. (2010). The effects of the use of video texts on ESL listening test performance. *Language Testing, 27*(4), 493–513. https://doi.org/10.1177/0265532210367486

Wang, Y., Huang, R., & Feng, Y. (2018). Eye-tracking study of university students' e-book reading behavior: The role of prior knowledge. *Journal of Educational Technology & Society, 21*(2), 228–240

62

Caoxi & Zhenni, M. (2025) • Colomb. Appl. Linguist. J.
Printed ISSN 0123-4641 Online ISSN 2248-7085 • July - December 2025. Vol. 27 • Issue 2 pp. 51-63.

Webb, S., & Chang, A. C.-S. (2014). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research, 18*(6), 713–733. https://doi.org/10.1177/1362168814559800

Winke, P., Godfroid, A., & Gass, S. (2013). Introduction to the special issue: Eye-movement research in second language acquisition and bilingualism. *Studies in Second Language Acquisition, 35*(2), 205–212. https://doi.org/10.1017/S0272263112000888

Yaneva, V., Ha, L. A., & Keane, M. T. (2021). Do successful test-takers look at the same things as unsuccessful ones? An eye-tracking study of TOEFL iBT® reading tests. *Language Testing, 38*(3), 378–401. https://doi.org/10.1177/0265532220960384

Zhu, L., Li, M., & Wang, X. (2021). Investigating discourse synthesis in integrated L2 writing assessment: Evidence from eye-tracking, keystroke logging, and stimulated recall. *Assessing Writing, 48*, 100521. https://doi.org/10.1016/j.asw.2021.100521