Eye tracking research on readers' interactions with multimodal texts: a mini-review

Al Ryanne Gabonada Gatcho[1]*, Jeremiah Paul Giron Manuel[2] and Rocky James Guevarra Sarasua[3]

[1]School of Foreign Languages and Literature, Hunan Institute of Science and Technology, Yueyang, China, [2]Sekolah Pelita Harapan, Bogor, Indonesia, [3]Faculty of Teacher Development, Philippine Normal University, Manila, Philippines

This mini-review advocates for the role of eye-tracking research in understanding readers' engagement with multimodal texts. Synthesizing findings from a variety of studies, the review reveals how eye-tracking gives insights into sophisticated interactions between the textual, visual, and auditory elements within reading environments that assist both cognitive processing and comprehension. Several gaps were revealed: limited demographic scope, integration of advanced technologies, and substantial impact to the area of eye tracking and multimodal literacy. Future directions must therefore include studies across diverse populations, innovative technologies, and cross-discipline research studies. These directions are critical for advancing literacy development in an increasingly multimodal digital world.

KEYWORDS

cognitive processes, eye tracking, literacy, multimodal reading, reading comprehension

Introduction

Numerous research domains have used eye tracking. In marketing, its presence is apparent in consumer experience research, where viewers' eyes were examined on various purchase stages (Duerrschmid and Danner, 2018; Ishibashi et al., 2019). In health care, it is acknowledged for its diagnostic (Sun et al., 2022), therapeutic (Harezlak and Kasprowski, 2018), and interactive properties (Tscholl et al., 2020). In education, examining learners' experience with learning materials is facilitated by such technology (Conley et al., 2020; Serrano-Mamolar et al., 2023; Susac et al., 2023). Throughout the years, eye tracking is regarded as a tool for examining reading engagement (Child et al., 2020; Liu and Yu, 2022), comprehension (Abundis-Gutiérrez et al., 2018; Mézière et al., 2023), and atypical patterns like dyslexia or ADHD (Klein et al., 2019).

The ever-evolving technological landscape popularized the use of multimodal texts enabling learners to absorb information regardless of their learning styles (Bearne, 2012; Jancsary et al., 2016; Smith, 2012). Readers need to interact with multiple modes of information. The development of frameworks relative to multimodal literacies has taken up space in this field (Serafini, 2015; Shin et al., 2020). This area demands scoping examinations, constructing appropriate methodologies and optimizing its function. Eye-tracking provides perspectives on how readers can interact efficiently with multimodal materials (Holmqvist et al., 2011a, 2011b). Contrarily, the researchers find the tension posed by the utility of multimodal materials in literacy, believing that such practices cause strain to the limit of cognitive processing of the learners, thus limiting its potency (Mayer and Moreno, 2010). The insights gained from eye-tracking studies may guide the development of effective multimodal learning resources (Alemdag and Cagiltay, 2018), enabling researchers to analyze how long

readers focus on particular elements of multimodal text (van der Sluis et al., 2018; Armfield, 2011; Schmidt-Weigand et al., 2010), which often demands transitions between different types of representations and careful consideration of information presented in graphs, as is the case with the Programme for International Student Assessment (PISA) items (Susac et al., 2018; Mason et al., 2015). Mason et al. (2022), showed how visual attention patterns can be used to inform the development of instructional video content. To illustrate the potential of eye-tracking in refining multimodal materials, consider its application in multimedia learning. Mayer and Fiorella (2021) discuss how students' eye movements in multimedia lessons can reveal whether they are efficiently coordinating attention between explanatory diagrams and narration, highlighting areas where learners struggle to connect visual and verbal information, prompting educators to revise the layout or sequence of these materials for enhanced learning (Wiegand et al., 2017). Thus, this mini review validates the body of research on eye tracking in multimodal reading to identify the arguments that drive this expanding field. It reveals gaps in research while providing contrasting perspectives that influence current understandings. The review does not only contextualize current knowledge but also prognosticates on future implications.

Eye tracking in reading research

The application of eye tracking technology in reading is anchored on Just and Carpenter's (1980) "eye-mind hypothesis," positing a connection between gaze location and cognitive processing. Whoever started the eye tracking movement during reading remains unsettled. However, the experiments of Louis Emile Javal in 1879 concluded that reading is a nonlinear process since readers' eyes exhibit a series of quick movements dubbed by brief moments of stillness on certain parts of the text (Płużyczka, 2018). Moving to the early 20th century, Edmund Huey invented the first yet intrusive eye tracker to understand reading behaviors (Walczyk et al., 2014). Following this, Buswell's seminal work in the 1920s discovered that eye movements are not smooth but composed of saccades and fixations (Wade, 2020), concepts that were previously identified by Javal but were not named during his time. Since then several researchers demonstrated how saccadic movements correlated with cognitive processing during reading (Rayner, 1978; Taylor, 1965).

The late 20th and early 21st centuries yielded profound insights into how readers interact with the texts through eye tracking technology within conventional reading environments. For example, skilled readers have shorter fixations and longer saccades as opposed to struggling readers (Boland, 2004; Weger and Inhoff, 2006). Shifting the focus to how engaged readers interact with texts, they display longer fixations on meaningful parts, while disengaged readers demonstrate quick eye movement patterns, shorter fixations, and frequent regressions—which signify comprehension difficulty (Holmqvist et al., 2011a, 2011b; Rayner and Pollatsek, 2006). Additionally, dyslexics display atypical eye movement patterns, finally shedding light on their word recognition and processing speed difficulties (Hyönä and Olson, 1995; Jones et al., 2008). Learners with ADHD, on the other hand, have erratic fixations with frequent saccadic movements (Karatekin and Asarnow, 1998), thereby clarifying why reading is a predicament for them. On a practical note,

eye tracking technology paved to the development of targeted interventions like improving text readability (Goldberg and Wichansky, 2003) or integrating assistive technologies in the teaching practice (van Gog and Scheiter, 2010).

Multimodal texts and reader interaction

A multimodal text is a combination of more than one of the "modes," pertaining to the method of communication being used: spatial, linguistic, visual, gestural, and audio, creating meaning far beyond the capacity of any single mode to do so (Moses and Reid, 2021; Sutrisno et al., 2023; Jewitt, 2013; Forceville, 2011). For example, with digital presentation, it can be read and seen coupled with animations and voice narration (Kress, 2010). Second is interactivity where many of the multimodal texts, especially those digital forms let users engage with the content through clicks, links, or swipes that determine how and what they navigate through in the text (Serafini, 2014). Third characteristic of the multimodal texts is non-linearity. Hypertexts and websites give readers liberty to navigate content through different pathways.

Kress and van Leeuwen (2006) explain that multimodal literacy relies on a reading operation where the reader decodes and syncretizes these diverse semiotic sources (Serafini, 2011; Forceville, 2010) unlike traditional print sources. For example, websites and new forms of digital media, such as digital comics strips or an infographics allow users to click on links to take them where they choose to within the content, thereby actively assuming meaning-making agency (Jewitt, 2013; Bezemer and Kress, 2008). In the case of an instructional video, they interpret the visual demonstrations as well as the auditory instructions (Serafini, 2014). For these reasons, it is necessary to delve into the literature about multimodal text reader interaction. To begin with, it widens to print literacy and encompasses various ways different people communicate within media environments (Cope and Kalantzis, 2009). With the growing incorporation of technology in educational systems, understanding how readers interact with multimodal texts can inform teaching practices so that students will be better prepared in the contemporary world (Walsh, 2010; Jewitt, 2009).

As noted by Shin (2023), "we live in a multimodal world," and this becomes evident with the integration of various modes of communication to convey complex information. Eye tracking, according to Holsanova (2014), is a potentially useful tool for gathering accurate visual data about how readers engage with multimodal texts.

Visual and textual synchronization

Eye tracking studies revealed that the temporal aspect of visuals within a text provides a context that primes readers' mind for new or important information (Gegenfurtner et al., 2011; Hoffman, 2016), which facilitates better comprehension (Huth et al., 2024; Lee and Révész, 2018; Loewen and Inceoglu, 2016). These visual cues also act as cognitive anchors, aiding information retention (Pjesivac et al., 2021). Recognizable images are recalled with fewer fixations at the center during recognition phases

(Borkin et al., 2015), and notably, animations are found to enhance readers' information recall as opposed to still visuals due to their sequential properties (Coskun and Cagiltay, 2022). Additionally, the frequent shift of eye focus between animated segments and texts insinuate a potentially fragmented reading experience (Foulsham et al., 2016).

Eye movement data reveals that visuals affect where and how long attention is held (Indrarathne and Kormos, 2017; Lee and Jung, 2021), determines whether readers are visual learners by comparing their length of gaze on images and texts (Koć-Januchta et al., 2017), and provides insights into strategies employed when interpreting visuals (Borkin et al., 2015). However, Huang et al. (2011) revealed otherwise since graph drawings on texts have minimal impact on readers' task performance.

In terms of affective response, dilated pupils on high resolution images imply heightened reading interest (Brunyé et al., 2019), while multiple and rapid side movements of the eyes when absurd images are encountered suggest discomfort (Gregory, 2015). In scenarios where readers are locating specific information, eye tracking has revealed that visuals act as reference points, which accelerates the process (Drew et al., 2017; van der Gijp et al., 2017). Additionally, differences in strategies of various demographics in using visuals as search cues were also explored (Józsa and Hámornik, 2012).

Recent eye tracking researches have explored how games integrated in digital texts can be more accessible for visually impaired people through gazed-controlled interfaces, bypassing the need for a traditional input device like keyboard or mouse (Deng et al., 2014; Munoz et al., 2011). Likewise, Krebs et al. (2021) and Gu et al. (2022) demonstrated that adaptive learning games designed with eye tracking feedback can improve reading comprehension of dyslexic students.

Texts with audio elements

Text processing may be impacted by audio integration. For example, voiced narratives caused readers to focus on text or image portions for longer (Kruger, 2012; Liu et al., 2011); explanatory audios helped readers focus on the relevant content and reduced the need to read the texts again (Conklin et al., 2020); and quiet music during passages can help people reflect more deeply (Kerchner, 2014; Holmqvist et al., 2011a, 2011b). When background music was played during brief passages, there was a decrease in visual wandering; however, lengthier passages showed the opposite pattern (Hyönä and Ekholm, 2016). Cognitive load may also be affected by variations in speech volume, tone, and tempo (Hvelplund, 2011). This can be seen when watching movies with subtitles because viewers' eye movements change based on whether they are simultaneously exposed to dynamic audio that supports or contradicts the textual information being displayed (Kruger and Steyn, 2014).

Moreover, auditory learners benefit from audio-enhanced texts as data showed that their eyes are less strained when processing information through listening rather than reading (Conklin et al., 2020; Pellicer-Sánchez et al., 2018). However, Kruger and Steyn (2014) noted that the design of audio elements within the texts should consider diverse readers, especially those with hearing impairments or those who are easily distracted by sounds. In light of this consideration, integrating both visual and auditory stimuli is recommended.

Combinations of textual, visual and auditory modes

The integration of texts, visuals, and audios altogether in multimodal materials has demonstrated strengthened contribution to reader engagement in the field of eye tracking. Schiavo et al. (2015) developed the GARY application, a text-to-speech multimedia application, supporting struggling readers in their progress. Similarly, the Zurich Cognitive Language Processing Corpus (ZuCo) demonstrated findings on the advancement of studies concerning literacy and language development at the brain and eye coordination fields (Hollenstein et al., 2018). While such innovations are considered a gamble, looking at both positive and negative effects observed in its utility (Bus et al., 2015; Dobler, 2015), its potential to help improve reading comprehension levels can no longer be ignored.

Recent innovations in the field

Current technological developments came as beneficial complementary tools for enhancing multimodal literacy development. Santos et al. (2016) demonstrated Augmented Reality (AR) as an effective tool in improving vocabulary of the learners, utilizing multimedia information in its setting. Placing buttons for translating, describing, and listening, the constructed environment allowed the students to immerse themselves in learning new words. In the case of reading comprehension, Danaei et al. (2020) revealed that children who used AR-based literature had better grasp of stories over those who had traditional books. Even among children with learning disabilities, AR-induced learning boosted reading comprehension (Shaaban and Mohamed, 2024). Moreover, virtual reality (VR) took a similar position in multimodal text processing, supporting learners to receive help and encouragement (Tai et al., 2020; Asad et al., 2021). These findings supported both platforms as a potent developer of multimedia text literacy (Liu et al., 2020; Bursali and Yilmaz, 2019). These recent progress in the field of AR and VR boosted developers to enhance eye tracking technology linked to these innovations (Dudinskaya et al., 2020). Head-mounted displays specific for AR and VR intersects for eye tracking were developed, emphasizing gaze-based interaction (Kapp et al., 2021).

On the other hand, AR considerations on optimizing its integration hold developers as risks to children were also observed (Li et al., 2018; Papanastasiou et al., 2018). Similarly, VR-based implementations are costly (Kamińska et al., 2019), successful integration requires demanding labor from the teacher (Alizadeh, 2019), and potential mental health risks (Richter et al., 2018). Moreover, it also keeps the field of eye tracking demanding for fresh insights relative to these practical gaps revealed in literature.

Discussion: unpacking the gaps

The area of eye tracking in the field of literacy remains significant, enticing developers and explorers to continually locate substantial materials towards its optimized implementation. As demonstrated by the GARY (Schiavo et al., 2015) and ZuCo (Hollenstein et al., 2018) applications, one can actually readily witness a variety of eye tracking-specific literacy-promoting methods. As Holsanova (2014) firmly

thought, the technology's current contribution to literacy cannot be dismissed. The visions of Just and Carpenter in 1980 relative to the eye-mind hypothesis paved the way to these innovations, integrating the nuances in different formats throughout its development. On the other hand, its sparsity in its field stands observable despite its prevalence in studies.

Extensive research on eye tracking in traditional settings has produced insightful findings, particularly regarding its impact on multimodal texts (Gegenfurtner et al., 2011; Hoffman, 2016; Huth et al., 2024; Lee and Révész, 2018; Loewen and Inceoglu, 2016). Additionally, the extent of the demographics of the participants includes dyslexics (Hyönä and Olson, 1995; Jones et al., 2008) and children with ADHD (Karatekin and Asarnow, 1998). Focusing on eye movements (Indrarathne and Kormos, 2017; Lee and Jung, 2021), including the dilation of the pupils (Brunyé et al., 2019), and length of gaze on images and texts (Koć-Januchta et al., 2017), this area of study has demonstrated interesting insights especially to helping readers. While present advancements in eye tracking have integrated technological innovations (Krebs et al., 2021; Gu et al., 2022), the speedy development of technology enhancements demands continual findings.

Another gap observed in this field was the lack of studies on long term effects of eye tracking such how repeated exposure to multimodal texts influences memory consolidation and cognitive development, seeking recommendations on pragmatic actions. Pjesivac et al. (2021) and Borkin et al. (2015) concurringly shared vital insights on visual cues but nevertheless maintained a narrow approach on recognition and retention stages. While Coskun and Cagiltay (2022) posited similar findings on the utility of visual cues towards enhanced information recall, the postulations remained unlinked with long term considerations. Even in the area of including audio with text elements, studies regarding this concern seemed lacking. Conklin et al. (2020), Kruger (2012), and Liu et al. (2011) demonstrated the benefits of audio for reading focus, but their studies only covered voice-assisted reading. The field's exclusive practices, which make it difficult for the discipline to provide long-term considerations in eye tracking research, are implied in this gap.

Studies on eye tracking in a cross-disciplinary lens demonstrated another demand to assist this field gain prominence. Technology was heavily considered in the studies visited (Krebs et al., 2021; Gu et al., 2022; Schiavo et al., 2015; Hollenstein et al., 2018) especially augmented and virtual reality integrations, but current literacy partnerships with psychology seemed ignored. Two of the closest cross-disciplinary findings were the studies concerning participants with ADHD (Karatekin and Asarnow, 1998), and dyslexia (Hyönä and Olson, 1995; Jones et al., 2008), which currently are demanding fresh explorations considering their dates of publication.

Recommendations and future directions

The following suggestions are seen to be advantageous to the field of eye tracking research about its application in reading multimodal texts.

There is a need in integrating more innovative applications to eye tracking activities and in reading multimodal texts. Considering the fast pace of the technological landscape in education, rearing today's children with the most advanced gadgets, latest developments towards multimodal literacy, especially the gamified ones are critical.

The narrow participant demographics in these studies call for expanding research to include mature readers and diverse needs. Similarly, the spectrum in ADHD has developed in recent years, demanding specific interventions and approaches for a particular spectrum type. Moreover, studies concerning children with socio-emotional learning needs should also be given attention in the field of eye tracking research, and how literacy can be addressed among these children especially in reading multimodal texts.

Challenges in AR and VR technologies linked with multimodal text literacy development demand explorations for eye tracking research. Institutions can explore feasible means to alleviate the costs and help communities to enjoy such innovations. Moreover, workshops and trainings may be arranged for utilizing such advancements toward eye tracking research. One may also observe and examine the suggestions of Mayer and Moreno (2010) in decreasing the cognitive processing load that multimedia materials place among learners.

These recommendations may not be exhaustive considering that this is a mini-review. Longer term consolidation of study results through systematic reviews and meta-analyses can provide greater insights into unexplored impact of eye-tracking on literacy development and multimodal reading in the digital era.

Conclusion

This paper synthesizes eye tracking studies to explore how readers interact with multimodal texts. Findings reveal that multimodal reading requires cognitive flexibility since readers navigate interplays of content. Key strategies like prioritizing information based on its perceived importance and confirming text with visual aids are essential for comprehension. However, research gaps persist, particularly in understanding how varied populations, such as those with reading difficulties or non-native language backgrounds, engage with multimodal texts in naturalistic settings. Future research should focus on these areas to enhance instructional designs and embrace the digital evolution of literacy practices.

Author contributions

AG: Conceptualization, Supervision, Writing – original draft. JM: Writing – review & editing. RS: Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abundis-Gutiérrez, A., González-Becerra, V. H., Del Rio, J. M., López, M. A., Ramírez, A. A. V., Sánchez, D. O., et al. (2018). Reading comprehension and eye-tracking in college students: comparison between low-and middle-skilled readers. *Psychology* 9, 2972–2983. doi: 10.4236/psych.2018.915172

Alemdag, E., and Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Comput. Educ.* 125, 413–428. doi: 10.1016/j.compedu.2018.06.023

Alizadeh, M. (2019). Virtual reality in the language classroom: theory and practice. *Comput. Assist. Lang. Learn. Electron. J.* 20, 21–30. Available at: https://callej.org/index.php/journal/article/view/280

Armfield, D. M. (2011). Multimodality: a social semiotic approach to contemporary communication. By Gunther Kress. *Tech. Commun. Q.* 20, 347–349. doi: 10.1080/10572252.2011.551502

Asad, M. M., Naz, A., Churi, P., and Tahanzadeh, M. M. (2021). Virtual reality as pedagogical tool to enhance experiential learning: a systematic literature review. *Educ. Res. Int.* 2021:7061623. doi: 10.1155/2021/7061623

Bearne, E. (2012). "Multimodal texts: what they are and how children use them" in Literacy moves on (Abingdon: David Fulton Publishers), 16–30.

Bezemer, J., and Kress, G. (2008). Writing in multimodal texts. *Writ. Commun.* 25, 166–195. doi: 10.1177/0741088307313177

Boland, J. E. (2004). "Linking eye movements to sentence comprehension in reading and listening" in The on-line study of sentence comprehension (London: Psychology Press), 51–76.

Borkin, M. A., Bylinskii, Z., Kim, N. W., Bainbridge, C. M., Yeh, C. S., Borkin, D., et al. (2015). Beyond memorability: visualization recognition and recall. *IEEE Trans. Vis. Comput. Graph.* 22, 519–528. doi: 10.1109/TVCG.2015.2467732

Brunyé, T. T., Drew, T., Weaver, D. L., and Elmore, J. G. (2019). A review of eye tracking for understanding and improving diagnostic interpretation. *Cogn. Res. Princ. Implic.* 4:7. doi: 10.1186/s41235-019-0159-2

Bursali, H., and Yilmaz, R. M. (2019). Effect of augmented reality applications on secondary school students' reading comprehension and learning permanency. *Comput. Hum. Behav.* 95, 126–135. doi: 10.1016/j.chb.2019.01.035

Bus, A. G., Takacs, Z. K., and Kegel, C. A. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. *Dev. Rev.* 35, 79–97. doi: 10.1016/j.dr.2014.12.004

Child, S., Oakhill, J., and Garnham, A. (2020). Tracking your emotions: an eye-tracking study on reader's engagement with perspective during text comprehension. *Q. J. Exp. Psychol.* 73, 929–940. doi: 10.1177/1747021820905561

Conklin, K., Alotaibi, S., Pellicer-Sánchez, A., and Vilkaitė-Lozdienė, L. (2020). What eye-tracking tells us about reading-only and reading-while-listening in a first and second language. *Second. Lang. Res.* 36, 257–276. doi: 10.1177/0267658320921496

Conley, Q., Earnshaw, Y., and McWatters, G. (2020). Examining course layouts in blackboard: using eye-tracking to evaluate usability in a learning management system. *Int. J. Hum.-Comput. Interact.* 36, 373–385. doi: 10.1080/10447318.2019.1644841

Cope, B., and Kalantzis, M. (2009). "Multiliteracies": new literacies, new learning. *Pedagogies* 4, 164–195. doi: 10.1080/15544800903076044

Coskun, A., and Cagiltay, K. (2022). A systematic review of eye-tracking-based research on animated multimedia learning. *J. Comput. Assist. Learn.* 38, 581–598. doi: 10.1111/jcal.12629

Danaei, D., Jamali, H. R., Mansourian, Y., and Rastegarpour, H. (2020). Comparing reading comprehension between children reading augmented reality and print storybooks. *Comput. Educ.* 153:103900. doi: 10.1016/j.compedu.2020.103900

Deng, S., Kirkby, J. A., Chang, J., and Zhang, J. J. (2014). Multimodality with eye tracking and haptics: a new horizon for serious games? *Int. J. Serious Games* 1, 17–34. doi: 10.17083/ijsg.v1i4.24

Dobler, E. (2015). E-textbooks: a personalized learning experience or a digital distraction? *J. Adolesc. Adult. Lit.* 58, 482–491. doi: 10.1002/jaal.391

Drew, T., Boettcher, S. E., and Wolfe, J. M. (2017). One visual search, many memory searches: an eye-tracking investigation of hybrid search. *J. Vis.* 17:5. doi: 10.1167/17.11.5

Dudinskaya, E. C., Naspetti, S., and Zanoli, R. (2020). Using eye-tracking as an aid to design on-screen choice experiments. *J. Choice Model.* 36, 100232. doi: 10.1016/j.jocm.2020.100232

Duerrschmid, K., and Danner, L. (2018). "Eye tracking in consumer research" in Methods in consumer research (Duxford: Woodhead Publishing), 279–318.

Forceville, C. J. (2010). The Routledge handbook of multimodal analysis. *J. Pragmat.* 42, 2604–2608. doi: 10.1016/j.pragma.2010.03.003

Forceville, C. J. (2011). Multimodality: a social semiotic approach to contemporary communication. *J. Pragmat.* 43, 3624–3626. doi: 10.1016/j.pragma.2011.06.013

Foulsham, T., Wybrow, D., and Cohn, N. (2016). Reading without words: eye movements in the comprehension of comic strips. *Appl. Cogn. Psychol.* 30, 566–579. doi: 10.1002/acp.3229

Gegenfurtner, A., Lehtinen, E., and Säljö, R. (2011). Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educ. Psychol. Rev.* 23, 523–552. doi: 10.1007/s10648-011-9174-7

Goldberg, J. H., and Wichansky, A. M. (2003). "Eye tracking in usability evaluation: a practitioner's guide" in In the mind's eye (Amsterdam: North-Holland), 493–516.

Gregory, R. L. (2015). Eye and brain: the psychology of seeing. Princeton, NJ: Princeton University Press.

Gu, C., Chen, J., Lin, J., Lin, S., Wu, W., Jiang, Q., et al. (2022). The impact of eye-tracking games as a training case on students' learning interest and continuous learning intention in game design courses: taking flappy bird as an example. *Learn. Motiv.* 78:101808. doi: 10.1016/j.lmot.2022.101808

Harezlak, K., and Kasprowski, P. (2018). Application of eye tracking in medicine: a survey, research issues and challenges. *Comput. Med. Imaging Graph.* 65, 176–190. doi: 10.1016/j.compmedimag.2017.04.006

Hoffman, J. E. (2016). Visual attention and eye movements. *Attention*, 119–153. Available at: https://psycnet.apa.org/doi/10.1037/xhp0000291

Hollenstein, N., Rotsztejn, J., Troendle, M., Pedroni, A., Zhang, C., and Langer, N. (2018). ZuCo, a simultaneous EEG and eye-tracking resource for natural sentence reading. *Sci. Data* 5, 1–13. doi: 10.1038/sdata.2018.291

Holmqvist, K., Andersson, R., Dewhurst, R., Jarodzka, H., and van de Weijer, J. (2011a). Eye tracking: a comprehensive guide to methods and measures. Oxford: Oxford University Press.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., and van de Weijer, J. (2011b). Eye tracking. Oxford: Oxford University Press.

Holsanova, J. (2014). "Reception of multimodality: applying eye tracking methodology in multimodal research" in Routledge handbook of multimodal analysis (Abingdon: Routledge), 285–296.

Huang, L. T., Chiu, C. A., Sung, K., and Farn, C. K. (2011). A comparative study on the flow experience in web-based and text-based interaction environments. *Cyberpsychol Behav Soc Netw*, 14, 3–11. doi: 10.1089/cyber.2009.0256

Huth, F., Koch, M., Awad-Mohammed, M., Weiskopf, D., and Kurzhals, K. (2024). Eye tracking on text reading with visual enhancements. Proceedings of the 2024 Symposium on Eye Tracking Research and Applications. 1–7

Hvelplund, K. T. (2011). Allocation of cognitive resources in translation: an eye-tracking and key-logging study. Copenhagen Business School (CBS): Frederiksberg.

Hyönä, J., and Ekholm, M. (2016). Background speech effects on sentence processing during reading: an eye movement study. *PLoS One* 11:e0152133. doi: 10.1371/journal.pone.0152133

Hyönä, J., and Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: effects of word length and word frequency. *J. Exp. Psychol. Learn. Mem. Cogn.* 21:1430. doi: 10.1037//0278-7393.21.6.1430

Indrarathne, B., and Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions: an eye-tracking study. *Stud. Second. Lang. Acquis.* 39, 401–430. doi: 10.1017/S027226311600019X

Ishibashi, K., Xiao, C., and Yada, K. (2019) Study of the effects of visual complexity and consumer experience on visual attention and purchase behavior through the use of eye tracking. 2019 IEEE International Conference on Big Data (Big Data) 2664–2673. IEEE

Jancsary, D., Höllerer, M., and Meyer, R. (2016). "Critical analysis of visual and multimodal texts" in Methods of critical discourse studies. eds. R. Wodak and M. Meyer (London: SAGE Publications), 180–204.

Jewitt, C. (Ed.). (2009). The Routledge handbook of multimodal analysis. London: Routledge. 340. Available at: https://techstyle.lmc.gatech.edu/wp-content/uploads/2012/08/Jones-2009.pdf

Jewitt, C. (2013). "Multimodal methods for researching digital technologies" in The SAGE handbook of digital technology research (London: SAGE Publications), 250–265.

Jones, M. W., Obregón, M., Kelly, M. L., and Branigan, H. P. (2008). Elucidating the component processes involved in dyslexic and non-dyslexic reading fluency: an eye-tracking study. *Cognition* 109, 389–407. doi: 10.1016/j.cognition.2008.10.005

Józsa, E., and Hámornik, B. P. (2012). Find the difference: eye tracking study on information seeking behavior using an online game. *Eye Tracking Vis. Cogn. Emotion* 2, 407–420. doi: 10.1093/iwc/iwv015

Just, M. A., and Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychol. Rev.* 87, 329–354. doi: 10.1037/0033-295X.87.4.329

Kamińska, D., Sapiński, T., Wiak, S., Tikk, T., Haamer, R., Avots, E., et al. (2019). Virtual reality and its applications in education: survey. *Information* 10:318. doi: 10.3390/info10100318

Kapp, S., Barz, M., Mukhametov, S., Sonntag, D., and Kuhn, J. (2021). ARETT: augmented reality eye tracking toolkit for head mounted displays. *Sensors* 21:2234. doi: 10.3390/s21062234

Karatekin, C., and Asarnow, R. F. (1998). Components of visual search in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder. *J. Abnorm. Child Psychol.* 26, 367–380. doi: 10.1023/A:1021903923120

Kerchner, J. L. (2014). Music across the senses: listening, learning, and making meaning. Oxford: Oxford University Press.

Klein, C., Seernani, D., Ioannou, C., Schulz-Zhecheva, Y., Biscaldi, M., and Kavšek, M. (2019). "Typical and atypical development of eye movements" in Eye movement research: an introduction to its scientific foundations and applications (Cham: Springer), 635–701.

Koć-Januchta, M., Höffler, T., Thoma, G. B., Prechtl, H., and Leutner, D. (2017). Visualizers versus verbalizers: effects of cognitive style on learning with texts and pictures—an eye-tracking study. *Comput. Hum. Behav.* 68, 170–179. doi: 10.1016/j.chb.2016.11.028

Krebs, C., Falkner, M., Niklaus, J., Persello, L., Klöppel, S., Nef, T., et al. (2021). Application of eye tracking in puzzle games for adjunct cognitive markers: pilot observational study in older adults. *JMIR Serious Games* 9:e24151. doi: 10.2196/24151

Kress, G. (2010). Multimodality: a social semiotic approach to contemporary communication. Abingdon: Routledge.

Kress, G. R., and van Leeuwen, T. (2006). Reading images: the grammar of visual design. Abingdon: Routledge.

Kruger, J. L. (2012). Making meaning in AVT: eye tracking and viewer construction of narrative. *Perspectives* 20, 67–86. doi: 10.1080/0907676X.2011.632688

Kruger, J. L., and Steyn, F. (2014). Subtitles and eye tracking: Reading and performance. *Read. Res. Q.* 49, 105–120. doi: 10.1002/rrq.59

Lee, M., and Jung, J. (2021). Effects of textual enhancement and task manipulation on L2 learners' attentional processes and grammatical knowledge development: a mixed methods study. *Lang. Teach. Res.*, 28:3621688211034640. doi: 10.1177/13621688211034640

Lee, M., and Révész, A. (2018). Promoting grammatical development through textually enhanced captions: an eye-tracking study. *Mod. Lang. J.* 102, 557–577. doi: 10.1111/modl.12503

Li, Z., Butler, E., Li, K., Lu, A., Ji, S., and Zhang, S. (2018). Large-scale exploration of neuronal morphologies using deep learning and augmented reality. *Neuroinformatics* 16, 339–349. doi: 10.1007/s12021-018-9361-5

Liu, H. C., Lai, M. L., and Chuang, H. H. (2011). Using eye-tracking technology to investigate the redundant effect of multimedia web pages on viewers' cognitive processes. *Comput. Hum. Behav.* 27, 2410–2417. doi: 10.1016/j.chb.2011.06.012

Liu, M.-H., Tseng, T.-F., Yeh, R. C., and Chung, P. (2020). "Research on multimedia community teaching platform into problem-based learning (PBL)—an empirical study on the teaching mode of creative innovation and entrepreneurship courses for technical and vocational students" in Cognitive cities. IC3 2019. Communications in computer and information science (Singapore: Springer), 256–261.

Liu, S., and Yu, G. (2022). L2 learners' engagement with automated feedback: an eye-tracking study. *Lang. Learn. Technol.* 26, 78–105. Available at: https://doi.org/10125/73480

Loewen, S., and Inceoglu, S. (2016). The effectiveness of visual input enhancement on the noticing and L2 development of the Spanish past tense. *Stud. Second Lang. Learn. Teach.* 6, 89–110. doi: 10.14746/ssllt.2016.6.1.5

Mason, L., Pluchino, P., and Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: effects on processing and learning. *Contemp. Educ. Psychol.* 41, 172–187. doi: 10.1016/j.cedpsych.2015.01.004

Mason, L., Moessnang, C., Chatham, C., Ham, L., Tillmann, J., Dumas, G., et al (2022). Stratifying the autistic phenotype using electrophysiological indices of social perception. *Sci Transl Med*, 14, eabf8987. doi: 10.1126/scitranslmed.abf8987

Mayer, R. E., and Fiorella, L. (2021). The Cambridge handbook of multimedia learning. Cambridge: Cambridge University Press.

Mayer, R. E., and Moreno, R. (2010). "Techniques that reduce extraneous cognitive load and manage intrinsic cognitive load during multimedia learning" in Cognitive load theory. eds. J. L. Plass, R. Moreno and R. Brünken (Cambridge: Cambridge University Press), 131–152.

Mézière, D. C., Yu, L., Reichle, E. D., von der Malsburg, T., and McArthur, G. (2023). Using eye-tracking measures to predict reading comprehension. *Read. Res. Q.* 58, 425–449. doi: 10.1002/rrq.498

Moses, L., and Reid, S. (2021). Supporting literacy and positive identity negotiations with multimodal comic composing. *Lang. Lit.* 23, 1–24. doi: 10.20360/langandlit29502

Munoz, J., Yannakakis, G. N., Mulvey, F., Hansen, D. W., Gutierrez, G., and Sanchis, A. (2011) Towards gaze-controlled platform games. 2011 IEEE Conference on Computational Intelligence and Games (CIG'11). 47–54. IEEE

Papanastasiou, G., Drigas, A., Skianis, C., Lytras, M., and Papanastasiou, E. (2018). Virtual and augmented reality effects on K-12, higher and tertiary education students' twenty-first century skills. *Virtual Real.* 23:425. doi: 10.1007/s10055-018-0363-2

Pellicer-Sánchez, A., Tragant, E., Conklin, K., Rodgers, M., Llanes, A., and Serrano, R. (2018). L2 reading and reading-while-listening in multimodal learning conditions: an eye-tracking study. *ELT Res. Pap.* 18, 1–28. Available at: https://englishagenda.britishcouncil.org/sites/default/files/attachments/pub_h191_elt_l2_reading_and_reading-while-listening_in_multimodal_final.pdf

Pjesivac, I., Wojdynski, B. W., and Geidner, N. (2021). Television infographics as orienting response: an eye-tracking study of the role of visuospatial attention in processing of television news. *Electron. News* 15:193124312110395. doi: 10.1177/19312431211039500

Płużyczka, M. (2018). The first hundred years: a history of eye tracking as a research method. *Appl. Linguist. Pap.* 25, 101–116. doi: 10.32612/uw.25449354.2018.4

Rayner, K. (1978). Eye movements in reading and information processing. *Psychol. Bull.* 85:618. doi: 10.1037/0033-2909.85.3.618

Rayner, K., and Pollatsek, A. (2006). "Eye-movement control in reading" in Handbook of psycholinguistics (San Diego: Academic Press), 613–657.

Richter, P., von Spiegel, W., and Waldern, J. (2018). 55-2: invited paper: volume optimized and mirror-less holographic waveguide augmented reality head-up display. *SID Symp. Dig. Tech. Pap.* 49, 725–728. doi: 10.1002/sdtp.12382

Santos, M. E. C., Lübke, A. I. W., Taketomi, T., Yamamoto, G., Rodrigo, M. M. T., Sandor, C., et al. (2016). Augmented reality as multimedia: the case for situated vocabulary learning. *RPTEL*, 11, 1–23. doi: 10.1186/s41039-016-0028-2

Schiavo, G., Osler, S., Mana, N., and Mich, O. (2015). Gary: combining speech synthesis and eye tracking to support struggling readers Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia 417–421)

Schmidt-Weigand, F., Kohnert, A., and Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learn. Instr.* 20, 100–110. doi: 10.1016/j.learninstruc.2009.02.011

Serafini, F. (2011). Expanding perspectives for comprehending visual images in multimodal texts. *J. Adolesc. Adult. Lit.* 54, 342–350. doi: 10.1598/JAAL.54.5.4

Serafini, F. (2014). Reading the visual: an introduction to teaching multimodal literacy. New York: Teachers College Press, 71–82.

Serafini, F. (2015). Multimodal literacy: from theories to practices. *Lang. Arts* 92, 412–423. doi: 10.58680/la201527389

Serrano-Mamolar, A., Miguel-Alonso, I., Checa, D., and Pardo-Aguilar, C. (2023). Towards learner performance evaluation in iVR learning environments using eye-tracking and machine-learning. *Comunicar* 31, 9–19. doi: 10.3916/C76-2023-01

Shaaban, T. S., and Mohamed, A. M. (2024). Exploring the effectiveness of augmented reality technology on reading comprehension skills among early childhood pupils with learning disabilities. *J. Comput. Educ.* 11, 423–444. doi: 10.1007/s40692-023-00269-9

Shin, J. K. (2023). "Developing primary English learners' visual literacy for a multimodal world" in Innovative practices in early English language education (Cham: Springer), 101–127.

Shin, D. S., Cimasko, T., and Yi, Y. (2020). Development of metalanguage for multimodal composing: a case study of an L2 writer's design of multimedia texts. *J. Second. Lang. Writ.* 47:100714. doi: 10.1016/j.jslw.2020.100714

Smith, F. (2012). Understanding reading: a psycholinguistic analysis of reading and learning to read. Abingdon: Routledge.

Sun, J., Liu, Y., Wu, H., Jing, P., and Ji, Y. (2022). A novel deep learning approach for diagnosing Alzheimer's disease based on eye-tracking data. *Front. Hum. Neurosci.* 16:972773. doi: 10.3389/fnhum.2022.972773

Susac, A., Bubic, A., Kazotti, E., Planinic, M., and Palmovic, M. (2018). Student understanding of graph slope and area under a graph: a comparison of physics and nonphysics students. *Phys. Rev. Phys. Educ. Res.* 14. doi: 10.1103/physrevphyseducres.14.020109

Susac, A., Planinic, M., Bubic, A., Jelicic, K., and Palmovic, M. (2023). Linking information from multiple representations: an eye-tracking study. *Front. Educ.* 8:1141896. doi: 10.3389/feduc.2023.1141896

Sutrisno, D., Abidin, N. A. Z., Pambudi, N., Aydawati, E., and Sallu, S. (2023). Exploring the benefits of multimodal literacy in English teaching: engaging students through visual, auditory, and digital modes. *Glob. Synth.Educ. J.* 1, 1–14. doi: 10.61667/xh184f41

Tai, T. Y., Chen, H. H. J., and Todd, G. (2020). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Comput. Assist. Lang. Learn.* 35, 892–917. doi: 10.1080/09588221.2020.1752735

Taylor, J. B. (1965). The bender gestalt as a measure of intelligence and adjustment in the lower intellectual range. *J. Consult. Psychol.* 29:595. doi: 10.1037/h0022741

Tscholl, D. W., Rössler, J., Handschin, L., Seifert, B., Spahn, D. R., and Nöthiger, C. B. (2020). The mechanisms responsible for improved information transfer in avatar-based patient monitoring: multicenter comparative eye-tracking study. *J. Med. Internet Res.* 22:e15070. doi: 10.2196/15070

van der Gijp, A., Ravesloot, C. J., Jarodzka, H., van der Schaaf, M. F., van der Schaaf, I. C., van Schaik, J. P., et al. (2017). How visual search relates to visual diagnostic performance: a narrative systematic review of eye-tracking research in radiology. *Adv. Health Sci. Educ.* 22, 765–787. doi: 10.1007/s10459-016-9698-1

van der Sluis, F., van den Broek, E. L., van Drunen, A., and Beerends, J. G. (2018). Enhancing the quality of service of mobile video technology by increasing multimodal synergy. *Behav. Inform. Technol.* 37, 874–883. doi: 10.1080/0144929X.2018.1505954

van Gog, T., and Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learn. Instr.* 20, 95–99. doi: 10.1016/j.learninstruc.2009.02.009

Wade, N. J. (2020). Looking at Buswell's pictures. *J. Eye Mov. Res.* 13. doi: 10.16910/jemr.13.2.4

Walczyk, J. J., Tcholakian, T., Igou, F., and Dixon, A. P. (2014). One hundred years of reading research: successes and missteps of Edmund Burke Huey and other pioneers. *Read. Psychol.* 35, 601–621. doi: 10.1080/02702711.2013.790326

Walsh, M. (2010). Multimodal literacy: researching classroom practice. Australia: Primary English Teaching Association.

Weger, U. W., and Inhoff, A. W. (2006). Attention and eye movements in reading: inhibition of return predicts the size of regressive saccades. *Psychol. Sci.* 17, 187–191. doi: 10.1111/j.1467-9280.2006.01683.x

Wiegand, I., Petersen, A., Finke, K., Bundesen, C., Lansner, J., and Habekost, T. (2017). Behavioral and brain measures of phasic alerting effects on visual attention. *Front. Hum. Neurosci.* 11:176. doi: 10.3389/fnhum.2017.00176