Using Eye Tracking for Research on Learning and Computational Thinking

Fengfeng Ke[1]([✉]) [iD], Ruohan Liu[2], Zlatko Sokolikj[1], Ibrahim Dahlstrom-Hakki[3] [iD], and Maya Israel[2] [iD]

[1] Florida State University, Tallahassee, FL 32306-4453, USA
fke@fsu.edu
[2] University of Florida, Gainesville, FL 32611, USA
[3] TERC, Cambridge, MA 02140, USA

Abstract. This paper presents a conceptual discussion of the theoretical constructs and perspectives in relation to using eye tracking as an assessment and research tool of computational thinking. It also provides a historical review of major mechanisms underlying the current eye-tracking technologies, and a technical evaluation of the set-up, the data capture and visualization interface, the data mining mechanisms, and the functionality of freeware eye trackers of different genres. During the technical evaluation of current eye trackers, we focus on gauging the versatility and accuracy of each tool in capturing the targeted cognitive measures in diverse task and environmental settings—static versus dynamic stimuli, in-person or remote data collection, and individualistic or collaborative learning space. Both theoretical frameworks and empirical review studies on the implementation of eye-tracking suggests that eye-tracking is a solid tool or approach for studying computational thinking. However, due to the current constraints of eye-tracking technologies, eye-tracking is limited in acting as an accessible and versatile tool for tracking diverse learners' naturalistic interactions with dynamic stimuli in an open-ended, complex learning environment.

Keywords: Eye-tracking · Literature and technology review · Computational thinking

1 Introduction

Computational thinking (CT) refers to a notion of using algorithmic thinking and computational solutions to represent complex tasks, solve problems, and design systems [2, 6, 52, 56]. Prior research has conceptualized and studied CT as a) a problem-solving practice (e.g., algorithm design, testing, and debugging, or data organization and analysis), b) an assortment of computational concepts (e.g., sequence, parallelism, control), c) foundational cognitive processes related to algorithmic thinking (e.g., abstraction, pattern recognition and generalization), and d) dispositions and perspectives important for the enactment of the aforesaid elements (e.g., tolerance for ambiguity, persistence, collaboration). Despite emerging as an area of growing educational significance in K-12

education, conceptual and empirical research on CT education is still under-researched [3]. In particular, there is a need to explore effective methods and tools for assessing CT as a multifaceted competency developed and enacted during dynamic, contextualized practices. Recent research is starting to suggest the use of eye tracking for assessing or validating learning of CT, attitudes toward, cognitive engagement with, or processes and states of development in CT practices (e.g., [4, 39, 50]). However, this research approach is still emerging and is not yet a common practice.

Eye tracking has long been used in the field of cognitive psychology to study underlying cognitive constructs such as attention, memory formation, and processing difficulty [45, 46]. It has also long been considered one of the best measures of visual attention allocation [42], and hence a prominent approach for tracking learners' interaction with the external environment or stimuli. The use of eye tracking to study students' attention and explore their cognitive processes or efforts in authentic educational environments has been more limited but has garnered increasing interest in recent years [12]. Advances in both hardware and software based eye tracking solutions along with a reduction in their cost has made deployment of these solutions at scale possible [41, 55]. There is also an increasing need for a non-intrusive assessment or analysis tool that can track learners' engagement, cognitive processing patterns, and their cognitive-affective states in a personalized, highly interactive, or collaborative learning environment. In partnership with other multimodal or action-oriented data resources, eye tracking enables learning scientists to better study affordances of a learning environment along with learners' agency and conscientiousness.

Therefore, in this paper we intend to provide a conceptual discussion of the theoretical constructs and perspectives in relation to using eye tracking as a CT assessment tool. We will also survey the current eye-tracking technologies—freeware desktop, mobile, and web-based eye trackers—to present a technical evaluation of their set-up, data capture and visualization interface, data mining mechanisms, and functionality. During the technical evaluation of current eye trackers, we focus on gauging the versatility and accuracy of each tool in capturing the targeted cognitive measures in diverse task and environmental settings—static versus dynamic stimuli, in-person or remote data collection, individualistic or collaborative learning space, and neurodiverse learners.

2 Theoretical Constructs and Perspectives

Much of the prior work involving eye tracking and learning has focused on looking at memory formation with a focus on the cognitive processes involved. This has included looking at the visual attention, processing difficulty, working memory, and long-term memory aspects of simple tasks including sentence reading, visual search, category formation and list recall [46]. Prior research on eye tracking and learning have provided a great deal of understanding of the various cognitive elements that make up learning but have had limited impact on classroom learning [12]. In the following section, we provide a review of major theoretical perspectives and related constructs that should shed light on using eye tracking in the assessment and research of CT.

2.1 Eye-Mind Assumption (EMA) and Visual Attention

The relationship between eye movements and cognitive processes are based on two assumptions established by Just and Carpenter [23]: the immediacy assumption and eye-mind assumption. Assuming a linkage between a person's visual focus and cognitive focus, the immediacy and eye-mind assumption were often used as an operational basis for interpreting eye-tracking data. Just and Carpenter contended, "there is no appreciable lag between what is being fixated and what is being processed" and "the interpretations at all levels of processing are not deferred; they occur as soon as possible" (p. 331).

People's visual attention behavior involves two types of attention: overt attention and covert attention [14]. Overt attention is the act of intentionally directing one's attention towards visual stimuli, it happens when one selectively attends to one stimulus while others are ignored. Covert attention is a neural process, it happens when one attends to something without moving the eyes towards the object attended [15]. In the literature, eye-tracking methodology is emphasized to provide a direct and objective measure of overt attention by capturing the timing and location of participants' visual focus during visual studies [17]. Covert attention, however, cannot be directly measured via eye-tracking technology, but it can be inferred by integrating eye-tracking measures with other types of measures such as behavioral data and physiological data [14, 30].

Eye-tracking has become a focus of interest in computer-supported collaborative learning (CSCL) research. It is considered a promising technique to examine and support visual attention coordination, or joint attention, in collaborative learning environments [47]. Joint attention is the ability to coordinate one's focus of attention with that of another person during a social interaction; it is "crucial for the development of social communication, learning and the regulation of interpersonal relationships" ([10], p. 502). The literature suggests that eye gaze, earlier than language and pointing gestures, is typically the initial communicative channel one develops and relies on to experience joint attention in social interactions [43]. With the enhancement of the eye-tracking technology in measuring subtle changes in visual attention, eye-tracking becomes a robust method to detect joint attention (e.g., [34]).

Previous studies emphasized that the ability to establish joint attention is crucial for a group to establish a common mental model and empathy in collaborative problem solving (e.g., [48]). Joint attention was measured in multiple ways in these studies. For example, Papavlasopoulou et al. [38] compared eye fixations of participants in two different groups and examined the level of their gaze similarity (e.g., spatial dispersion). Schneider et al. [48] constructed a metric for joint attention by incorporating the captured gaze points with two additional parameters: latency and distance between gazes. Pietinen, Bednarik, and Tukiainen [44] developed a new joint visual attention metric using the number of overlapping fixations and fixation duration of overlapping fixation to examine the quality of collaboration.

2.2 Engagement

Engagement and its measurement have long been a focus of research in human-computer interaction [37]. Traditionally, engagement is measured through self-report instruments [16]. Eye-tracking can compensate for the weaknesses of self-report measures (e.g.,

honesty, sampling bias) and is gaining growing popularity in research of engagement. At a lower end, engagement can be measured by the act of paying attention [33]. O'Brien and Toms [36] argued that user engagement manifests through the observed interest and visual attention when the user interacts with a designated tool. Extending this view, Bassett and Green [7] concluded that visual attention provides an important lens to understand cognitive engagement. Based on the eye-mind assumption that people tend to engage with what they visually attend to, eye-tracking metrics such as fixation frequency and fixation duration are widely used to indicate the level of engagement in learning. Specifically, higher fixation frequency and longer fixation duration are linked to higher levels of engagement in learning. As an effective tool to measure micro-level engagement [33], eye-tracking has been utilized along with other types of measures (e.g., performance, self-report, and physiological measures) to capture and assess multifaceted engagement in learning (e.g., [24]).

2.3   Inferring Cognitive Processes, States, and Traits via Eye Tracking

Current eye-tracking technologies allow researchers to trace participants' eye movements with minimum intrusiveness, which makes the eye-tracking data a solid inference of the cognitive or information processing patterns, states, or traits [1, 54]. In the research of multimedia learning, eye-tracking is frequently used to study how learners visually process different formats of information. The literature has established a number of eye-tracking metrics that are commonly used to infer participants' cognitive processes during learning [22]. For example, the number of fixations overall is a widely used eye-tracking measure in HCI studies, it is thought to be negatively associated with participants' searching efficiency [19]. A greater number of fixations indicate lower searching efficiency which possibly due to the poor design of visual elements display [22]. Frequency of fixations on a specific area or element demonstrates the importance of the fixated area or element. Additionally, the time one spends gazing at a particular component of a visual scene designates the content he/she is visually engaged with [32].

Fixation duration is a commonly used metric to measure the level of processing difficulty in learning [51]. A longer fixation duration on a stimulus indicates greater processing difficulty associated with the stimulus (e.g., [35]). Saccades, the quick movements between fixations, is another cognitive measure related to eye-tracking [18]. Saccade is believed to relate to the change of focus in visual attention and interest in learning (e.g., [18]). Typical saccade-based metrics include number of saccades, saccadic amplitude (i.e. saccadic distance), saccade regressions, saccadic duration, or saccadic velocity [9]. Previous studies suggested that fewer saccades are associated with less mental effort during task performance (e.g., [39]). Saccade amplitude, specifically, is used to gauge the cognitive processes that involve planning and hypothesis testing [11].

2.4   Cognitive Load and Effort

Cognitive load is a commonly examined cognitive construct in eye-tracking research (e.g., [29]). The cognitive load theory implies that humans have a limited capacity and duration of working memory, and the amount of information one is able to process and temporarily store in working memory cannot exceed the limit of the capacity [53].

Fixation counts and fixation duration are typical metrics used to infer the mental effort participants exert or the cognitive load they experience during a task (e.g., [58]). According to Obaidellah et al. [35], low fixation time and counts link to less effort in mental processing, while long fixation time and high fixation counts indicate more effort is warranted for the task. In addition to these fixation-based measures, another important metric in identifying cognitive load is pupil size [32]. Pupil size has in recent years been used to examine the degree of cognitive load participants experience when accomplishing a task. Prior research found that pupil dilation increases when a task is perceived cognitively demanding (e.g., [21]).

3   Prior Eye-Tracking Reviews

A number of prior eye-tracking review studies have been conducted to provide an overview of eye-tracking methodology, how it connects to research of HCI, learning, and cognitive science, as well as the merits and challenges associated with using eye-tracking for educational and research purposes (e.g., [1, 8, 25, 31, 49]). In this section, we provide a synthesis of the major findings and discussions of these review studies.

Alemdag and Cagiltay [1], for example, conducted a systematic review of eye-tracking research in the domain of multimedia learning. This work revealed the popularity of temporal and count scales of eye-tracking measurements in multimedia research, providing three cognitive processes: selecting, organizing, integrating. The authors emphasized the necessity to bridge the current gap of eye-tracking studies with young participants, and advocated more effort to be invested in research in the K-12 contexts. Additionally, future studies should consider including qualitative analysis of eye movement measures to supplement and support quantitative eye-tracking measures.

Focusing on the application of eye-tracking methods in spatial cognition, Kiefer et al. [25] provided a review of recent literature and claimed that the potentials of mobile eye-trackers in real-world studies have only just started to be exploited. However, two main challenges of current mobile eye-tracking studies should be acknowledged: a) the processing of mobile eye-tracking data is labor-intensive; b) the real-world environment is hard to control. Therefore, the trade-off between internal and external validity is a particular challenge for eye-tracking studies in authentic environments. For future eye-tracking research in spatial cognition, the author suggests that more effort be devoted to advance the traditional eye-tracking analysis, beyond the classic fixation, saccade or scan-path measures, and consider deconstructing the complexity of the interplay between ambient and focal attention [27], and participants' switching patterns across areas of interest [28].

Blascheck et al. [8] performed a survey of 90 publications to examine visualization techniques used in eye-tracking studies. Overall, nine types of visualization techniques were identified while some existing challenges and unanswered questions were revealed. For example, capturing participants' interactions with dynamic stimuli still emerged as a major challenge in current eye-tracking studies. Given the growing trend of research involving multimodal data analysis, developing multimodal data visualization techniques (e.g., combining eye-tracking data with other sensor information from EEG devices, or skin-resistance measurements) can be a potential focus in future eye-tracking research.

3.1 Summary

Multiple salient and common themes have emerged from the findings of these previous eye-tracking review studies. First, the diversity and richness of quantitative metrics of eye-tracking as inferences of the aforementioned theoretical constructs on cognition and learning are unanimously reported by prior reviews. The parameters involved in these quantitative eye-tracking metrics are generally specified in the previous studies. On the other hand, an integral analysis with these eye-tracking metrics to infer on a comprehensive or multifaceted cognitive process, state, or trait of learners is still lacking. The previous reviews also reported a concern on the lack of cross-validation and in-depth analysis with the current quantitative eye-tracking data [1, 8, 25, 31, 49]. Second, even though eye-tracking has been frequently used to investigate the dynamics between a learner (user) and an interactive, computerized learning environment, the designated stimuli or areas of interest in such an environment are typically fixed, pre-defined, and constrained. Due to the limited functionality of current eye-trackers, eye tracking is generally used in a highly controlled lab setting and falls short of capturing and measuring the learners' interactions with dynamic or emerging stimuli in an open-ended learning space. Ultimately, prior research on how the implementation of eye-tracking will address the needs of neurodiverse users or learners with diverse cognitive or physiological characteristics is generally missing. As such, sampling bias is an innate issue in eye-tracking research.

4 A Survey and Evaluation of Existing Eye-Tracking Technologies

4.1 Introduction

From the onset of the 20th century, scientific research has been using basic eye movements such as saccadic suppression, saccade latency, and the size of the perceptual span to make deductions in the field of behavioral experimental psychology. The fast stream of technological development has allowed eye tracking to trickle into other fields of research, with hardware and software algorithms allowing researchers of different scientific fields to incorporate eye tracking into their scientific endeavors. In addition, the availability of eye tracking has made it present not only in scientific research but also in the commercial world. Initially, eye tracking relied on hardware specifically built for eye tracking, such as Charles Judd's eye movements camera [59]. However, the fast opportunities that eye tracking offered to the commercial world has pushed for faster development and a shift from the expensive, intrusive and specially built apparatus to software that works with any general webcam.

With the introduction of eye-tracking in both the commercial world and scientific research, a unique field of research has emerged to improve the accuracy of the eye tracking algorithms as well as the real-time information they provide. Various companies are interested in tracking the gaze of potential customers on their infomercials, while the scientific community develops a similar goal of pinpointing the gaze of a user of a specific platform. One way of estimating this gaze is using real-time images of certain facial features and specific interactions with a platform to estimate the part of the platform at which the user is looking. This appearance-based method does not require specific

cameras and can be done using a regular webcams, but is limited by the computational power of the estimation algorithm. Conversely, a different method relies on a more technologically advanced camera to extract features not available to the naked, and runs specific models with these unseen features to estimate gaze. One such example of this model based method is an infrared camera that uses the IR glint of the eye to estimate the gaze. These methods may not be computationally expensive but require specific hardware to function.

The appearance-based method has opened up an interdisciplinary field of computational mathematics that aims to output the most accurate gaze estimation, given an input of images of specific facial features of a user and interactions of said user with the medium. This goal has been approached from all the different corners of computational mathematics and data mining. The model is generally a two-step process, consisting of an algorithm extracting facial features, such as the pupils of the eyes, and the gaze prediction algorithm, correlating those features with specific points of the screen. The simplest gaze prediction algorithm is the interpolation of interactions with a known gaze point to produce a curve from which the gaze point of other interactions can be estimated. The model is a good starting point as it makes certain assumptions about the user's interactions and has been used in certain papers where these assumptions are valid. In addition, this method is not computational expensive so therefore can be done on various platforms and mediums.

4.2   Evaluation of Freeware Eye-Trackers

WebGazer. An exemplary embodiment of the appearance based method is WebGazer [40] with the additional feature of data validation, in the form of data points acquired during specific mouse movements. WebGazer uses the facial recognition algorithms to identify the pupils in the image which chronologically corresponds to a specific mouse click. This facial recognition algorithm converts the pixels that make up the pupils into a 120D feature vector. Afterward, this feature vector, along with screen coordinates of the corresponding click, form the linear interpolation with which future gaze points are estimated. WebGazer does not require an initial calibration. The parameters of the linear interpolation are formed as the software is being used, meaning that it might take some time for the said parameters to stabilize around a specific value and consequentially for the program to make stable gaze predictions.

WebGazer algorithm is implemented in JavaScript, which can be included on the HTML side of any website. Therefore any activity that would require the use of WebGazer would have to be hosted on a specific website. WebGazer uses Clmtrackr for the facial recognition portion of the algorithms but has the option to incorporate other tools or libraries. Lastly, WebGazer has been designed with the anonymity of the website visitor in mind, therefore as it is being hosted on a website, it only records the estimated gaze coordinates of a website visitor and aggregates them with data collected of other users. This aggregated data provides a general overview of the estimated gaze of all visitors. This feature makes WebGazer ideal for discovering what captures the attention of most visitors of a website, but also makes WebGazer very difficult to use in applications where the gaze of a specific user needs to be tracked continuously and known at each time point.

PACE. The algorithm behind PACE [20] works in a similar manner as that of WebGazer: the gaze estimation algorithm is trained as the software is being used. It also uses the Clmtrackr to acquire the facial features, and packages them into 120D feature vectors, expanding the data collected to the whole face instead of being limited to the pupils. PACE also collects the corresponding gaze screen coordinates from mouse clicks and other mouse interactions. But before those mouse-click coordinates are used as part of the training data, they undergo a rigorous process that validates whether the position of the mouse matched the user's gaze at that moment through behavioral and data-driven validation. The first hundred validated data points, both screen coordinates and corresponding feature vectors are fed into a random decision forest to train the gaze prediction algorithms that estimates the gaze from then on. In conclusion, PACE, like WebGazer, does not require an initial calibration from the user but does not produce any gaze estimation until a specific amount of acceptable data has been gathered. Unlike WebGazer, PACE does not constantly retrain its algorithm and therefore the gaze estimation parameters stabilize conclusively. PACE is a standalone system that can extract data from both live webcam recordings and pre-recorded webcam videos, and therefore can easily be used in situations where the exact movement of the gaze of a specific user needs to be known.

TuckerGaze. Venturing further from the simple design of the gaze prediction with linear interpolation and random forest, there are other algorithms that use more complex mathematical methods but require a bit more computational power. This trade-o limits the applications and settings in which they could be used. Like the previously mentioned software, TuckerGaze [57] extracts facial features of images from its training data using Clmtrackr. However, the training data is not gathered throughout the use of the software; instead it is collected before the software is used, through an initial calibration. The collected training data is then trained on an algorithm with Ridge Regression and the resulting model is used for future gaze prediction. After its use online, TuckerGaze reevaluates its prediction model using SVC with Gaussian Kernel. TuckerGaze has been hosted on websites and has only been used for specific gamelike interactions, to produce gaze estimation during preordained stimuli. It does not have the variability of use as the aforementioned software, but presents an important evolutionary landmark in appearance based gaze estimation.

Gaze Capture. Another important evolutionary landmark in eye-tracking algorithm is the introduction of artificial neural networks. The recent widespread utilization of artificial neural networks (ANNs) in the computational world for the purposes of classification and regression has inspired attempts of incorporating ANNs in gaze estimation. The team behind Gaze Capture [26] built a training data set from collecting mobile phones camera face recordings of a large number of participants, following guided movements on screen. The cropped up images of faces and eyes from each recording and the assumed corresponding gaze point were used to train a neural network. This neural network is the basis for the gaze prediction software behind Gaze Capture. Such a setup, through its incorporation of mobile phones, allows for fast training of neural networks. But it does not translate into accurate predictions outside of the environment in which the training data was made. Therefore it may be unsuitable for gaze estimations in environments where the screen used is significantly larger than the screen of a mobile device or tablet.

OpenFace 2.2. While not specifically designed for eye tracking, OpenFace [5] is a popular open source toolkit that is capable of identifying eye location, eye landmarks, head orientation, and gaze direction. OpenFace is primarily designed for facial behavior analysis. The software can identify 68 facial landmarks and based on those landmarks can identify 12 facial Action Units (AUs). While it doesn't directly detect emotions the detected AUs can be mapped onto emotion states based on the Facial Action Coding System (FACS) [13].

OpenFace doesn't provide actual eye tracking, that is, it does not provide a direct estimate of what the subject is viewing on the screen. Instead of training a neural network to map visual input onto screen locations, OpenFace instead tries to model the eye itself and maps the visual data onto eyeball orientation. OpenFace is also estimates head location and head orientation which in combination with the eye orientation data should allow for an estimate of the gaze location on any surface relative to the camera location.

In order to assess the accuracy of OpenFace's gaze orientation algorithm, we collected data from four participants. All participants had normal vision. Two participants were adults and two were children ages 8 and 10. Participants were asked to sit two feet in front of a laptop computer running OpenFace. Participants were asked to perform the follow series of head and eye movements:

– Fixate the center of the left, right, top and bottom edges of the screen without moving their heads
– Move their heads up, down, left, and right while fixating the center of the screen

Changes in the magnitude and direction of their gaze was measured as they performed these actions. The results were consisted across all four participants, see Table 1 for details.

Table 1. Results of observed impact on detected gaze orientation in OpenFace 2.2 as eye and head orientation is manipulated independently.

Action	Prediction	Observation
Look left, head fixated	Gaze shifts to the left	Gaze shifts to the left
Look right, head fixated	Gaze shifts to the right	Gaze shifts to the right
Look up, head fixated	Gaze shifts up	Little to no shift in gaze
Look down, head fixated	Gaze shifts down	Little to no shift in gaze
Fixate center, head moves left	No change in gaze	Small gaze shift to the left
Fixate center, head moves right	No change in gaze	Small gaze shift to the right
Fixate center, head moves up	No change in gaze	Gaze shifts up
Fixate center, head moves down	No change in gaze	Gaze shifts down

The findings of our small assessment of the accuracy of gaze orientation detection in OpenFace 2.2 indicates that the software algorithm has a difficult time distinguishing

between head and eye movements, even though it is able to detect each independently. This challenge is a common one in eye tracking research which is why most lab research involving eye tracking uses a protocol that eliminates head movement. More troubling however is that OpenFace was unable to detect vertical shifts in eye orientation if they were not accompanied by head movements. This makes OpenFace's algorithm only useful for very coarse estimates of visual attention allocation.

5   Conclusion and Implication

Both theoretical frameworks and empirical review studies on the implementation of eye-tracking suggest that eye-tracking is a solid tool or approach for not only capturing observed engagement behaviors or states (e.g., visual engagement and joint attention coordination), but also inferring covert cognitive processes or traits (e.g., information processing patterns, cognitive effort or commitment). A rich set of eye-tracking metrics have been delineated and infield tested as measures for these cognitive variables in the research of learning and HCI. All these prominent facets of eye-tracking apply to CT, a componential area of both learning and HCI. However, due to the current constraints of eye-tracking technologies, eye-tracking is limited in acting as an accessible and versatile tool for tracking diverse learners' naturalistic interactions with dynamic stimuli in an open-ended, complex learning environment. Such a limit imposes a conflict between eye-tracking and the current CT education that highlights an inclusive or adaptive design for neurodiversity as well as authentic, constructionism-oriented learning activities. Notably, the recent development of deep learning and data mining can potentially push the boundary of eye-tracking algorithms to enhance its accuracy and versatility in gaze capturing and prediction in a more naturalistic and versatile way. It is also warranted for cognitive and learning scientists to conduct more systematic research on the methods of data fusion and integral analysis that will enhance and cross-validate the interpretation of the eye-tracking data along with that of other multimodal behavioral and physiological data.

References

1. Alemdag, E., Cagiltay, K.: A systematic review of eye tracking research on multi- media learning. Comput. Educ. **125**, 413–428 (2018)
2. Anderson, N.D.: A call for computational thinking in undergraduate psychology. Psychol. Learn. Teach. **15**(3), 226–234 (2016)
3. Angeli, C., Giannakos, M.: Computational thinking education: issues and challenges (2020)
4. Arslanyilmaz, A., Corpier, K.: Eye tracking to evaluate comprehension of computational thinking. In: Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education, p. 296 (2019)
5. Baltrusaitis, T., Zadeh, A., Lim, Y.C., Morency, L.P.: Openface 2.0: facial behavior analysis toolkit. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). pp. 59–66. IEEE (2018)
6. Barr, D., Harrison, J., Conery, L.: Computational thinking: a digital age skill for everyone. Learn. Lead. Technol. **38**(6), 20–23 (2011)

7. Bassett, D., Green, A.: Engagement as visual attention: a new story for publishers. In: Publishing and Data Research Forum, London, pp. 17–20 (2015)

8. Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., Ertl, T.: State- of-the-art of visualization for eye tracking data. In: EuroVis (STARs) (2014)

9. Borys, M., Plechawska-Wójcik, M.: Eye-tracking metrics in perception and visual attention research. EJMT **3**, 11–23 (2017)

10. Caruana, N., et al.: Joint attention difficulties in autistic adults: an interactive eye- tracking study. Autism **22**(4), 502–512 (2018)

11. Cowen, L., Ball, L.J., Delin, J.: An eye movement analysis of web page usability. In: People and Computers XVI-Memorable Yet Invisible, pp. 317–335. Springer (2002). https://doi.org/10.1007/978-1-4471-0105-5_19

12. Dahlstrom-Hakki, I., Asbell-Clarke, J., Rowe, E.: Showing is knowing: the potential and challenges of using neurocognitive measures of implicit learning in the classroom. Mind Brain Educ. **13**(1), 30–40 (2019)

13. Ekman, R.: What the Face Reveals: Basic and Applied Studies of Spontaneous Expression using the Facial Action Coding System (FACS). Oxford University Press, Oxford (1997)

14. Ellis, N.C., Hafeez, K., Martin, K.I., Chen, L., Boland, J., Sagarra, N.: An eye- tracking study of learned attention in second language acquisition. Appl. Psycholinguist. **35**(3), 547–579 (2014)

15. Findlay, J.M., Findlay, J.M., Gilchrist, I.D., et al.: Active Vision: The Psychology of Looking and Seeing, vol. 37, Oxford University Press, Oxford (2003)

16. Fredricks, J.A., McColskey, W.: The measurement of student engagement: a compartive analysis of various methods and student self-report instruments. In: Handbook of Research on Student Engagement, pp. 763–782. Springer (2012). https://doi.org/10.1007/978-1-4614-2018-7_37

17. Godfroid, A.: Eye tracking. Routledge encyclopedia of second language acquisition, pp. 234–236 (2012)

18. van Gog, T., Jarodzka, H.: Eye tracking as a tool to study and enhance cognitive and metacognitive processes in computer-based learning environments. In: Azevedo, R., Aleven, V. (eds.) International Handbook of Metacognition and Learning Technologies. SIHE, vol. 28, pp. 143–156. Springer, New York (2013). https://doi.org/10.1007/978-1-4419-5546-3_10

19. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. **24**(6), 631–645 (1999)

20. Huang, M.X., Kwok, T.C., Ngai, G., Chan, S.C., Leong, H.V.: Building a personalized, auto-calibrating eye tracker from user interactions. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 5169–5179 (2016)

21. Hyönä, J., Tommola, J., Alaja, A.M.: Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. Q. J. Exp. Psychol. **48**(3), 598–612 (1995)

22. Jacob, R.J., Karn, K.S.: Eye tracking in human-computer interaction and usability research: ready to deliver the promises. In: The Mind's Eye, pp. 573–605. Elsevier (2003)

23. Just, M.A., Carpenter, P.A.: A theory of reading: from eye fixations to comprehension. Psychol. Rev. **87**(4), 329 (1980)

24. Kaakinen, J.K., Ballenghein, U., Tissier, G., Baccino, T.: Fluctuation in cognitive engagement during reading: evidence from concurrent recordings of postural and eye movements. J. Exp. Psychol. Learn. Mem. Cogn. **44**(10), 1671 (2018)

25. Kiefer, P., Giannopoulos, I., Raubal, M., Duchowski, A.: Eye tracking for spatial research: Cognition, computation, challenges. Spat. Cogn. Comput. **17**(1–2), 1–19 (2017)

26. Krafka, K., et al.: Eye tracking for everyone. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016

27. Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A., Kopacz, A.: Discerning ambient/focal attention with coefficient k. ACM Trans. Appl. Perception (TAP) **13**(3), 1–20 (2016)
28. Krejtz, K., et al.: Gaze transition entropy. ACM Trans. Appl. Perception (TAP) **13**(1), 1–20 (2015)
29. Kruger, J.L., Doherty, S.: Measuring cognitive load in the presence of educational video: towards a multimodal methodology. Australas. J. Educ. Technol. **32**(6) (2016)
30. Kulke, L.V., Atkinson, J., Braddick, O.: Neural differences between covert and overt attention studied using EEG with simultaneous remote eye tracking. Front. Hum. Neurosci. **10**, 592 (2016)
31. Lai, M.L., et al.: A review of using eye-tracking technology in exploring learning from 2000 to 2012. Educ. Res. Rev. **10**, 90–115 (2013)
32. Liu, H.C., Lai, M.L., Chuang, H.H.: Using eye-tracking technology to investigate the redundant effect of multimedia web pages on viewers' cognitive processes. Comput. Hum. Behav. **27**(6), 2410–2417 (2011)
33. Miller, B.W.: Using reading times and eye-movements to measure cognitive engagement. Educ. Psychol. **50**(1), 31–42 (2015)
34. Navab, A., Gillespie-Lynch, K., Johnson, S.P., Sigman, M., Hutman, T.: Eye- tracking as a measure of responsiveness to joint attention in infants at risk for autism. Infancy **17**(4), 416–431 (2012)
35. Obaidellah, U., Al Haek, M., Cheng, P.C.H.: A survey on the usage of eye-tracking in computer programming. ACM Comput. Surv. (CSUR) **51**(1), 1–58 (2018)
36. O'Brien, H.L., Toms, E.G.: What is user engagement? A conceptual framework for defining user engagement with technology. J. Am. Soc. Inform. Sci. Technol. **59**(6), 938–955 (2008)
37. O'Brien, H.L., Cairns, P., Hall, M.: A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. Int. J. Hum Comput Stud. **112**, 28–39 (2018)
38. Papavlasopoulou, S., Sharma, K., Giannakos, M., Jaccheri, L.: Using eye-tracking to unveil differences between kids and teens in coding activities. In: Proceedings of the 2017 Conference on Interaction Design and Children, pp. 171–181 (2017)
39. Papavlasopoulou, S., Sharma, K., Giannakos, M.N.: How do you feel about learning to code? Investigating the effect of children's attitudes towards coding using eye- tracking. Int. J. Child-Comput. Interact. **17**, 50–60 (2018)
40. Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., Hays, J.: Webgazer: scalable webcam eye tracking using user interactions. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, pp. 3839–3845 (2016)
41. Park, S., Aksan, E., Zhang, X., Hilliges, O.: Towards end-to-end video-based eye-tracking. In: Vedaldi, A., Bischof, Horst, Brox, T., Frahm, J.-M. (eds.) ECCV 2020. LNCS, vol. 12357, pp. 747–763. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58610-2_44
42. Peterson, M.S., Kramer, A.F., Irwin, D.E.: Covert shifts of attention precede involuntary eye movements. Percept. Psychophys. **66**(3), 398–405 (2004)
43. Pfeiffer, U.J., Vogeley, K., Schilbach, L.: From gaze cueing to dual eye-tracking: novel approaches to investigate the neural correlates of gaze in social interaction. Neurosci. Biobehav. Rev. **37**(10), 2516–2528 (2013)
44. Pietinen, S., Bednarik, R., Tukiainen, M.: Shared visual attention in collaborative programming: a descriptive analysis. In: Proceedings of the 2010 ICSE Workshop on Cooperative and Human Aspects of Software Engineering, pp. 21–24 (2010)
45. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychol. Bull. **124**(3), 372 (1998)
46. Rayner, K.: The 35th sir frederick bartlett lecture: eye movements and attention in reading, scene perception, and visual search. Q. J. Exp. Psychol. **62**(8), 1457–1506 (2009)

47. Schneider, B., Pea, R.: Real-time mutual gaze perception enhances collaborative learning and collaboration quality. Int. J. Comput.-Support. Collab. Learn. **8**(4), 375–397 (2013). https://doi.org/10.1007/s11412-013-9181-4

48. Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., Pea, R.: Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. Int. J. Comput.-Supported Collab. Learn. **13**(3), 241–261 (2018)

49. Sharafi, Z., Soh, Z., Guéhéneuc, Y.G.: A systematic literature review on the usage of eye-tracking in software engineering. Inf. Softw. Technol. **67**, 79–107 (2015)

50. Sharma, K., Papavlasopoulou, S., Giannakos, M.: Coding games and robots to en- hance computational thinking: How collaboration and engagement moderate children's attitudes? Int. J. Child-Comput. Interact. **21**, 65–76 (2019)

51. Shojaeizadeh, M., Djamasbi, S., Trapp, A.C.: Density of gaze points within a fixation and information processing behavior. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2016. LNCS, vol. 9737, pp. 465–471. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40250-5_44

52. Shute, V.J., Sun, C., Asbell-Clarke, J.: Demystifying computational thinking. Educ. Res. Rev. **22**, 142–158 (2017)

53. Sweller, J.: Cognitive load during problem solving: effects on learning. Cogn. Sci. **12**(2), 257–285 (1988)

54. Underwood, G., Radach, R.: Eye guidance and visual information processing: reading, visual search, picture perception and driving. In: Eye Guidance in Reading and Scene Perception, pp. 1–27. Elsevier (1998)

55. Valliappan, N., et al.: Accelerating eye movement research via accurate and affordable smartphone eye tracking. Nat. Commun. **11**(1), 1–12 (2020)

56. Wing, J.M.: Computational thinking. Commun. ACM **49**(3), 33–35 (2006)

57. Xu, P., Ehinger, K.A., Zhang, Y., Finkelstein, A., Kulkarni, S.R., Xiao, J.: Turkergaze: crowdsourcing saliency with webcam based eye tracking. arXiv preprint arXiv:1504.06755 (2015)

58. Zagermann, J., Pfeil, U., Reiterer, H.: Measuring cognitive load using eye tracking technology in visual computing. In: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization, pp. 78–85 (2016)

59. Judd ,C.H.: Psychol. Rev. Monoh. Suppl. **VII**(35) (1907)