Eye-tracking in archaeological practice: applications, potential, and challenges

Abstract: Eye-tracking is a powerful tool that enables the understanding of visual attention through analysis of gaze patterns. Although sight is one of the senses crucial for learning about the past, the eye-tracking technology is still rarely used in archaeology. This paper discusses current applications, potential, and challenges of eye-tracking in archaeological practice in order to foster a broader use of this method. In particular, the article investigates the influence of subjective, technical, and methodological factors on the feasibility of eye-tracking research in archaeology based on experience gained during the project "Psychological Aspects of Creating and Acquiring Knowledge about the Past". Observations and analyses conducted as part of this project have shown that while technical limitations pose minimal obstacles to research, methodological issues may present a more significant challenge. The discussion of various aspects of eye-tracking in archaeology is intended to aid archaeologists considering the integration of this technology into their research.

Keywords: eye-tracking, archaeology, applications, feasibility, methodology, Sudan, Poland

INTRODUCTION

The central role of "the act of discovery" in archaeology prompts questions about its underlying processes. How do archaeological discoveries happen? Why are some archaeological objects readily discernible, while

Tomasz Michalik

University of Warsaw, Polish Centre of Mediterranean Archaeology

Polish Archaeology in the Mediterranean 33 | Michalik 2024: 433–458
https://doi.org/10.37343/uw.2083-537X.pam33.07
received 25 March 2024 | received in revised form 3 June 2024 | accepted 22 June 2024 | available online 14 November 2024

Acknowledgments
The author wishes to thank all the participants and institutions involved in the realization of the project “Psychological Aspects of Creating and Acquiring Knowledge about the Past”. Special thanks are due to: Prof. Agnieszka Szarkowska for providing the ATV Laboratory for the eye-tracking study on visual analysis of magnetic data; Prof. Łukasz Gaweł and Prof. Aleksandra Bełczowska-Sulikowska for enabling research in the Faras Gallery, National Museum in Warsaw; Prof. Artur Obłuski for the opportunity to conduct research in Old Dongola, Sudan; Dr. Mariusz Drzewiecki and Robert Ryndziewicz for providing magnetic data from Soba, Sudan; the authorities of the Faculty of Archaeology, University of Warsaw, as well as Dr. Dobrochna Zielińska and Agnieszka Ryś-Jarmużek for the opportunity to conduct eye-tracking educational research with students of Archaeology.
This work was funded within the framework of the Excellence Initiative — Research University Programme Action I.3.10. “Archaeooriental Studies” implemented by the University of Warsaw. The research in Old Dongola was co-funded by the grant “Kyriakós oíkos? Diachronic conceptualisation of the space and function of Medieval Nubian churches” from the National Science Centre, Poland (grant No. 2021/42/E/HS3/00381).

others remain more difficult to identify, and some evade attention for extended periods? Also, how do socially and culturally disparate audiences perceive archaeological remains?

A possible approach to answering these questions is through the analysis of visual attention. Eye-tracking —recording eye movements and fixations— can be a powerful tool for achieving this objective. The advantages of eye-tracking have been recognized in numerous fields. For example, in medicine, it allows for the early recognition of social disorders in children who may not yet communicate verbally, facilitating prompt treatment (Jones and Klin 2013). In forensic science, it aids understanding factors that can influence judgments (Watalingam et al. 2017). In aviation, it helps prevent pilot mistakes (Peysakhovich et al. 2020; for other applications see Stuart 2022). However, the potential of eye-tracking as evidenced in other disciplines is not readily transferable to archaeological research, the pivotal concern being the vulnerability of eye-tracking to a multitude of contextual variables. As emphasized in visual-cognitive literature, technical parameters of the devices, the nature of the tasks, as well as the conditions under which the research is conducted all impact the outcome of an eye-tracking study (e.g. Hüttermann, Noël, and Memmert 2018; Vehlen et al. 2021; Ulas, Emirzeoglu, and Burgert 2022).

The objective of this paper is to provide a framework for evaluating the feasibility of eye-tracking in archaeological practice, with a view to informing future research. The presentation of the potential and challenges of eye-tracking is based on the experience gained in the "Psychological Aspects of Creating and Acquiring Knowledge about the Past" project conducted in Poland and Sudan between 2021 and 2023. The project afforded an opportunity to test eye-trackers in a variety of settings, including fieldwork, visits to original monuments on archaeological sites and in museums, as well as desk-based work with artifacts and computer-aided analysis of archaeological data. Based on these experiences coupled with a review of relevant literature, this article discusses three issues: subjective feasibility, technical feasibility, and the internal and external validity of eye-tracking studies in archaeological practice. These considerations are preceded by a brief overview of applications of eye-tracking research in archaeology, offering insights into how this method may enhance the study of the past.

EYE-TRACKING IN ARCHAEOLOGY

The history of eye-tracking research in archaeology spans the last 20 years. To the best of the author's knowledge, the use of eye-tracking in archaeology goes back to 2004, when Yuko Tokitsu (2004) published a study on visual exploration of pottery, concluding that experts in the field of archaeology were more perceptive of the morphological features and proportions of pottery compared to non-experts. In the following years, studies on the role of expertise in visual exploration of archaeological data also utilized aerial photographs and three-dimensional excavation models. Seeking to determine how archaeological training influences detection of archaeological features in aerial photographs, Michalik (2014) found that students who had completed a course in Aerial Archaeology were more cued by geometric shapes in aerial photographs compared to untrained participants. In another study, which investigated the impact of prior knowledge on the visual processing of videos of three-dimensional models of excavations, Forte and colleagues (2022) proved that experienced archaeologists, in comparison to non-archaeologists, tended to spend more fixation time and engaged more with foreground archaeological features and artifacts.

Understanding the analytical strategies employed by experts proves useful in the teaching of archaeology. Educational materials are also crucial to this process; by investigating the visual engagement of students with learning aids, Ramkumar and colleagues (2019) demonstrated that eye-tracking can facilitate the assessment of educational solutions by teachers of archaeology. The subjects in the study engaged visually in a comparable manner when interacting with both virtual and tangible archaeological artifacts. Thus, the study suggested that virtual artifacts can serve as promising educational tools when real artifacts are inaccessible.

In the late 2010s, eye-tracking began to be used in archaeology to unravel past cognition. In a study on the role of dark zones of caves in rituals, Tabatabaeian (2018) demonstrated that the subjects' gaze patterns in darkness resembled those observed during creative thinking, suggesting that darkness may stimulate creativity, which, in turn, plays a significant role in ritual behaviors. In another study, Hirst (2019) analyzed gaze patterns when looking at Upper Paleolithic Venus figurines to empirically test theories regarding their function and develop a gaze-based classification of the figurines. This study showed that looking at artifacts through the eyes of others can significantly contribute to reconsiderations of archaeological interpretations.

An interesting application of eye-tracking was also demonstrated by Criado-Boado and colleagues (2019; 2023), who found that pottery made by prehistoric complex societies stimulated in modern viewers more vertical eye movements than pottery produced by egalitarian societies. The above research opened a new chapter in the study of relationships between social complexity, its material expressions, and their influence on the cognitive system.

In a series of experiments devoted to affordances of Lower Palaeolithic flint tools, Silva-Gago and colleagues (2021; 2022a; 2022b; 2022c; 2022d) present eye-tracking as a promising methodology for research on the evolution of brain–body–tool interactions (see also Silva-Gago and Bruner 2023). The studies revealed, among others, that, when looking at Lower Palaeolithic stone tools, the subjects tended to visually focus on aspects related to tool manipulation regardless of prior training. In contrast, archaeologists experienced with this type of artifacts tended to focus on the tools' functional parts (Silva-Gago et al. 2022b). These results lead to an intriguing question: Are grasping cues evolutionarily more deeply ingrained in our relationship with stone tools than functional ones?

So far, Palaeolithic artifacts have been the most commonly studied materials in archaeological eye-tracking research. However, eye-tracking can be just as effective for studying more recent historical topics, such as life in ancient Roman houses. In a project integrating GIS and eye-tracking data during virtual tours of Pompeian houses, Campanaro and Landeschi (2022) demonstrated that immersive virtual reality (VR) combined with eye-tracking can provide deeper insights into the experience of ancient houses, for example in various lighting conditions. This versatile approach holds promise with regard to its applicability to other types of monuments and periods, opening up new opportunities for archaeological research.

Without a doubt, in recent years, eye-tracking has been recognized as a useful tool for identifying factors influencing archaeological interpretations as well as enhancing our knowledge about the past. At the same time, despite the growing number of eye-tracking studies in archaeology, the feasibility aspects of experiments have not been considered thoroughly enough. The discussion in this paper is grounded in the results of eye-tracking research conducted in various conditions and focuses on the opportunities and challenges posed by eye-tracking in the field of archaeology with the aim to facilitate future research.

MATERIALS AND METHODS

Four eye-tracking experiments were conducted within the project "Psychological Aspects of Creating and Acquiring Knowledge about the Past".[1]

Experiment No. 1 sought to assess the impact of subject-specific knowledge on the visual analysis of magnetic data. The study was conducted between August 2021 and March 2022 at the ATV Laboratory, University of Warsaw, using the SMI RED250 eye-tracker. Three groups of participants took part in the study: non-experts (n=21), archaeologists who were not experts in magnetic data analysis (n=22), and experts in magnetic data analysis (n=5). The task of the participants was to silently analyze and later verbally describe the anomalies that might indicate archaeological features in ten magnetic images from the site of Soba in Sudan (Drzewiecki et al. 2021; 2022). During the silent analysis and verbal descriptions, the gaze patterns of the participants were recorded. Besides assessing the influence of subject-specific knowledge on visual exploration of the magnetic images, another objective of the study was to discern the types of anomalies that were noticeable even without prior archaeological knowledge from the ones whose detection and recognition required expert training [*Fig. 1*].

Experiment No. 2 involved research into the applicability of Eye Movement Modeling Examples (EMME) in the teaching of archaeology. EMME is

1 All studies were approved by the Committee for the Ethics of Research Involving Human Participants, University of Warsaw (Nos 97/2021; 112/2021; 150/2022; 203/2023). Prior to the study, all participants granted their written consent to participate.

an educational method that uses recordings of eye movements and fixations of experts performing analyses to guide the students' attention (for an overview, see Tunga and Cagiltay 2023). Within the study, instructional videos were prepared with the use of Tobii Pro Glasses 3, showing experts (n=13) in the process of analyzing artifacts, medieval paintings, and excavations. In order to create a reference group, five non-experts were also invited to the study. The task of both experts and non-experts was first to analyze artifacts and excavations in silence (natural conditions), and then to verbally describe their thoughts during the analysis (didactic conditions). Selected educational eye-tracking materials (concerning analysis of artifacts and paintings) were presented during classes held at the Faculty of Archaeology, University of Warsaw, in the summer semester of 2022 and evaluated by the students (n=27) with regard to their usefulness in understanding the experts' thinking (for results, see Michalik 2024).

Experiment No. 3 used eye-tracking to measure the development of iconographical analytical skills as a result of gaining knowledge in the field of archaeology. Twenty-eight students from the Faculty of Archaeology, University of Warsaw, were invited to participate in the experiment. The study, conducted with the use of Tobii Pro Glasses 3 eye-tracker, took place at the Faras Gallery of the National Museum in Warsaw and was organized in two sessions, one at the beginning and one at the end of the 2023 summer semester. The students' task was to look at Nubian wall paintings and offer an iconographical description and interpretation

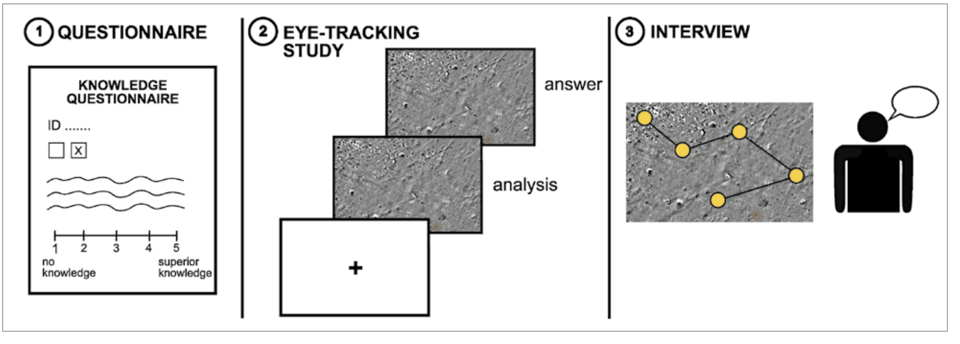
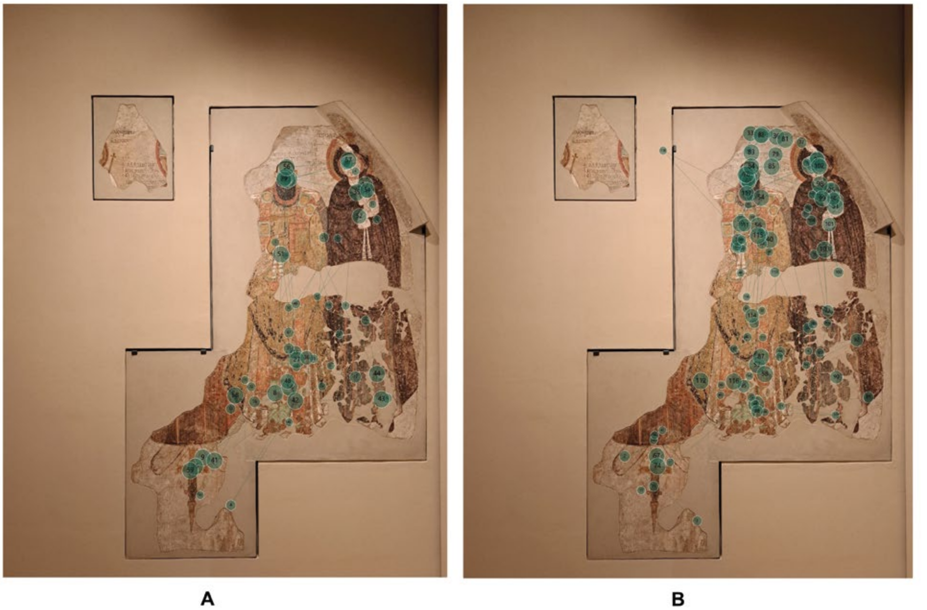


Fig. 1. Scheme of Experiment 1. (1) During the first step, participants were asked to fill out a questionnaire containing questions regarding socio-demographic data, as well as their knowledge and experience in archaeology and magnetic data analysis. Based on that, three groups of participants were distinguished: non-experts, archaeologists who were not experts in magnetic imaging, and experts in the magnetic method. (2) In the second step, participants were asked to analyze magnetic images from Soba, Sudan. Before viewing each image, the participants saw a fixation point, after which a randomly assigned image was displayed. When they were ready to describe anomalies that might indicate archaeological features, they were asked to press the spacebar. At this point, the same image was displayed on the monitor, and the description phase began. (3) After conducting the eye-tracking research, interviews were held to discuss the participants' experiences during the experiment, their strategies for analyzing the images, and the challenges they encountered (Processing T. Michalik)

of each painting. The aim of the study was to measure their gaze patterns and assess verbal descriptions of the paintings at the beginning and at the end of the semester, in order to evaluate the effects of archaeological teaching on iconographic interpretation [*Fig. 2*].

Lastly, Experiment No. 4 was designed to assess the influence of cultural factors on the perception of medieval Nubian wall paintings. The study was conducted at the Monastery on Kom H at the archaeological site of Old Dongola, Sudan, with Sudanese (n=48) and Western (n=19) visitors to the site. During the study, the participants were asked to view 17 wall paintings in a free-viewing paradigm similar to a museum experience. The gaze patterns of the subjects were recorded during their visit. After collecting the eye-tracking data, interviews focusing on the participants' understanding of each painting were conducted. The aim of the study was to identify similarities and differences in the visual engagement and understanding of paintings among subjects with various cultural backgrounds. The study was conducted between November and December 2022, with the use of the Tobii Pro Glasses 3 mobile eye-tracker [*Fig. 3*].



A B

Fig. 2. An example of an Archaeology student's scan path during silent observation (A) and during iconographical description (B) of the painting "Christ, Virgin Mary and Child, Bishop Marianos" at the beginning of the summer semester in 2023. The task given to the student was as follows: "Please look at the painting in silence and, when ready, describe it according to your knowledge." The data collected during the project provided insights into the order of visual analysis of paintings. The example in the figure shows that inscriptions were processed only during the stage of verbal description of the painting, and not during the initial stage of analysis. This opens a question about how students prioritize different aspects of a painting when analyzing it (Processing T. Michalik)

All four experiments provided valuable insights into the opportunities and challenges in using eye-tracking in archaeological practice. The key factors influencing the feasibility of the research, as well as summary information about the experiments are provided in [*Table 1*]. Maintaining a stable position was a common challenge with screen-based eye-tracking, as detecting eye movements requires the subjects to remain at a fixed distance from the screen. Whenever wearable eye-trackers were used, slippage proved to be a significant challenge, particularly in experiments that involved speaking, which affected the eye-tracker's position on the face. The third factor, lighting conditions, disrupted the measurements by interfering with infrared light, which is emitted by the eye-trackers and is crucial for detecting eye positions. Additionally, variation in pupil diameters affected the precise detection of the pupil/iris boundary, particularly in changing lighting conditions. This was a challenge in the monastery (Experiment No. 4), where disparities in room lighting were noted.

Table 1. Summary of information about Experiments Nos 1–4 and factors affecting research feasibility

Experiment No.	Device and software	No. of participants	Research setting	Challenges
1	SMI RED250/ BeGaze 3.0	48	Laboratory	Maintaining a stable position
2	Tobii Pro Glasses 3 (100 Hz)/ Tobii Pro Lab 1.194	18	Outdoors (excavations), museum, university offices	Natural light, slippage
3	Tobii Pro Glasses 3 (100 Hz)/Tobii Pro Lab 1.194	28	Museum	Slippage
4	Tobii Pro Glasses 3 (100 Hz) Tobii Pro Lab 1.194	67	Original monument	Variable lighting

RESULTS AND DISCUSSION

To explore the potential of eye-tracking in archaeology, challenges listed in [*Table 1*] were analyzed in the context of subjective feasibility, technical feasibility, and internal and external validity.

The impact of subjective factors on the success of eye-tracking research was particularly notable in Experiment 1, which called for a focused qualitative analysis of the participants' feedback. This experiment was the most challenging in terms of subjective comfort because the participants were required to maintain a stable position. Feedback on the comfort levels during the research tasks in Experiment 1, along with the notes concerning the participants' perception of eye-tracking devices from other experiments, and the number of participants who completed all experiments (task completion rate), served as a basis for identifying the subjective factors affecting the willingness to engage in the study. The identified factors were discussed in the first section in relation to the data collected in the project.

In the second section, data related to technical feasibility (i.e., traceability) were analyzed in the broader context of outcomes from other experiments reported in the literature. Supplementary data were provided by case studies conducted during the preparatory phase of the experiments to delve into other factors that could impact data quality (e.g. accuracy and precision). The aim was to determine whether in certain conditions obtaining high-quality data is more challenging than in others, since the project research was conducted in different environments and followed diverse research schemes [see *Table 1*, the columns "Research setting" and "Challenges"].

In the third section, a short discussion of socio-psychological factors influencing visual processing was offered. The aim of this part was to give a general overview of internal and external factors influencing quality of the research. Examples from the literature were limited to archaeological applications of eye-tracking research [*Fig. 4*].

SUBJECTIVE FEASIBILITY

Subjective factors are pivotal to the success of the eye-tracking research. Data quality is strongly affected by the participants' willingness to take part in the study and their engagement in performing the research task. Many factors may influence the subjects' decisions regarding participation in the study (for an overview, see Sheridan et al. 2020). This section focuses on the level of understanding the research and the comfort level during the research procedure.

Fig. 3. Eye-tracking research on the visual processing of medieval Nubian paintings in Old Dongola, Sudan. During the study, Tobii Pro Glasses 3 were used (Photo A. Chlebowski)

From notes and interviews conducted during the experiments, it is clear that for the majority of the subjects this was their first contact with eye-tracking. The lack of general knowledge about this method is a critical aspect to consider in eye-tracking research. As reported by Chu and colleagues (2015), the better the understanding of the research, the more positive the attitude toward participation. Most participants were curious about eye-tracking, yet some expressed concerns, as they imagined the eye-tracker to resemble a refractometer (a device used in ophthalmology to examine visual impairments) or goggles used in VR technology. Associations of this kind can lead to questions concerning the collection of health data during the study, as well as about limiting the freedom of vision and movement. Hence, information about the research device, the type of data collected, and the extent of data anonymization is fundamental for ensuring participation. The project experience also showed that a general presentation about the experiment supplemented with written information was a particularly effective way to encourage subjects to participate in the research.

None of the subjects reported a need to stop the study due to inconveniences associated with the experiment procedure or the testing equipment, which indicates a considerable potential for eye-tracking in archaeology. This is further supported by a qualitative analysis of responses in interviews conducted after Experiment 1. Most subjects (90%) indicated that the instructions were understandable, that they had no difficulty engaging in the study, and the research device was not a significant obstacle:

> "[Understanding the task] was not a problem. It seems to me that not only for me, but also for other participants. There shouldn't be any serious challenges here." (Subject A)
>
> "The instructions were clear, it was an interesting experience in general." (Subject B)
>
> "The [eye-tracker] program itself and the whole study was very fine, it wasn't cumbersome at all — neither staying in this [fixed] position nor calibration." (Subject C)

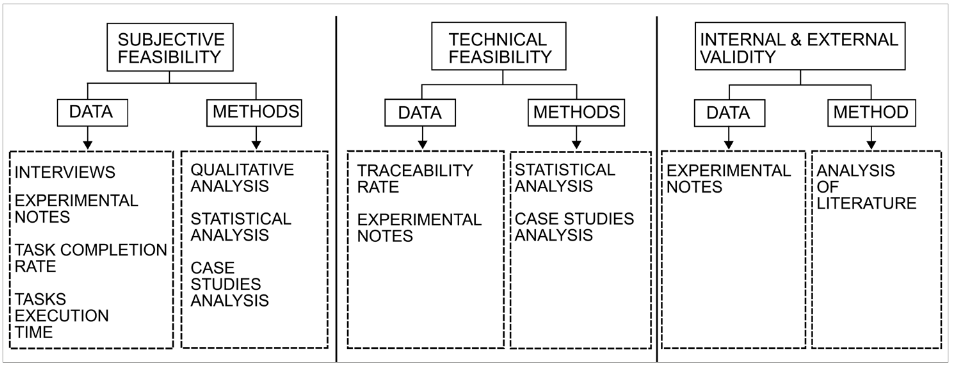


Fig. 4. Data and methods used to assess various aspects of the feasibility of eye-tracking research in archaeology based on the project "Psychological Aspects of Creating and Acquiring Knowledge about the Past" (Processing T. Michalik)

When discussing their feelings about the experiment, the subjects most often (85%) focused on the correctness of the analyses they had performed and expressed a desire to learn more about the images, indicating a focus on performing the research task rather than on the eye-tracker:

> "The task itself was easy. I watched and described what I saw, but in terms of archaeological interpretation, since I don't deal with geophysics at all, I found it difficult." (Subject D)

An issue reported during Experiment 1 was visual fatigue. In the study, which centered on magnetic imaging analysis, three subjects (out of 48) indicated that the task was too long (even though they were free to decide how long they wanted to engage in it). This outcome revealed a challenge related to the design of eye-tracking experiments, consisting in finding a balance between reliably studying a given aspect of archaeological knowledge and maintaining the subjects' ability to focus on the task. There are no general standards regarding the duration of eye-tracking experiments; as indicated by Sharafi and colleagues (2020) in the context of software engineering studies, the duration of the study should not exceed 90 minutes, with breaks every 30 minutes to counteract fatigue. In turn, Holmqvist and colleagues (2023) showed that 31% of the analyzed eye-tracking papers reported recording times ranging between 16 and 30 minutes, 28% between 31 and 45 minutes, and 20% between 46 and 60 minutes. Given the lack of uniform guidelines regarding experiment duration, conducting pilot studies might be particularly useful in planning archaeological eye-tracking research.

In the studies conducted by the author, three variables had a particular impact on the duration of the experiments. The first one was the level of expertise of the subjects. In the study related to the analysis of magnetic images (Experiment 1), the researched groups spent a similar amount of time on the silent analysis of ten magnetic images: between 5 minutes and 23 seconds and 6 minutes and 10 seconds ($\chi 2(2)=0.249$, $p=.883$) [*Table 2*]. However, the time devoted to describing the anomalies that potentially indicated archaeological objects was statistically longer in groups with prior archaeological expertise, ranging from 6 minutes and 29 seconds for non-experts to 30 minutes and 19 seconds for experts in magnetic surveys ($\chi 2(2)=23.2$, $p<.001$).[2] Thus, on average, the experts participated in the experiment the longest. One takeaway from this is that the descriptive phase significantly enriches our understanding of how subject-specific knowledge of the participants impacts their interpretation of magnetic images. All participants spent a similar amount of time on silent analysis, but the time devoted to sharing their interpretation differed. On the other hand, the experiment designs in which the descriptive phase is separate from the eye-tracking experiment can significantly reduce the time spent in a fixed position.

2 Statistically significant results were obtained for all groups as shown by pairwise comparisons: non-experts–archaeologists ($p<.001$), non-experts–experts ($p=.004$), archaeologists–experts ($p=.002$).

Table 2. Average time for silent analysis and verbal description of anomalies in Experiment 1 by research group

	Non-experts (n=21)	Archaeologists (n=22)	Experts (n=5)
Silent analysis	6 min 8 sec	6 min 10 sec	5 min 23 sec
Description of anomalies	6 min 29 sec	12 min 48 sec	30 min 19 sec

Table 3. Average time devoted to the silent analysis and to the think-aloud protocol in Experiment 2 by object type

	Artifacts	Wall paintings	Excavations
Silent analysis	34 sec	3 min 50 sec	3 min 31 sec
Think-aloud protocol	1 min 23 sec	6 min 27 sec	5 min 11 sec

Another factor influencing the length of the experiment was the research procedure. In Experiment 2, experts were asked to analyze a number of archaeological remains, first in silence (in the way they usually do when analyzing a given artifact/excavation), and then thinking aloud. The participants spent different amounts of time on the silent analysis relative to the verbal description. However, in general, the analysis took longer when the subjects were asked to verbalize their thoughts [*Table 3*]. For artifacts (stone tools, flint tools, pottery sherds), the experts (n=3) devoted an average of 34 seconds to the silent analysis of a single artifact, whereas the duration of the think-aloud protocol was approximately 2.4 times longer. In the case of medieval Nubian wall paintings from Kom H in Old Dongola, the time allotted to verbal description of the state of preservation of a painting (a task assigned to three experts) was approximately 68% longer than the silent analysis. For excavations (assigned to four experts), the time increase was by 47%. In the case of this experiment, no statistical analyses were performed due to the small number of participants. However, the observations align with other studies reporting that verbalization of thoughts increases the duration of experiments (e.g. Hertzum, Hansen, and Sønderstrup-Andersen 2009).

An interesting observation related to the duration of the analysis comes from a comparison of results of Experiments 2 and 4. Both involved looking at the same three paintings from the Monastery on Kom H in Old Dongola (Nos 18, 65, and 68, see catalog by Martens-Czarnecka 2011). In Experiment 2, described above, the task of the three experts was to assess the state of preservation of the paintings, while in Experiment 4 the participants (n=67) were asked to view them as in a museum setting (free viewing paradigm). The experts in Experiment 2 spent an average of 3 minutes and 50 seconds analyzing each painting [see *Table 3*], whereas the "visitors" in Experiment 4 devoted an average of 36 seconds to each composition. Therefore, an important variable affecting experiment duration is the task itself, which, in this case, was tailored to the participants' level of expertise [*Fig. 5*].

TECHNICAL FEASIBILITY

Since the 19th century, when the eye-tracking technology began to be developed, eye-trackers have changed beyond recognition. From invasive devices, causing fatigue and requiring time-consuming data processing, eye-trackers evolved into non-invasive, user-friendly tools offering fast insights into visual processing (Płużyczka 2018).

However, technological progress has not eliminated all technical limitations of eye-trackers. Various technical constraints that still arise are related to the nature of stimuli, the individual characteristics of the participants, and the research conditions.

Three eye-tracking metrics are particularly useful for describing the technical limitations of the equipment: traceability (data loss), accuracy, and precision. The first parameter refers to the percentage of samples used to calculate gaze points in relation to the theoretical maximum of samples collected by the device.[3] The second is related to the correspondence between the actual position of the gaze and the position indicated by the eye-tracker. The third metric is associated with the consistency in obtaining of similar measurements when the subject looks at the same point. The better the results obtained for these parameters, the higher the quality of the data.

The main question was whether there existed any notable disparities in data lost between the results of the study and the results reported in the literature.

Fig. 5. Influence of tasks on the visual processing of the wall painting "Christ and the College of Apostles" from the Monastery on Kom H in Old Dongola (A): The painting was viewed by Sudanese and Western visitors, whose task was to look at it as if in a museum — a free viewing paradigm (B) and by conservators whose task was to assess its state of preservation (C). As shown in the heat maps (B and C), the Sudanese and Western visitors (B) focused mostly on faces and attributes (the more red/intensive the color, the longer the fixation duration time on a particular part of the painting), whereas the conservators (C) conducted a broader visual inspection of the paintings and their context, focusing on such elements as flaking paint, traces of animal activity or influence of weather conditions (Photo and processing T. Michalik)

3 For example, an eye-tracker operating at a 100 Hz sampling frequency produces 100 samples per second. If the software used all these samples to determine gaze points, the traceability would be 100%. However, this value is rarely reached because factors such as blinking, distractions during the experimental task (e.g. looking away from the computer screen), or unstable external conditions can negatively impact the amount of the available data.

Taking into account the differences in lighting conditions and research procedures, five data samples were distinguished from the experiments [*Table 4*] and analyzed to determine the degree of traceability.

Table 4. Characteristics of samples used to analyze traceability. Note: depending on the experiment, varying amounts of recordings were made with each participant

	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Experiment	1	2	3	4	2
Conditions	Stable artificial lighting, maintaining stable position	Stable artificial lighting, slippage	Stable artificial lighting, slippage	Changeable lighting (difference in lighting between rooms)	Natural light, slippage
Research setting	Laboratory	University offices	Museum	Monastery	Excavation
Object	Magnetic images	Artifacts (pottery, flint tools, stone tools)	Wall paintings	Wall paintings	Trench
Number of recordings	48	29	209	134	8
Number of subjects	48	9	28	67	4

The results showed that despite the differences in environmental conditions and research schemas, the average traceability remained satisfactory. In all cases, the average amount of available data was higher than 90%, ranging between 92.13% and 97.48%. The highest traceability rates were obtained for Sample 4 (m=97.12%, mdn=98%, SD=2.07), Sample 1 (m=96.88%, mdn=97%, SD=2.03), and Sample 3 (m=96.65%, mdn=98%, SD=2.80). Lower results were recorded for Sample 2 (m=93.45%, mdn=94%, SD=4.28) and Sample 5 (m=92.13%, mdn=93.5%, SD=5.08). Most of the participants obtained results higher than 90%; only in individual cases was the traceability rate between 80% and 90% [*Fig. 6*].

What might have caused the lower values in Samples 2 and 5? An explanation in the case of Sample 2 may be the nature of the researched objects and the task itself. During the analysis of artifacts, the participants mostly directed their gaze downward, towards the artifacts held in their hands, which, when combined with speaking (one of the tasks was to describe their analytical strategies), increased the chance of slippage, thus affecting traceability. Interestingly, in Sample 3, similar to Sample 2 in terms of conditions (i.e., stable lighting and verbal responses) but different in terms of research objects (wall paintings), the traceability rate was higher (96.65%). Among other factors, this difference might result from the fact that the paintings, due to their size and location, encourage a vertical position of the head, preventing slippage. In the case of Sample 5, the most probable explanation for the lower traceability was the exposure to natural light during research. As discussed in the literature, natural light may interfere with infrared light sources emitted by the eye-tracker to detect the pupil and calculate gaze angle (Binaee et al. 2021). Another factor that may have affected the outcomes for Samples 2 and 5 was the small number of participants (Sample 2 — four subjects, Sample 5 — nine subjects). In small sample sizes, the occurrence of outliers may significantly affect the results (here especially in the case of Sample 5). Thus, further analysis is needed to determine if the nature of the researched objects and natural lighting systematically decreases traceability or if the obtained results are rather related to the subjects' idiosyncrasies. For example, one could imagine that, in individual cases, lower traceability might result from atypical eye anatomy rather than external factors, or that a combination of both causes was at fault.

Overall, the obtained results were acceptable as far as traceability was concerned. First of all, full traceability is difficult to achieve due to blinking, which may cause a loss of several percent of the data (Grootjen, Weingärtner, and Mayer 2023; Nyström et al. 2024). In addition, studies utilizing wearable eye-trackers have shown that experiments conducted outside the laboratory are typically characterized by lower data quality (e.g. Wang et al. 2010; Evans et al. 2012). Thus, the traceability rates observed in the project are consistent with those reported in the literature.

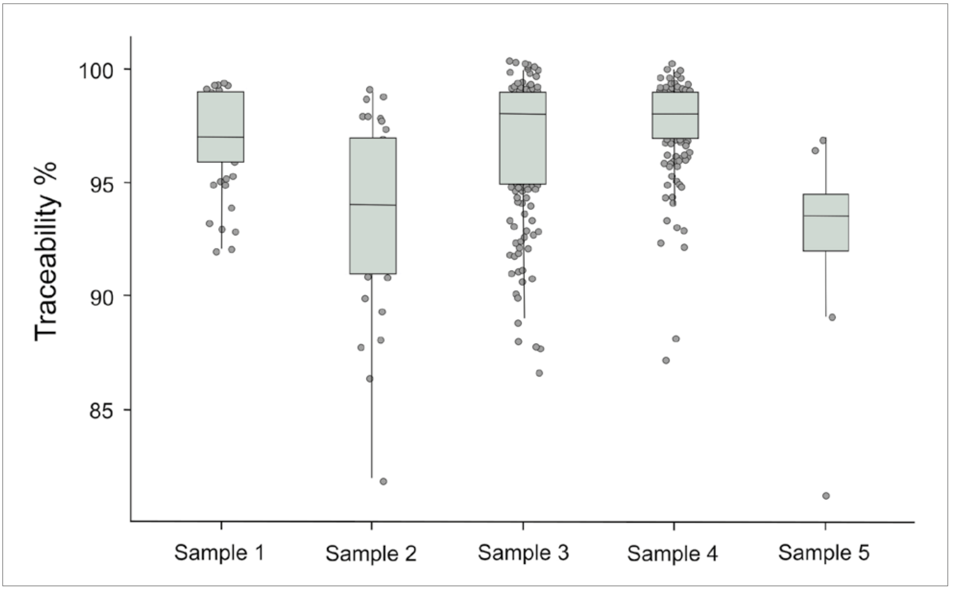


Fig. 6. Traceability results obtained in the study. Despite different external conditions and research schemas, the average traceability exceeded 90% (Processing T. Michalik)

While the results regarding traceability are satisfactory, a pilot study to Experiment 3 has shown that data loss may have been caused by fixations beyond the scope of the scene camera (the camera showing what the subject is looking at) [*Fig. 7:A*]. The problem occurred when the scene camera did not cover the whole object of interest (in this case, a painting), and the subjects fixated their gaze on its periphery without moving their head. The issue lies in the fact that the field of scene cameras is smaller than the scope of eye movements. Therefore, in the case of eye-trackers in the form of eyeglasses, data quality may improve if the subjects are asked to move their heads freely instead of moving only their eyes.

Another possible limitation was the size of the stimuli in relation to the manufacturing accuracy and precision of the device. This issue is particularly crucial when studying small, closely spaced features of artifacts (e.g. elements of vessel profiles or retouching in the case of flint tools). An eye-tracker merely approximates the real gaze location, making it difficult to determine which of the closely spaced elements the gaze is directed at. For example, an eye-tracker with a manufacturing accuracy of 0.5°, when used for an object viewed from a distance of 60 cm, may deviate by about 0.5 cm from the actual position of the gaze. Additionally, the precision error of 0.2° may increase this disparity by approximately 0.2 cm, rendering this device inadequate for research on the perception of elements spaced closer than 0.7 cm [*Fig. 7:B*]. Particularly useful in determining the feasibility of using a specific eye-tracker to study a given stimulus are visual angle calculators.[4] Knowing the magnitude of the accuracy error prior to the study can be helpful in obtaining data of sufficient quality. The problem can be avoided by implementing appropriate countermeasures, for example by manipulating the distance between the object and the participant, or by choosing an object with an optimal size.

INTERNAL AND EXTERNAL VALIDITY

While eye movements provide behavioral data, their interpretation can be a challenge, mainly because visual attention depends on a variety of external as well as internal factors (Orquin and Holmqvist 2018). For example, looking intensely at a particular part of an image may be linked either to its attractiveness (Velazquez and Pasch 2014) or to difficulties understanding it (Jacob and Karn 2003). Similarly, the speed of detecting a given element may be attributed to its visual properties (bottom-up guidance) or to the subject's intentional search for the item (top-down guidance) (Wolfe and Horowitz 2017). To avoid misinterpretations, a key aspect of eye-tracking research is to ensure its internal validity (control of factors influencing visual attention, like environmental conditions or personal characteristics) as well as external validity (enabling the extension of research results into real-life situations). Internal validity is easiest to achieve in a laboratory setting. However, laboratory-based research cannot always be translated into

4 See, e.g., https://elvers.us/perception/visualAngle/ or https://www.sr-research.com/visual-angle-calculator/.

Fig. 7. Examples of technical limitations of eye-trackers during visual analysis of artifacts: (A) the limited range of the scene camera's visual field, compared to the scope of eye movements, causes some fixations to fall outside the camera's field, resulting in lower data quality (example from the Faras Gallery, National Museum in Warsaw, painting "Christ, Virgin Mary and Child, Bishop Marianos"); (B) the manufacturing accuracy of the eye-tracker makes it impossible to distinguish which aspects of the vessel profile are being analyzed by the subject (Processing T. Michalik)

real-world situations. For instance, as demonstrated by Brieber and colleagues (2014), viewing artwork in a laboratory, as opposed to a museum, results in different visual behaviors and aesthetic experiences. Differences are also reported in the context of daily tasks (such as going out to get coffee) depending on whether they are observed on a computer screen or performed live (Foulsham, Walker, and Kingstone 2011).

Throughout the project discussed in this paper, the goal was to find a compromise between internal and external validity. External validity was maximized through contact with real monuments and archaeological data, real-life experimental tasks, and mostly non-laboratory research conditions. Simultaneously, efforts were made to maximize internal validity. The key aspects of these measures are discussed below.

One of the main challenges impacting internal validity in field eye-tracking studies is the lighting conditions. Sunlight is important from a psychological perspective, and changing illumination, among other factors, may guide attention and affect the recognition of object properties (Murray and Adams 2019; Ai et al. 2021). Thus, for example, if lighting conditions are unstable during a study of experts' visual attention, the observed disparities can be attributed not only to variations in expertise, but also to the differences in lighting. This was a factor in Experiment 2, in the study devoted to excavation analysis. The main objective of this part of the project was to record videos demonstrating the various ways in which experts analyze excavations. Therefore, it was important to ensure alike conditions for all subjects conducting the analyses. The impact of lighting conditions on the research outcomes was in this case minimized by conducting the study over a short period of time, in stable weather conditions (sunny, cloudless sky).

However, one could pose another research question — for instance, how do lighting conditions influence the detection of archaeological objects? In such a case, the lighting conditions would no longer be an obstacle but a variable helping to find the answer to the research question.

The second factor taken into account during the project was the social environment. Whether a task is performed alone or in the presence of others influences behavior in a number of ways. One of the main differences is that, in social situations, the eyes not only gather information from the environment, but also serve a communicative function (Freeth, Foulsham, and Kingstone 2013: 1–3). Since most of the research involved analyzing archaeological data, efforts were made to reduce social distractions and help participants stay focused on the task. The solution here was to minimize the presence of others during the experiments. In most cases, the subjects were accompanied only by the researcher. Experiments 3 and 4 differed slightly in this respect. In Experiment 4 (visiting medieval Nubian paintings in Old Dongola), the researcher was accompanied by an Arabic-to-English interpreter, while in Experiment 3, conducted in a museum, other visitors were also present. However in the latter case, the recordings were mostly made when no other individuals were near the paintings.

The influence of the social environment involves not only the physical presence of others but also the desire to meet the others' expectations. Thus, an important aspect of all experiments was to inform the participants that the primary focus of the research was to understand the visual processing of stimuli and not to assess the subjects' knowledge. Considering the initial state of eye-tracking research in archaeology and the absence of a frame of reference, there was no specific level of task performance expected from the subjects.

Yet another factor that can influence internal validity is personal characteristics. Here, among the relevant issues may be prior contact with the stimulus (for example, a past visit to the archaeological site), or expertise in a particular topic. Familiarity with an object usually results in faster processing of visual stimuli (Manahova, Spaak, and de Lange 2020), while expertise is related to, among other things, the focus on task-relevant elements and overlooking the irrelevant ones (Reingold and Sheridan 2011; Gegenfurtner, Lehtinen, and Säljö 2011). Selected factors influencing visual processing are presented in [*Table 5*].

Obtaining a high internal validity depends on the appropriate selection of participants based on their characteristics. The most common way of identifying the specific traits of the subjects are questionnaires and pre-research interviews. However, the method of collecting information is not neutral to the outcomes. Experiment 1 provides an interesting example. The relationship between self-assessed knowledge and the visual processing of magnetic data was one of the research questions that the experiment sought to answer. To investigate this relationship, the participants were asked to assess their knowledge of the magnetic method. Interestingly, in individual cases, archaeologists who did not professionally specialize in the analysis of magnetic images declared, on a Likert scale ranging from 1 (limited knowledge) to 5 (superior knowledge), a high level of knowledge about magnetic analysis despite their lesser experience compared to experts in the field. The questionnaire used in the experiment was designed to minimize response bias by collecting various metrics, such as the number of years of archaeological experience and years of experience in magnetic data analysis. The use of only one method of knowledge assessment (e.g. the Likert scale) might have led to misleading results. According to psychological literature, when using the Likert scale, some participants tend to choose middle or extreme values (Austin and Brunner 2003), which might influence outcomes. Another factor to consider is the Kruger-Dunning effect, a cognitive bias that causes people with a low level of expertise at a task to overestimate their ability, while those with high expertise underestimate their own competence (Kruger and Dunning 1999). Among several guides for designing research surveys and measuring social characteristics, Wolf and colleagues (2016) and Leeuw, Hox, and Dillman (2008) may serve as useful starting points.

Table 5. Selected factors influencing visual processing

Factor	Possible influence on visual processing
Gender	Gender may impact comprehensiveness of information processing (Meyers-Levy 1989) — in some circumstances, women show more exploratory visual behavior than men (Sargezeh, Tavakoli, and Daliri 2019)
Age	Saccade frequency, amplitude, peak velocity, and mean velocity are reduced with age (Dowiasch et al. 2015). Additionally, older adults exhibit a decline in the range in horizontal and upward gazes in comparison to younger subjects (Lee et al. 2019)
Knowledge	Prior training may influence the speed of information processing, selection of information, as well as the ability to extract information from parafoveal regions — experts often achieve better results in all of these aspects compared to non-experts (Gegenfurtner, Lehtinen, and Säljö 2011)
Cultural background	The level of collectivism or individualism within a culture may influence sensitivity to objects and backgrounds — participants from collectivistic cultures tend to be more sensitive to contextual information, whereas participants from individualistic cultures tend to center more on focal objects (e.g. Nisbett and Masuda 2003; Goh, Tan, and Park 2009)
Stimulus familiarity	Familiar objects often require a shorter processing time than novel objects (Gobbini et al. 2013). Moreover, some research suggests that detecting changes to familiar objects is easier (Goktepe and Schütz 2023)

CONCLUSIONS

Eye-tracking research has led to significant advancements in various fields. Archaeology is no exception, as eye-tracking may offer valuable insights and solutions to problems related to studying the past. Firstly, eye-tracking can enhance archaeological research by helping to understand the psychological factors that influence the detection of archaeological objects and the analysis and interpretation of archaeological data. Secondly, by enhancing the understanding of culturally and socially grounded ways of perceiving archaeological remains it can contribute to more efficient management of archaeological heritage. Finally, eye-tracking can provide valuable insights into the evolution of the human mind and enhance our understanding of past cognition.

However, a simple application of the technology is not sufficient to achieve scientific progress. Attention must be paid to methodological aspects and feasibility of eye-tracking during planning and conducting the eye-tracking research. Subjective, technical as well as internal and external factors play a significant role.

The outcomes of the "Psychological Aspects of Creating and Acquiring Knowledge about the Past" project indicated considerable potential for the use of eye-tracking in archaeological practice. Analysis of applications of eye-tracking in various research conditions revealed that, in terms of traceability, the results were comparable to studies in other fields, thus asserting the feasibility of integrating eye-tracking into archaeology. However, these promising findings should be complemented by reflections on three key issues to ensure high-quality research.

The first issue is to determine what we want to investigate, and what we do not want to investigate. The greatest challenge in planning the eye-tracking research is the selection of a research topic and determining factors that may hinder the research. Visual attention is susceptible to environmental and social factors, as well as the participants' personal characteristics [see *Table 5*]. These factors may influence research outcomes, potentially leading to errors in interpretation. In the context of analysis of archaeological data, crucial aspects are an accurate assessment of the subject's level of expertise, as well as stable and comparable research conditions, particularly during fieldwork. In the management of archaeological heritage, key aspects are the social and psychological characteristics of participants. Visitors to archaeological sites may differ in knowledge about heritage, and, as psychological literature suggests, first- or second-time viewing also strongly influences visual processing.

The second aspect concerns data selection and methods of their acquisition. The analyses demonstrated the importance of choosing a device compatible with the stimulus (e.g. size of the researched elements versus the manufacturing accuracy and precision of the eye-tracker) and the research conditions (e.g. sensitivity to variable lighting conditions). The repertoire of archaeological analyses is vast and each requires a careful selection of the eye-tracker best-suited for the task. However, knowing the technical constraints of each device can aid in mitigating problems. For example, manipulating the distance between the eyes and the stimulus influences the angular size of the image, limiting accuracy errors. In the case of artifact analysis, a mitigating approach might involve combining the study of real artifacts with their models viewed with screen-based eye trackers that typically offer better parameters.

Finally, the success of eye-tracking research depends on the organization of the research process. Interviews conducted in the project demonstrated a positive subjective perception of eye-tracking. However, in order to ensure subjective comfort and engagement of the participants, constraints in maintaining visual attention during the experimental procedure should be taken into consideration. This is where pilot studies are invaluable. Conducting them may help prevent misunderstanding of the research task, reduce task fatigue, or eliminate weaknesses in the research procedure. Sometimes simple solutions, such as breaks between tasks, or instructions encouraging participants to move their head freely [see *Fig. 7:A*] may significantly improve data quality.

As with any other scientific method, eye-tracking is not without its challenges. However, its scientific value can be bolstered with awareness of its limitations. This paper may help gain a better understanding of this exciting yet demanding method.

Dr. Tomasz Michalik
https://orcid.org/0000-0003-1369-5887
University of Warsaw
Polish Centre of Mediterranean Archaeology
t.michalik@uw.edu.pl

How to cite this article: Michalik, T. (2024). Eye-tracking in archaeological practice: applications, potential, and challenges. *Polish Archaeology in the Mediterranean,* 33, 433–458. https://doi.org/10.37343/uw.2083-537X.pam33.07

References

Ai, X., Wu, Z., Guo, T., Zhong, J., Hu, N., and Fu, C. (2021). The effect of visual attention on stereoscopic lighting of museum ceramic exhibits: A virtual environment mixed with eye-tracking. *Informatica*, 45, 713–729

Austin, P.C. and Brunner, L.J. (2003). Type I error inflation in the presence of a ceiling effect. *The American Statistician*, 57(2), 97–104

Binaee, K., Sinnott, C., Capurro, K.J., MacNeilage, P., and Lescroart, M.D. (2021). Pupil tracking under direct sunlight. *ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Adjunct)*, *Article 18*, 1–4. https://doi.00.1145/3450341.3458490

Brieber, D., Nadal, M., Leder, H., and Rosenberg, R. (2014). Art in time and space: Context modulates the relation between art experience and viewing time. *PLoS ONE*, 9(6), Article e99019

Campanaro, D. and Landeschi, G. (2022). Re-viewing Pompeian domestic space through combined virtual reality-based eye tracking and 3D GIS. *Antiquity*, 96(386), 479–486

Chu, S.H., Kim, E.J., Jeong, S.H., and Park, G.L. (2015). Factors associated with willingness to participate in clinical trials: A nationwide survey study. *BMC Public Health*, 15, Article 10(2015)

Criado-Boado, F., Alonso-Pablos, D., Blanco, M.J., Martínez, L.M., Porto, Y., and del Barrio-Álvarez, E. (2023). Gazed pottery: An archaeometric-cognitive approach to material culture visuality. *Journal of Archaeological Science*, 154, Article 105770

Criado-Boado, F., Alonso-Pablos, D., Blanco, M.J., Porto, Y., Rodríguez-Paz, A., Cabrejas, E., ... Martínez, L.M. (2019). Coevolution of visual behaviour, the material world and social complexity, depicted by the eye-tracking of archaeological objects in humans. *Scientific Reports*, 9, Article 3985

Dowiasch, S., Marx, S., Einhäuser, W., and Bremmer, F. (2015). Effects of aging on eye movements in the real world. *Frontiers in Human Neuroscience*, 10(9), Article 46

Drzewiecki, M., Kurcz, M., Ciesielska, J., Michalik, T., Czyżewska-Zalewska E., Kiersnowski, K., and Ryndziewicz, R. (2021). Interdisciplinary research into the legacy of the medieval metropolis of Soba in a modern Khartoum suburb. *African Archaeology Review*, 38, 597–623

Drzewiecki, M., Ryndziewicz, R., Ciesielska, J., Kurcz, M., Michalik, T., and Czyżewska-Zalewska, E. (2022). The spatial organisation of Soba: A medieval capital on the Blue Nile. *Antiquity*, 96(385), 213–220

Evans, K.M., Jacobs, R.A., Tarduno, J.A., and Pelz, J.B. (2012). Collecting and analyzing eye tracking data in outdoor environments. *Journal of Eye Movement Research*, 5(2), Article 6

Forte, M., White, L.E., Straneva, K., and Woytowitz, S. (2022). Virtual reality and neuroarchaeology: Visual perception and cognition of an archaeological excavation. In A. Banzi (ed.), *The brain-friendly museum: Using psychology and neuroscience to improve the visitor experience* (pp. 147–158). London: Routledge

Foulsham, T., Walker, E., and Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, 51(17), 1920–1931

Freeth, M., Foulsham, T., and Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PloS ONE*, 8(1), Article e53286

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. *Educational Psychology Review*, 23(4), 523–552

Gobbini, M.I., Gors, J.D., Halchenko, Y.O., Rogers, C., Guntupalli, J.S., Hughes, H., and Cipolli, C. (2013). Prioritized detection of personally familiar faces. *PLoS ONE*, 8(6), Article e66620

Goh, J.O., Tan, J.C., and Park, D.C. (2009). Culture modulates eye-movements to visual novelty. *PLoS ONE*, 4(12), Article e8238

Goktepe, N. and Schütz, A.C. (2023). Familiar objects benefit more from transsaccadic feature predictions. *Attention, Perception, and Psychophysics*, 85(6), 1949–1961

Grootjen, J.W., Weingärtner, H., and Mayer, S. (2023). Highlighting the challenges of blinks in eye tracking for interactive systems. *2023 Symposium on Eye Tracking Research and Applications (ETRA '23), Article 63*, 1–7. https://doi.org/10.1145/3588015.3589202

Hertzum, M., Hansen, K.D., and Sønderstrup-Andersen, H.H.K. (2009). Which is the better prompt in thinking-aloud studies, "What are you trying to achieve?" or "Keep talking"? In O.W. Bertelsen and A.M. Kanstrup (eds), *DHRS2009: Proceedings of the Ninth Danish HCI Research Symposium* (pp. 30–33). Aarhus: Aarhus University

Hirst, S.J.H. (2019). *Upper Palaeolithic female representations: an eye tracking study* (MA thesis). Bournemouth University. Retrieved from: https://eprints.bournemouth.ac.uk/32317/

Holmqvist, K., Örbom, S.L., Hooge, I.T.C., Niehorster, D.C., Alexander, R.G., Andersson, R., ... Hessels, R.S. (2023). Eye tracking: Empirical foundations for a minimal reporting guideline. *Behavior Research Methods*, 55, 364–416

Hüttermann, S., Noël, B., and Memmert, D. (2018). Eye tracking in high-performance sports: Evaluation of its application in expert athletes. *International Journal of Computer Science in Sport*, 17(2), 182–203

Jacob, R.J.K. and Karn, K.S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In R. Radach, J. Hyona, and H. Deubel (eds), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573–603). Amsterdam: Elsevier Science

Jones, W. and Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-

old infants later diagnosed with autism. *Nature*, 504, 427–431
Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134
Lee, W.J., Kim, J.H., Shin, Y.U., Hwang, S., and Lim, H.W. (2019). Differences in eye movement range based on age and gaze direction. *Eye*, 33, 1145–1151
Leeuw, E.D., Hox, J.J., and Dillman, D.A. (eds). (2008). *International handbook of survey methodology*. New York: Lawrence Erlbaum Associates
Manahova, M.E., Spaak, E., and de Lange, F.P. (2020). Familiarity increases processing speed in the visual system. *Journal of Cognitive Neuroscience*, 32(4), 722–733
Martens-Czarnecka, M. (2011). *The wall paintings from the Monastery on Kom H in Dongola* (=*Nubia* 3; *Dongola* 3). Warsaw: PCMA UW
Meyers-Levy, J. (1989). Gender differences in information processing: A selectivity interpretation. In P. Cafferata and A.M. Tybout (eds), *Cognitive and affective responses to advertising*. Lexington, MA: Lexington Books/DC Heath and Com
Michalik, T. (2014). Between eye and the mind. Technology, cognition and knowledge development—eye-tracking study report. *AARG News*, 48, 24–32
Michalik, T. (2024). Enhancing archaeological teaching through eye-tracking: A pilot study on eye movement modelling examples and teaching artefact analysis. *Archaeologies. Journal of the World Archaeological Congress*. https://doi.org/10.1007/s11759-024-09510-y
Murray, R.F. and Adams, W.J. (2019). Visual perception and natural illumination. *Current Opinion in Behavioral Sciences*, 30, 48–54
Nisbett, R.E. and Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences*, 100(19), 11163–11170
Nyström, M., Andersson, R., Niehorster, D.C., Hessels, R.S., and Hooge, I.T.C. (2024). What is a blink? Classifying and characterizing blinks in eye openness signals. *Behavioral Research Methods*, 56, 3280–3299
Orquin, J.L. and Holmqvist, K. (2018). Threats to the validity of eye-movement research in psychology. *Behavior Research Methods*, 50, 1645–1656
Peysakhovich, V., Rudi, D., Causse, M., Kiefer, P., Dehais, F., and Raubal, M. (eds). (2020). *Eye-tracking in aviation: Proceedings of the 1st International Workshop (ETAVI)*. Zurich: ETH Library
Płużyczka, M. (2018). The first hundred years: A history of eye tracking as a research method. *Applied Linguistics Papers*, 25(4), 101–116
Ramkumar, N., Fereydooni, N., Shaer, O., and Kun, A.L. (2019). Visual behavior during engagement with tangible and virtual representations of archaeological artifacts. In J. Cauchard, V. Gentile, M. Khamis, and S. Sorce (eds), *Proceedings of the 8th ACM International Symposium on Pervasive Displays* (Article 21). Association for Computing Machinery
Reingold, E.M. and Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. In S.P. Liversedge, I.D. Gilchrist, and S. Everling (eds), *The Oxford*

handbook of eye movements (pp. 523–550). Oxford: Oxford University Press

Sargezeh, B.A., Tavakoli, N., and Daliri, M.R. (2019). Gender-based eye movement differences in passive indoor picture viewing: An eye-tracking study. *Physiology and Behavior*, 206, 43–50

Sharafi, Z., Sharif, B., Guéhéneuc, Y.-G., Begel, A., Bednarik, R., and Crosby, M. (2020). A practical guide on conducting eye tracking studies in software engineering. *Empirical Software Engineering*, 25, 3128–3174

Sheridan, R., Martin-Kerry, J., Hudson, J., Parker, A., Bower P., and Knappe, P. (2020). Why do patients take part in research? An overview of systematic reviews of psychosocial barriers and facilitators. *Trials*, 21, Article 259

Silva-Gago, M. and Bruner, E. (2023). Cognitive archaeology, attention, and visual behavior. In E. Bruner (ed.), *Cognitive archaeology, body cognition, and the evolution of visuospatial perception* (pp. 213–239). Cambridge, MA: Academic Press

Silva-Gago, M., Fedato, A., Hodgson, T., Terradillos-Bernal, M., Alonso-Alcalde, R., and Bruner, E. (2021). Visual attention reveals affordances during Lower Palaeolithic stone tool exploration. *Archaeological and Anthropological Science*, 13, Article 145

Silva-Gago, M., Fedato A., Hodgson, T., Terradillos-Bernal, M., Alonso-Alcalde, R., and Bruner R. (2022a). The influence of tool morphology on visual attention during the interaction with Lower Palaeolithic stone tools. *Lithic Technology*, 47(4), 328–339

Silva-Gago, M., Fedato, A., Terradillos-Bernal, M., Alonso-Alcalde, R., Hodgson, T., and Bruner, E. (2022b). Does knowledge influence visual attention? A comparative analysis between archaeologists and naïve subjects during the exploration of Lower Palaeolithic tools. *Archaeological and Anthropological Science*, 14, Article 114

Silva-Gago, M., Ioannidou, F., Fedato, A., Hodgson, T., and Bruner E. (2022c). Visual attention and cognitive archaeology: An eye-tracking study of Palaeolithic stone tools. *Perception*, 51(1), 3–24

Silva-Gago, M., Terradillos-Bernal, M., Hodgson, T., and Bruner E. (2022d). The role of vision during Lower Palaeolithic tool-making. *Journal of Paleolithic Archaeology*, 5, Article 19

Stuart, S. (ed.). (2022). *Eye tracking: Background, methods, and applications*. New York: Humana Press

Tabatabaeian, S. (2018). *Eyes in the dark: Using eye-tracking technology to investigate the effects of darkness on human cognition and implications for cave archaeology* (MA thesis). University of California. Retrieved from: https://escholarship.org/uc/item/65k1c6zr

Tokitsu, Y. (2004). Visual scanning patterns of skilled archaeologists when observing pottery. *The Japanese Journal of Cognitive Psychology*, 1(1), 75–84

Tunga, Y. and Cagiltay, K. (2023). Looking through the model's eye: A systematic review of eye movement modeling example studies. *Education and Information Technologies*, 28, 9607–9633

Ulas, B.S., Emirzeoglu, M., and Burgert, O. (2022). Is eye-tracking a feasible interaction paradigm for large displays in the sterile field? A technical prototype for the presentation of HIS data and for supporting surgical hand disinfection. *Current*

Directions in Biomedical Engineering, 8(1), 78–80
Vehlen, A., Spenthof, I., Tönsing, D., Heinrichs, M., and Domes, G. (2021). Evaluation of an eye tracking setup for studying visual attention in face-to-face conversations. *Scientific Reports*, 11, Article 2661
Velazquez, C.E. and Pasch, K.E. (2014). Attention to food and beverage advertisements as measured by eye-tracking technology and the food preferences and choices of youth. *Journal of the Academy of Nutrition and Dietetics*, 114(4), 578–582
Wang, Y., Mehler, B., Reimer, B., Lammers, V., D'Ambrosio, L.A., and Coughlin, J.F. (2010). The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics*, 53(3), 404–420
Watalingam, R.D., Richetelli, N., Pelz, J.B., and Speir, J.A. (2017). Eye tracking to evaluate evidence recognition in crime scene investigations. *Forensic Science International*, 280, 64–80
Wolf, C., Joye, D., Smith, T.W., and Fu, Y. (2016). *The SAGE handbook of survey methodology*. London: Sage
Wolfe, J. and Horowitz, T. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1, Article 0058