Training, Education, and Learning Sciences, Vol. 59, 2022, 316–322
https://doi.org/10.54941/ahfe1002416

Eye-Tracking in Simulator Training and Assessment: A Semi-Structured Meta-Review

Charlott Sellberg[1], Gesa Praetorius[2,3], and Markus Nivala[1]

[1]University of Gothenburg, 405 30, Gothenburg, Sweden
[2]Swedish National Road and Transport Institute, 102 15 Stockholm, Sweden
[3]University of South-Eastern Norway, 3199 Borre, Norway

ABSTRACT

The aim of this meta-review is to chart out a) how eye-tracking is currently used in training and assessment, and b) the barriers and benefits of using eye-tracking reported in the literature. The results show that eye-tracking is used for identifying differences between novices and experts and for capturing expert gaze patterns. Different gaze patterns can be used for gaze training, formative assessment of non-technical skills, as well as summative assessment of technical skills. Further, it is appreciated as an unobtrusive technology that enables a quantitative assessment of objective parameters. Hence, the promise of using eye-tracking for informing instructional design as well as for designing training systems and improve learning environments is clear. However, it is important to consider that implementing eye-tracking is a rather expensive and time-consuming endeavor that requires carefully designed tasks and task analyses to fulfil its potential.

Keywords: Eye-tracking, Simulator training, Simulator assessment, Meta-review

INTRODUCTION

The potential for using eye-tracking as quantitative and objective tools for training and assessment has gained substantial attention across domain with high standards for safety (Rosch & Vogel-Walcutt, 2013). An initial and unstructured search for literature on eye-tracking applications for training and assessment found a large number of reviews that explore the potentials of using eye-tracking technologies in socio-technical working environments (Rosch & Vogel-Walcutt, 2013; Ashraf et al. 2018; Fox & Faulkner-Jones, 2017; Merali et al., 2017; Tien et al. 2014). Given the considerable number of studies conducted on this topic, the aim of this study is to synthesize results from previous reviews on eye-tracking support in training and assessment for work in complex socio-technical domains. In particular, the objective is to map out a) how eye-tracking is used in training and assessment, and b) the barriers and benefits of using eye-tracking reported in the literature, in order to formulate guidelines for implementing eye-tracking in the study of visual expertise in socio-technical environments.

The increasing availability of relatively inexpensive, user-friendly, unobtrusive and non-intrusive eye-tracking technology has resulted in increasing number of eye-tracking studies across scientific disciplines and domains. Improvements in eye-tracking technology (e.g., sampling rate, accuracy, fewer physical restrictions) have enabled researchers to capture previously unavailable data and measures (Klein & Ettinger, 2019). Technological development has enabled both higher sample rates and accuracy which, in turn, enable researchers to capture a more detailed and accurate representation of eye movements. The most frequently reported eye-tracking measures have at least traditionally been fixations and saccades. The term fixation denotes an event in which the eye, or gaze, stops moving for a certain period of time, whereas saccade refers to the rapid eye movements between fixations (Holmqvist et al. 2011). While both stimuli-driven and volitional saccades are often considered as exploration of the visual environment, fixations are hypothesized to be an indicator of attention to a certain position (Klein & Ettinger, 2019; Holmqvist et al. 2011). However, it is debatable to what degree our eye movements and fixations influence, and are influenced by, our cognitive system in various situations. Furthermore, it is somewhat unclear to what extent our eye movements are affected by training and if, and how, eye movement measures are related to cognitive processes, performance or learning outcomes (Mayer, 2010). While these methodological and theoretical issues call for caution in interpretation of eye movement data, they also constitute an exciting research agenda which, in turn, has resulted in a thriving field of research.

This study is designed as a meta-review, that is, a review of previous reviews or meta-analyses (Sarrami-Foroushani et al. 2015), following a semi-structured review approach to be able to synthesize the state of knowledge between disciplines and the variety of methodological contributions across different domains. Semi-structured approaches are suitable for mapping a field of research and synthesize the state of knowledge to set an agenda for future studies (Snyder, 2019). While the approach might take on systematic search strategies, studies often lack the scientific rigor of systematic or scoping reviews. What is considered important also for the semi-structured review method is a transparent and developed research strategy for the audience to determine the worth and value of the chosen topic, method used and findings from the study (Snyder, 2019). In accordance with best practices, the study design is guided by the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) protocol for systematic reviews (Moher et al. 2015).

METHOD

The search for literature was conducted in February 2021. Advanced search options for discriminating type of articles (reviews) and timespan (2010-2021) were used on three databases: Scopus, Web of Science and Science Direct. Using search words “eye tracking”, “simulator”, “training” and “assessment” when searching through article titles, abstracts and keywords, a total of 403 studies were identified (Figure 1).

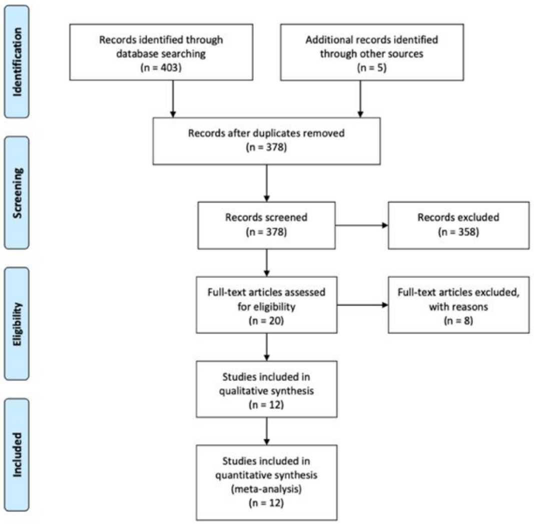

Figure 1: PRISMA 2009 Flow diagram representing the process of identifying, screening, and excluding/including relevant reviews in the meta-review.

Table 1. Studies included in the meta-review.

Authors	Domain	No of studies
Ashraf et al. (2018)	Medicine	33
Castillo-Segura et al. (2021)	Medicine	101
Fox & Faulkner-Jones (2017)	Medicine	61
Garden et al. (2015)	Medicine	8
Hermens et al. (2013)	Medicine	21
Levin et al. (2019)	Medicine	76
Limbu et al. (2018)	Cross-disciplinary	78
Merali et al. (2017)	Medicine	9
Robbins & Chapman (2019)	Transport	13
Rosch & Vogel-Walcutt (2013)	Cross-disciplinary	319
Tien et al. (2014)	Medicine	24
Ziv (2016)	Transport	50

After removing duplicates, 378 studies were screened. In this step, articles that didn't meet the inclusion criteria or lacked an explicit focus on eye-tracking support in simulator training and/or assessment were excluded. In all, 20 full-text articles were assessed for eligibility. After reading full-text, 12 studies remained and are included in the meta-review (Table 1).

RESULTS

The results of the meta-review show how eye-tracking is used for identifying differences between novices and experts and/or capturing expert gaze patterns. Gaze patterns can then be used for gaze training, formative assessment of non-technical skills as well as summative assessment of technical skills. Moreover, the results show how eye-tracking is appreciated as an unobtrusive technology that enables a quantitative assessment of objective parameters. The different uses as well as the barriers and benefits reported in the literature are described in the sections that follows.

Uses of Eye-Tracking for Training and Assessment

For training purposes, expert gaze patterns can be used for gaze training (Ashraf et al. 2018; Hermens et al. 2013). Gaze training in this context originates from laparoscopic surgery training where trainees are taught to adopt the gaze strategies of expert laparoscopic surgeons. While gaze training traditionally has been a matter of detailed, verbal instruction, recent eye-tracking studies has shown that gaze training by adding visual cues to a virtual environment, for example by using augmented reality, is a fruitful approach to teach novices to adopt expert gaze patterns (Tien et al. 2014; Limbu et al. 2015). Moreover, showing expert gaze patterns on a screen during a simulated task can be used for guiding novices to complete the tasks (Tien et al. 2014). Adopting expert gaze patterns have shown to lead to improved performance in terms of more efficient eye and hand movements as well as fewer errors in several medical tasks, including laparoscopy surgery and surgical knot tying (Ashraf et al. 2018; Tien et al. 2014). Hermens et al (2013) put forward the argument that gaze training may be especially effective since eye movements mainly rely on subconscious thought, which in turn are supposed to be less susceptible to the influences of stress. However, it is still largely unknown if the effect of gaze training is persistent over a prolonged period, or if it represents a permanent shift in gaze pattern (Fox & Faulkner-Jones, 2017). Expert gaze patterns can also be used for formative assessment of non-technical skills (Garden et al. 2015). By combining eye-tracking data with video records from the training situation, gaze behavior becomes visible for both instructors and trainees (Ashraf et al. 2018). Providing the means for showing where the participants guided their attention during training, through the combination of eye-tracking data and video recordings from a simulated scenario, opens for reflections on attention, situation awareness and decision-making (Rosch & Vogel-Walcutt, 2013). Hence, the combination of using eye-tracking data with videos from the simulation show potential as a fruitful approach to facilitate post-simulation debriefing.

Eye-tracking is a method for valid, reliable, and objective assessment of proficiency and therefore eye-tracking has gained most attention in summative assessment of technical skills (Rosch & Vogel-Walcutt, 2013; Fox & Faulkner-Jones, 2017; Tien et al. 2014; Levin et al. 2016). Traditionally, summative assessment of technical skills has been done by direct observation and feedback from an expert (Rosch & Vogel-Walcutt, 2013). Over the years, the search for objective measures have led to the development of a variety of rating scales and efforts to standardize the procedures that should be assessed (Rosch & Vogel-Walcutt, 2013). In the quest for reliable, quantitative assessment methods, eye-tracking has been used to establish objective metrics, e.g., path length, dwell time or number of movements as a metric to define surgical skill (Ashraf et al. 2018; Tien et al. 2014). Since experts and novices seem to demonstrate distinct differences in eye behavior during critical stages of task performance, eye-tracking is considered a useful assessment tool in medical areas such as surgery, pathology, and radiology, well suited for identifying skill level of trainees (Ashraf et al. 2018). However, in research on transportation, i.e., in driver behavior and aviation, the usefulness of eye-tracking for skill assessment is less clear (Robbins & Chapman, 2019; Ziv, 2016). In a review on eye-tracking in aviation, Ziv (2016) concluded that there was inconsistency in the differentiation between expert and novice pilots gaze patterns in the reviewed studies. While these inconsistencies could be attributed to methodological differences between studies, Ziv (2016) still conclude that expert pilots have more defined visual scan patterns with more frequent visits to instruments, and shorter dwell times on each instrument than novices. Similarly, gaze patterns between novice and expert drivers varied between included studies in Robbin and Chapmans' review (Robbins & Chapman, 2019). While it was difficult to differentiate between novices and experts in relation to fixation durations, vertical search and number of fixations, the most noticeable result was that novice drivers display a narrower horizontal search compared to experienced drivers (Robbins & Chapman, 2019). Hence, a conclusion that can be made is that albeit eye-tracking for training and assessment is a large and growing field of research, there is still need of studies that explore gaze patterns in settings characterized broad peripheral fields and the need to focus attention on moving targets.

Barriers and Benefits Reported in the Literature

While modern eye-tracking technologies are relatively inexpensive, user-friendly, unobtrusive and non-intrusive, there are still barriers for implementing eye-tracking in training and assessment reported in the literature. Several studies report costs as a barrier (Rosch & Vogel-Walcutt, 2013; Merali et al. 2014; Castillo-Segura et al. 2021), and several studies report challenges in taking them into use, including difficulties in selecting suitable tasks and/or uncertainty of how to measure specific skills (Fox & Faulkner-Jones, 2017; Merali et al. 2017; Hermens et al. 2013; Limbu et al. 2018). Barriers related to difficulty in programming the equipment are also reported (Rosch & Vogel-Walcutt, 2013). Moreover, studies report that eye-tracking is a rather time-consuming method, generating large data sets that need to be properly analyzed (Merali et al. 2017; Castillo-Segura et al. 2021). Hence, using eye-tracking technologies in training and assessment is still not an off-the-shelf pedagogical method, easily available for the everyday educational practices in complex socio-technical domains. However, there is still a number of benefits associated with using eye-tracking technology for training and assessment that encourage the field to continue studying its worth. As already mentioned, eye-tracking technologies can enable quantitative assessment of objective parameters (Ashraf et al. 2018; Tien et al. 2014; Hermens et al. 2013; Levin et al. 2019) and has shown valuable for training purposes (Merali et al. 2017; Tien et al 2014; Limbu et al. 2018; Robbins & Chapman, 2019). Modern eye-tracking devices are unobtrusive and enable users' natural movement, an important aspect when training or assessing manual skills (Castillo-Segura et al. 2021).

DISCUSSION

As outlined above, the current state of research clearly identifies the potential of using eye-tracking as a tool to support training and assessment across domains. However, there is the need to acknowledge some of the limitations that have been identified in the literature included in this review.

Due to the fact that eye-tracking studies are conducted in controlled environments where instrumentation may set limits to the design of tasks, internal and external validity have been reported to be rather low, such in the studies included in Garden et al. (2015). This is also in line with Robbins and Chapman (2016) who discuss that identified differences in visual search patterns between novices and experts may have been identified due to the fact that comparisons are partially made to extremely experienced drivers (e.g., policemen). This may limit the generalizability of the studies towards the differences between experienced and novice drivers. Further, as noted by Levin et al. (2019) and Robbins and Chapman (2016), there is a limited generalizability of the results across due to the sampling methods in many studies. Sample sizes are reported to be small and the recruitment of subjects, for example using students, may also impact on the potential to generalize from sample to the general public. Moreover, task design and especially the use of eye-tracking to measure cognitive load are mentioned as critical across the reviewed articles. To obtain information on cognitive load during tasks, the study design needs to be built on detailed task analyses and needs to take into concern that certain professions require spatial movements (Rosch & Vogel-Walcutt, 2013) or use peripheral vision (Merali et al. 2017), which both can be problematic to capture in eye-tracking studies. Further, due to the focus on the individual gaze patterns, collaborative tasks and team work remains unexplored (Ziv, 2016). Similarly, Hermens et al. (2013) notes that there is not yet a clear understanding of what characterizes best pattern in terms of skills, which then also may also be impacted by the choice of task under study. In addition, despite the increasing body of knowledge, there is still a limitation with eye-tracking being used to measure cognitive abilities according to Levin et al. (2019) as studies usually lack comparisons to other secondary measurements or psychophysiological responses approaches (Rosch & Vogel-Walcutt, 2013).

This study used a semi-structured review approach based on the PRISMA-protocol. While the initial set consisted of 378 studies, only 12 articles met the inclusion criteria. The included studies do not only show the wide variety of eye-tracking applications across domains, but also exemplify the heterogeneity with which reviews can be conducted. The time spam of included studies differed widely, as did the degree of structure with which the reviews were conducted. Furthermore, it is noted that the number of scientists and experts in this field of research is limited, which means that some research collaborations and clusters may have been overrepresented in the included review articles, in turn affected the obtained results from the systematic analysis.

ACKNOWLEDGMENT

This study is part of the project *Evaluation of eye-tracking as support in simulator training for maritime Pilots* financed by the Swedish Transport Administration.

REFERENCES

Ashraf, H., Sodergren, M. H., Merali, N., Mylonas, G., Singh, H., & Darzi, A. (2018). Eye-tracking technology in medical education: A systematic review. *Medical teacher*, 40(1), 62–69.

Castillo-Segura, P., Fernández-Panadero, C., Alario-Hoyos, C., Muñoz-Merino, P. J., & Kloos, C. D. (2021). Objective and automated assessment of surgical technical skills with IoT systems: A systematic literature review. *Artificial Intelligence in Medicine*, 112, 102007.

Fox, S. E., & Faulkner-Jones, B. E. (2017). Eye-tracking in the study of visual expertise: methodology and approaches in medicine. *Frontline Learning Research*, *5*(3), 29–40.

Garden, A. L., Le Fevre, D. M., Waddington, H. L., & Weller, J. M. (2015). Debriefing after simulation-based non-technical skill training in healthcare: a systematic review of effective practice. *Anaesthesia and Intensive Care*, 43(3), 300–308.

Hermens, F., Flin, R., & Ahmed, I. (2013). Eye movements in surgery: A literature review. *Journal of Eye movement research*, 6(4).

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Klein, C., & Ettinger, U. (Eds.). (2019). *Eye movement research: An introduction to its scientific foundations and applications*. Springer Nature.

Levin, M., McKechnie, T., Khalid, S., Grantcharov, T. P., & Goldenberg, M. (2019). Automated methods of technical skill assessment in surgery: a systematic review. *Journal of surgical education*, 76(6), 1629-1639.

Limbu, B. H., Jarodzka, H., Klemke, R., & Specht, M. (2018). Using sensors and augmented reality to train apprentices using recorded expert performance: A systematic literature review. *Educational Research Review*, 25, 1–22.

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and instruction*, 20(2), 167–171.

Merali, N., Veeramootoo, D., & Singh, S. (2017). Eye-tracking technology in surgical training. *Journal of Investigative Surgery*.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, 4(1), 1–9.

Robbins, C., & Chapman, P. (2019). How does drivers' visual search change as a function of experience? A systematic review and meta-analysis. *Accident Analysis & Prevention*, 132, 105266.

Rosch, J. L., & Vogel-Walcutt, J. J. (2013). A review of eye-tracking applications as tools for training. *Cognition, technology & work*, 15(3), 313–327.

Sarrami-Foroushani, P., Travaglia, J., Debono, D., Clay-Williams, R., & Braithwaite, J. (2015). Scoping meta-review: introducing a new methodology. Clinical and translational science, 8(1), 77–81.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of business research*, 104, 333–339.

Tien, T., Pucher, P. H., Sodergren, M. H., Sriskandarajah, K., Yang, G. Z., & Darzi, A. (2014). Eye tracking for skills assessment and training: a systematic review. *journal of surgical research*, *191*(1), 169–178.

Ziv, G. (2016). Gaze behavior and visual attention: A review of eye tracking studies in aviation. *The International Journal of Aviation Psychology*, 26(3-4), 75–104.