Eye Movements in VR Training: Expertise Measurement and it's Meaning for Adaptive Chess Training

Carolin Hainke(✉) and Thies Pfeiffer

Faculty of Technology, University of Applied Sciences Emden Leer, Constantiaplatz 4, 26723 Emden, Germany
{carolin.hainke, thies.pfeiffer}@hs-emden-leer.de

Abstract. Observing behavior can provide insights into mental states and cognitive processes. In the context of expertise measurement, research is taking advantage of analyzing gaze-based data to understand those underlying processes. The movements of the eyes represent search strategies in problem-solving scenarios, which can be used to distinguish a novice from a person that is more experienced with the topic and the solving process. Applications such as learning environments can be improved by taking expertise into account, as expertise-related instructions and design-decisions could then be adjusted to the learner's individual needs. In the following, we will discuss the meaning of expertise and instruction design for learning in more detail. Prior work has shown that different groups of expertise can be distinguished using eye movements. We verify this for virtual reality-based chess trainings by presenting a study on chess problem solving in VR. Based on these findings, we discuss suggestions for the implementation of adaptive learning applications on the example of chess.

Keywords: Adaptive training · Eye-tracking · Expertise

1 Introduction

Human expertise emerges from knowledge that is structured and stored in form of cognitive schemas in our long-term memory [15]. Learning is a process that aims to build those schemas. With increasing expertise, the schemas are improved and can be loaded into our working memory during the learning process to, additionally to the information presented by the learning material, allow the learner to rely on the prior gained knowledge. This reduces the amount of cognitive load that is produced by the learning material itself (intrinsic cognitive load) or the way the information to learn is presented to the learner (extraneous cognitive load) and leaves more capacities for the germane load [13]. The latter is said to represent the effort willingly put into learning and therefore is essential to building schemas [15].

However, novices do not have that structured form of knowledge and therefore have to work with the information given during learning. Without additional guidance, they have little chance to engage in strategic problem-solving, as experts would do, but rather have to follow a cognitively demanding trial-and-error approach [7]. But external

T. Z. Ahram and C. S. Falcão (Eds.): AHFE 2021, LNNS 275, pp. 123–130, 2021.
https://doi.org/10.1007/978-3-030-80091-8_15

guidance that can help novices finding relevant information and knowing how to use it can on the other side hinder the learning outcome of experts. For them, it adds unnecessary information that they additionally have to process. Further, if the guidance does not match their schema in the long-term memory, a process of cross-referencing is needed that requires additional working-memory resources. This phenomenon is called the “expertise reversal effect” [8].

Therefore, for example Kalyuga suggests adaptive learning environments that take the expertise of the learner into account [6]. During the learning with problem-solving tasks, elaborated knowledge and built schemas can be used for improving skills and performance. However, a certain level of expertise is required to truly benefit from this learning technique. Novices for example would have to deal with both basic knowledge about the problem and additionally the solving as it is. This can easily lead to a cognitive overload. Therefore, learning with worked examples is a more appropriate technique for learners new to a certain domain. By studying examples of prior solved problems, novices can gain a basic understanding of the topic and solving strategies. They can build schemas and increase their knowledge and expertise until they are ready to use their knowledge for the actual problem-solving.

Our goal is to follow the approach of expertise-related training adaption with building a training application that adapts the learning technique to the individual users’ needs by taking their level of expertise into account. With the use of eye tracking, the user’s expertise should be assessed online right during training. Hence the application should be able to adapt not only to the general expertise but to the one directly related to the current situation.

In the following, we give an overview of the positive effects expertise-adapted learning can have on the training and how expertise has been assessed (Sect. 2). In the presented experiment (Sect. 3), we want to verify those findings for expertise assessment in a VR chess application. For the context of chess, we hypothesize that experts focus more on the solving process, represented by more frequently gazing at free positions while looking for the correct target position. Contrary, novices probably will look around the board and the pieces, resulting in a rather generally distributed gaze behavior. We conclude with considerations on how our findings can be used in our future work to build an online-adapting chess training application.

2 Related Work

In order to analyze the impact of prior knowledge on learning, Clarke and collegues designed an experiment in which students were asked to learn mathematical content by using spreadsheet technology [2]. Students novel to spreadsheet learning seemed to be overwhelmed by being presented simultaneously the spreadsheet introduction and the mathematical learning material. When the material for the technology was presented prior to the actual learning content, the results of those students when testing the learned mathematical skills were significantly better.

Kalyuga and Sweller analyzed the effect of learner-adapted instructional design on learning efficiency [7]. They define efficiency (E) as the impact of task performance (P) on subjective mental effort rating (R) as follows: E = P / R. In an adapted training session, learners were based on their prior assessed expertise allocated to three different instructional stages: worked examples, faded worked examples and conventional problems. A control group did not learn with adapted instructions. Results show that learner-adapted training leads to higher efficiency gains meaning that in this condition learners were able to use their cognitive capacities more efficiently.

In contrast to the pre-tests done in the mentioned studies to assess the user's expertise, Lai and collegues compared the problem-solving behavior of individual users against those of experts [9]. They used the information to build a system that recommends learning activities intending to maximize a user's expertise in the specific domain. The feedback to this system was constantly positive and all questioned users would use such a system to build their expertise.

Those results show that expertise-adapted learning is very promising. But the assessment of expertise has been done in pre-tests or by an effortful data-driven comparison of user behavior. Reingold and Charness were able to show that, at least in the context of chess, experts can also be distinguished from novices by their gaze behavior [11]. They could show that experts do have a greater visual span when perceiving the scene. Further, experts process the chessboard in form of information chunks rather than individual pieces, shown by fewer fixations and more refixations on related pieces. This results in experts being faster done with scene perception and starting earlier with the actual problem-solving, compared to novices.

Those findings suggest that an online-assessment of expertise using eye-tracking data could be feasible. Additionally, other researchers did take eye-tracking for adaptive learning into account. For example, AdeLe [10] detects skipped content and reacts on the wrong visual focus of the learner. Similarly, e5Learning [1] dynamically displays additional content depending on the main content being assessed.

To tie in with the literature, we want to bring those findings in the context of online expertise assessment using eye-tracking. Our first step is the study presented in the following, using eye-tracking data to distinguish chess experts from novices. A technique that could later be used for an online assessment of expertise level and related training adaption.

3 Method

The study should be conducted in a problem-solving environment in which the level of difficulty is easy to modify and the distinct states are easy to evaluate. Therefore, we decided to build a VR chess application, including a training mode and a study mode with chess tasks of different levels of difficulty (see Fig. 1).

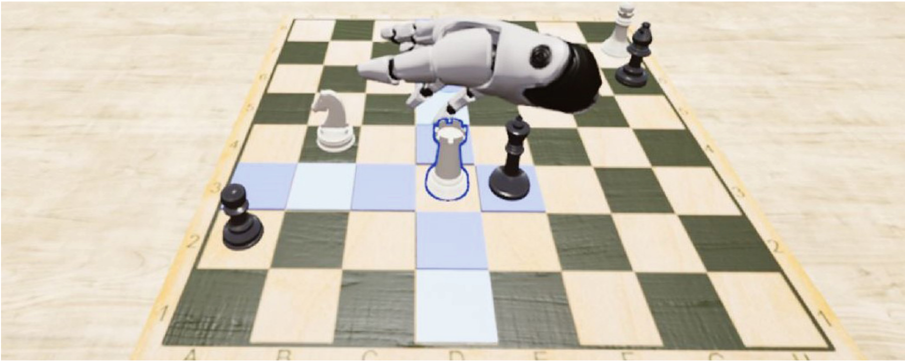

Fig. 1. Screenshot from the tutorial part, where pieces can be moved freely, with highlighted movement patterns

Chess is an example of a complex visual problem-solving task. Players have to examine the overall situation, followed by their choice of the piece to move as well as its destination position. Those situations can be very diverse. The 64 positions and six different types of chess pieces per player offer the opportunity to create chess scenarios of many different levels of difficulty and ambiguity. Those scenarios can easily be evaluated since chess has clearly defined rules that result in discrete states for every piece and position. Further, chess is a valid environment in which the expertise of players can be distinguished by analyzing their eye movement behavior [11].

All participants had to solve the same tasks in different order to enable a within-subject comparison of the resulting data. Tasks have been created by a chess trainer in the chess club Hückeraschen [4]. He created 6 tasks, 18 in total, for each of three levels of difficulty: easy, medium and hard.

Overall, 34 participants took part in this study. Of these, 14 participants are members of the chess club, ranked with a DWZ [3], a German ranking scheme comparable to the international Elo. This group is referred to as experts in the following. The group of novices consisted of 19 participants who had less or no chess experience. The study lasted 45 min at max. Participants have been compensated with 8€ each. An SMI-HTC Vive system was used for VR interaction and eye-tracking device.

3.1 Study Design

Before the experiment started, the participants were given the instructions in the form of written text. They were told about the tasks and the handling of the application. In particular, they were asked to immediately draw after they decided on a solution. Lastly, they signed an informed consent about the procedure, the data use and their option to stop the experiment at any time. If there have been no further questions, the participants continued with filling out a questionnaire about personal data such as their dominant eye and their chess experience. Following, the participants were seated on a chair and the calibration of the application started. The position of the chessboard was adjusted to the position of the current participant as well as the individual height measured by the headset. Afterward, the eye tracker was calibrated using the five-point-calibration procedure, provided by SMI [12].

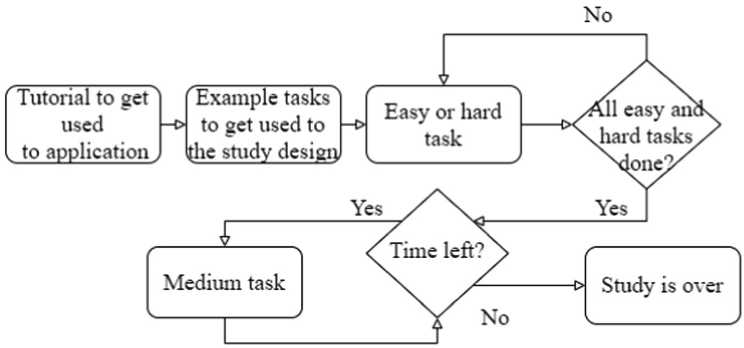

Fig. 2. Overview of the study procedure.

After the calibration processes, the actual experiment started. An overview of the study process is given in Fig. 2. Every participant started with a tutorial which consists of two phases. First, one chess piece of every kind was placed on the board to enable the participants to freely move the pieces, learn their movement patterns and get familiar with the handling of the application. All positions that can be reached by the selected piece were highlighted during the tutorial. This phase was ended by the experimenter as soon as the participant stated to feel confident in the use of the application. The second phase contained two exemplary tasks to give the participant the chance to get used to the task design. The difficulty was below the easiest tasks that were designed for the study. Similarly to the later process, the tasks were finished after a piece was moved. After the second exemplary task, the experimenter asked the participant if there were any further questions. If not, the actual study started. The tutorial phase was designed that way to reduce the extraneous cognitive load [13] that might be generated by the use of an unknown application, in particular given the virtual reality experience, that might be new for many participants.

By finishing the tutorial phase, the actual study started. Participants had to solve all easy and hard tasks in random order. For every task there was a time limit of three minutes. After one, two and two minutes and 30 s a sound was played to let the participant know how much time was left to solve the task. If the time was over, the next task started automatically. To not overwhelm the participants, the study finished in all cases after 36 min. This time was reached if the maximum available time was needed for every of the 12 easy and hard tasks. In case participants were able to finish these tasks earlier, they were prompted with tasks of medium difficulty until the time limit was hit. This study design was chosen to make sure to get data that probably shows greater differences in the viewing behavior between experts and novices. The additional medium tasks have been included to also obtain data that might be less clearly differentiable to make the evaluation more robust. When all tasks were solved or the time was over, the study was finished.

3.2 Results

To test the results for significance, the "Welch Two Sample t-test" has been used since it is more robust to unequal variances and sample sizes.

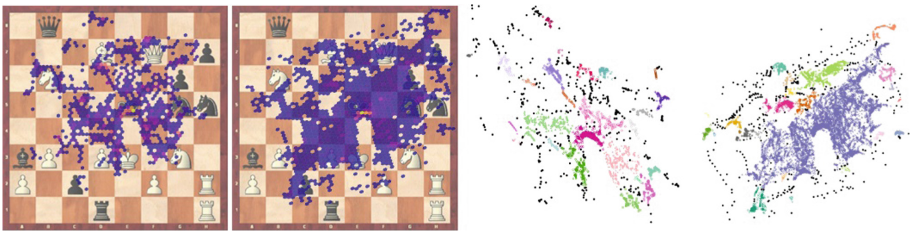

Fig. 3. Example of the gaze distribution for experts (1) and novices (2) and the resulting clustering for experts (3) and novices (4)

Figure 3 exemplarily shows the gaze distribution for one expert and one novice. While the gaze of novices is more equally distributed over a large area on the chessboard, experts particularly investigate smaller areas that are more likely related to the solution of the task. Although the size of the area investigated by experts and novices can be visually distinguished, no significant difference has been found comparing the areas covered ($p = 0.41$). Therefore, the gaze distribution has been further analyzed using the clustering algorithm DBSCAN [5]. The parameters used to cluster the gaze positions were 1 for the epsilon neighborhood and 20 as the minimal amount of points that should lay within one epsilon region.

It is known that experts do have a more strategic solving-behavior [7] and, especially in chess, they are faster done with scene perception and start with the solving process earlier [11]. Further, the longer the task takes, the more gaze positions are logged what might automatically lead to more clusters.

To prevent the results from potentially being influenced by that and to see, if experts and novices can already be distinguished in an early state of the task, only data of the first five seconds for every task have been analyzed. The gaze positions of novices can be clustered in fewer groups (Mean = 11.16, SD = 5.91) than the positions of experts (Mean = 16.06, SD = 7.14). This difference is statistically significant ($p < 0.01$).

Further, the amount of pieces and positions the participants gazed at during their decision process has been evaluated. Experts gazed on average 55% of the task time at free positions and only 45% of the time at pieces (SD = 12.4). Whereas the gaze of novices landed on free positions only 47% of the time (SD = 15.2). Hence, experts gazed significantly ($p < 0.01$) more at free positions, while novices spent significantly ($p < 0.01$) more time gazing at pieces. Furthermore, on average experts fixated their later chosen position after 11.6 s (SD = 16.6) the first time, while the first fixation of novices on their later chosen position occurred after 27.12 s (SD = 29.5) on average. This difference is statistically significant ($p < 0.01$).

3.3 Discussion

The results show that significant differences in gaze behavior could be found that can be attributed to expertise. Since experts are already familiar with solving chess-related tasks, they know where to look for relevant information that leads to the correct solution. This results in gazing at fewer, but more distinct areas on the chessboard. In contrast to that, novices investigate a broader area to get more information about the whole situation.

This can also be a reason why experts spend more time gazing at free positions compared to novices. While novices need more time to get an overview of the situation and gather information about their available pieces and options, experts focus on finding the solution by primarily examining free positions. They also might be able to identify pieces in their peripheral vision, what could be made possible by their greater visual span, as mentioned by Reingold and Charness [11]. Those results probably lead to the last presented finding. Because experts know where to find relevant information and focus on possible solutions, they fixate their later chosen position earlier and take it into account for a possible solution. However, novices process more information spread over the chessboard and therefore need more time to find and fixate the position that might solve the task.

4 Conclusion and Future Work

With the presented results we are able to show an overview of parameters to consider for expertise measurement in the context of chess. We could show that it is possible to distinguish experts from novices by their gaze behavior within the first five seconds after the start of solving a chess task. This expertise assessment does not only represent the overall chess expertise, but rather the expertise related to the very specific task.

In our future work, we want to use these results for online expertise assessment to build an adaptive chess training application. Based on our findings from the literature an adaptive training should consider worked examples to train novices [6]. In the context of chess, this could mean to show a possible solution for a task with additional explanations. More specifically, the target piece and its movement pattern as well as the target position can be highlighted to let novices know where to find information relevant for the solution. A faded example [7] could be to highlight only the piece or the target position or show the movement pattern of fixated pieces – similar to the approaches of showing content based on the gaze behavior [1, 10]. Experts would get no additional guidance for the actual problem-solving. Further, to enable sequential learning for novices, they should get the opportunity to learn the movements of the pieces before starting to solve the chess tasks. This can be achieved by our provided tutorial. To test this approach, we will conduct a study with chess novices to see if this kind of adapted guidance can improve novices' chess training.

References

1. Calvi, C., Porta, M., Sacchi, D.: e5Learning, an E-Learning environment based on eye tracking. In: Díaz, P. (ed.) Proceedings / the 8th IEEE International Conference on Advanced Learning Technologies. Santander, Spain, 1–5 July 2008, pp. 376–380. IEEE Computer Soc, Los Alamitos (2008)
2. Clarke, T., Ayres, P., Sweller, J.: The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications. Educ. Tech. Res. Dev. **53**, 15–24 (2005)
3. DSB-Wertungsordnung - Deutscher Schachbund - Schach in Deutschland (2021). https://www.schachbund.de/wertungsordnung.html. Accessed 08 Feb 2021
4. Ellenberg, G.: Schachverein Hückeraschen (2016). http://www.xn–hcker-aschen-dlb.de/html/schach.htm. Accessed 08 Feb 2021
5. Ester, M., Kriegel, H.-P., Sander, J., Xiaowei, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd **96**(34), 226–231 (1996)
6. Kalyuga, S.: Knowledge elaboration: a cognitive load perspective. Learn. Instr. **19**, 402–410 (2009)
7. Kalyuga, S., Sweller, J.: Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. Educ. Tech. Res. Dev. **53**(3), 83–93 (2005)
8. Kalyuga, S., Ayres, P., Chandler, P., Sweller, J.: The expertise reversal effect. Educ. Psychol. **38**, 23–31 (2003)
9. Lai, J., et al.: Expediting expertise: supporting informal social learning in the enterprise. In: Kuflik T (ed) Compilation publication of IUI'14 proceedings & IUI'14 companion. pp 133–142 Haifa, Israel. ACM, New York (2014)
10. Pivec, M., Trummer, C., Pripfl, J.: Eye-tracking adaptable e-learning and content authoring support. Informatica **30**, 83–86 (2006)
11. Reingold, E.M., Charness, N.: Perception in chess: evidence from eye movements. In: Underwood, G. (ed.) Cognitive Processes in Eye Guidance, pp. 325–354, Repr. Oxford Univ. Press, Oxford (2009)
12. SMI Eye Tracking HMD HTC Manual.pdf on Google Docs (2021). https://drive.google.com/file/d/0B2BduJfBcR6RVEo5UmkxTzM5WnM/view. Accessed 08 Feb 2021
13. Sweller, J.: Cognitive Load Theory. In: Mestre, J. (ed.) Cognition in Education, pp. 37–76, vol. 55. Elsevier, Amsterdam (2011)
14. van Gog, T., Ericsson, K.A., Rikers, R.M.J.P., Paas, F.: Instructional design for advanced learners: establishing connections between the theoretical frameworks of cognitive load and deliberate practice. ETR&D **53**, 73–81 (2005)
15. van Merrienboer, J.J.G., Ayres, P.: Research on cognitive load theory and its design implications for e-learning. Educ. Tech. Res. Dev. **53**(3), 5–13 (2005)