Eye tracking in Educational Science: Theoretical frameworks and research agendas

Halszka Jarodzka
Open University of the Netherlands, the Netherlands
& Lund University, Sweden

Kenneth Holmqvist
Lund University, Sweden
& UPSET, NWU Vaal, South Africa

Hans Gruber
University of Regensburg, Germany
& Turku University, Finland

Eye tracking is increasingly being used in Educational Science and so has the interest of the eye tracking community grown in this topic. In this paper we briefly introduce the discipline of Educational Science and why it might be interesting to couple it with eye tracking research. We then introduce three major research areas in Educational Science that have already successfully used eye tracking: First, eye tracking has been used to improve the *instructional design* of computer-based learning and testing environments, often using hyper- or multimedia. Second, eye tracking has shed light on *expertise* and its development in visual domains, such as chess or medicine. Third, eye tracking has recently been also used to promote visual expertise by means *of eye movement modeling examples*. We outline the main educational theories for these research areas and indicate where further eye tracking research is needed to expand them.

Keywords: applied eye tracking, education, learning, expertise

Introduction

Eye tracking has been developed to measure 'where we look at'. For a long time and up until now, optimizing the apparatuses to measure accurately and unobtrusively how the eyes move, considerations which eye movements can be distinguished from a neurological perspective (cf. the discussion of whether post-saccadic oscillations are separate eye movements or belong to saccades), and developing software to detect these different types of eye movements were in focus. These topics are still ongoing and there is still plenty room for this fundamental eye tracking research. But already from the beginning, these apparatuses were used – irrespective of the many fundamental unknowns and imperfections – to apply them to answer research questions from other fields. This applied eye tracking research field began with letting people view art paintings (Yarbus, 1967). Quickly linguistics jumped onto the eye tracking train and this became probably the best investigated field of applied eye tracking research (Rayner, 1998, 2009). Later on, usability and human-computer-interaction researchers discovered the value of eye tracking for their purposes (Jacob & Karn, 2003). A rather young field of applied eye tracking research is the one of Educational Science that we would like to introduce here to the readers. Let us begin with what Educational Science actually entails.

Educational Science investigates how people learn and how this learning can be fostered with instruction. But what is learning? Kids at school begin with deciphering single letters and end up analyzing complex texts and relate these to accompanying graphs or pictures. University students begin with studying countless facts over years to finally become highly specialized experts who effortlessly diagnose complex problems. Hence, learning is the act of acquiring or improving knowledge, skills or behavior. Its result is a persistent change of these. Learning follows a trajectory from an initial encounter with a topic or task, such as studying a textbook page for 30 minutes, to mastering it on high levels of expertise, in

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

professional development lasting for decades. Thus, learning is rather a process than merely an outcome, such as a grade or a diploma. Researchers in Educational Science investigate this process to understand how learning is constituted and how it can be fostered through instruction.

Eye tracking (Holmqvist et al., 2011) has become an important tool to investigate learning processes over the past years. The reason for this is that we take most information in via our eyes; this is true when we learn, but also when we execute a professional task. Consider for instance scientific illustrations. Such illustrations on the composition or functioning of diverse systems have been around since hundreds of years. Below you see an example from the 19th century (Lilienthal, 1889) on the flight of birds (Figure 1). Not only professionals had to deal with such illustrations, but also students had to use them to study the subject matter. Nowadays, with increasing possibilities to create visualizations, their use, but also their variability has mushroomed. For instance, professionals have to operate complex computer-generated simulations (e.g., interactive 3D medical images), while students have to learn from all sorts of visualizations, such as videos, and often they have to integrate information from many sources. And these are just few examples of where eye tracking can aid in understanding and even improving learning and its instruction within Educational Science.
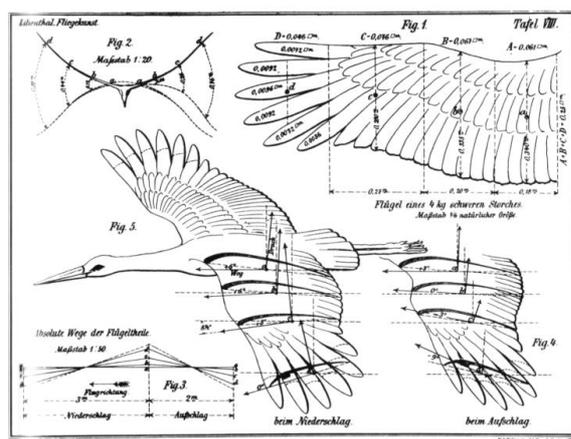

Figure 1. Scientific illustration on the flight of birds. Otto Lilienthal, Der Vogelflug als Grundlage der Fliegekunst, Berlin, 1889.

Nowadays, learning often takes place in environments that are rich in information. These environments may be learning materials, such as textbooks or e-learning settings. But they may also be working environments, such as a surgical room for medical residents or a flight simulator for pilots. Often, they can be so information-rich that they can easily overwhelm the learner. Basically, there are two possibilities to deal with this issue. First, the environment can be adapted to the learner. This approach is most effective for initial stages of learning and is called *Instructional Design*. Instructional material that is designed to optimally make use of the human cognitive information processing system as well as the abilities of the learner enables the learner to autonomously and efficiently make progress. In later stages of learning, it is important to encounter the environments in their full complexity. This is for instance the case in workplace learning. In such cases, the second option comes into play, namely, scaffolding the learner to the environment. This part of educational research is called *expertise development.* Again with the long-term aim to enable the learner to autonomously develop. The theories used in Educational Science are based on findings from fundamental research on cognition and perception, but are at the same time applicable to concrete educational practice.

In the following we will describe these two areas of research in education with concrete examples from our own research. Next we will show how both areas can be integrated into a training method of visual expertise, called eye movement modeling examples.

Instructional Design – adapting the environment to the learner's abilities

Theories of human learning – the working memory perspective

Let us begin with the initial stage of learning: a person who has little prior knowledge on a topic wants to learn new facts from a textbook, for instance about the functioning of a car engine. The material presented in this book contains a text describing the functioning of this engine, but also several graphs that show how the different elements of the engine would move at different stages of the stroke cycle. This person might experience quite some difficulties to relate all this information into one coherent mental model in his or her mind. He or she might be also distracted by a picture of a fancy car placed on this page. The research area of Instructional Design

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

investigates how to construct learning material that optimally supports the learner. One very important aspect of this is how the material is visually presented.

The strongest focus in Instructional Design lies on the (visual) flow and processing of information to and within working memory. This view is based on (a simple version of) Baddeley's working memory model (Baddeley, 2012) and Paivio's dual coding theory (Paivio, 1991). The two most influential theories on Instructional Design are the Cognitive Theory of Multimedia Learning (CTML: Mayer, 2009) and the Cognitive Load Theory (CLT: Chandler & Sweller, 1991). Both theories assume that (a) the working memory capacity is limited and learning can only take place if enough capacity is available and not consumed by 'bad' Instructional Design. Moreover, (b) learning only takes place if the learner actively engages with the learning material or the task. The *Cognitive Load Theory* (Chandler & Sweller, 1991) mainly states that the working memory capacity can be consumed by different types of load that can either be attributed to the difficulty of the task itself (known as intrinsic load), ineffective layout of the instructional material (known as extraneous load), or active elaborations on the task content (known as germane load). Only the latter results in learning. The *Cognitive Theory of Multimedia Learning* (Mayer, 2009) focuses on the working memory's dual coding in interaction with the instructional material and long-term memory. This theory predicts how pictures and words are processed in working memory depending on their modality (written or spoken) and integrated with long-term memory content. For learning to occur, relevant information from the material must be visually selected and integrated, organized in mental models and integrated with prior knowledge. If this happens, a person learned. It is easy to see that the theories include statements on perceptual processes (e.g., visual search of relevant information; integration of information from different sources), although these processes were not directly tested when these theories were formed.

Both theories result in astonishingly similar guidelines on how to design (the layout of) instructional material (Mayer & Moreno, 2003; Sweller, Van Merriënboer, & Paas, 1998). The aim of these guidelines is to decrease unnecessary cognitive processes (i.e., extraneous load) and to foster cognitive processes leading to learning (i.e., germane load). These guidelines shall make learning *efficient* (i.e., as much content learnt within as little time as possible). The most established guidelines include

- Seeking **coherence of information**. First and foremost it is crucial to avoid unnecessary information presented in instructional material, such as decorative pictures. As the learner tries to make sense out of every information given and integrate it with the other presented information and with own prior knowledge, irrelevant information will only unnecessarily consume cognitive capacities.

- Avoiding **redundant information**. The exact same information should not be given in different formats, because the learner tries to integrate all information with each other as well as with prior knowledge. This in turn costs cognitive capacities, which are not available for learning any more. One common 'bad' example is presenting a text on the slides and reading it out loud at the same time.

- Making use of **multimedia**. Even though the exact same information should not be presented in different modalities, preferable the same subject matter should be presented in different ways. For instance, an explanation of a car engine is easier to understand with an accompanying picture or animation.

- Making use of **different modalities**. To account for the dual-coding characteristics of working memory, instructional material should present related information in different modalities. For instance, a graph accompanied by an audio text instead of a written text.

- Avoiding **split attention** by seeking contiguity. Instructional material should present related information that needs to be integrated in closely, both in space and time. For instance, the legend of a graph should better be incorporated in the graph itself than presented on the side.

More principles were developed over time and fill entire textbooks (Mayer, 2009), but these are the most fundamental ones. These guidelines sound valid and were often supported by empirical studies – but not always.

Testing learning theories in educational practice

The above described theories were developed based on many empirical studies that were conducted under specific circumstances. We will exemplify this with the

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

studies of Mayer (for an overview of 15 years of studies: Mayer, 2009) and describe how new studies should enrich these findings. First, most studies were conducted with psychology students as participants. This is common research practice, as psychology students have to participate in research for course credits and form the backbone of a lot of psychological research. For many research topics that should be equal across humans (e.g., perception, memory) psychology students are valid participants. For educational research, however, they represent a preselected group with very specific characteristics that may influence the outcomes (e.g., in Germany only students with very high grades are allowed to enter psychology study). Thus, we argue that it is crucial to test the actual target group of a learning material when investigating educational principles. Second, the illustrations used were very specific. Mayer used in most studies short black and white drawings (animated or static) showing the formation of lightning (or a bicycle pump). Of course it was important to keep the material constant when investigating different principles. Nowadays, however, we must acknowledge that this was a very specific format (simple black & white drawings) and a specific topic (shouldn't lighting formation be known to university students?). Third, these studies used short, one sentence texts in English. This may have caused artefacts in the findings. For instance, research suggests that a modality effect only occurs for short sentences, while for long sentences only the last part is affected (Rummer, Schweppe, Fürstenberg, Scheiter, & Zindler, 2011) or that it might even occur only for English text (Lindow et al., 2011). We argue that it is necessary to test the guidelines and principles found thus far on diverse material that probably uses more up-to-date multimedia.

In the following, we present two examples, where eye tracking shed light on the processes underlying these effects that were carried out in ecologically valid scenarios. In the first example, we tested the **split-attention effect** (Jarodzka, Janssen, Kirschner, & Erkens, 2015). In our study, we used multimedia material on the topic of arts that is used nation-wide for assessment of all Dutch pupils at secondary school level. Moreover, our participants were 16 years old pupils. So, we used ecologically valid material that was tested with the actual target group. The material itself consisted not only of one task, but of eight tasks. Each task consisted of a text paragraph describing the task background and additional multimedia material, such as pictures, text or videos. We compared

two versions of this material (Figure 2): In one version, all additional material was presented on one side of the screen and the task text on the other. This is a classic split-attention design as the pupils must visually search for the related information. In the second version, all additional material was placed within the text, right where it was referred to. This corresponds to a classic integrated design as it allows the pupils to process the multimedia information right when it is needed.

Figure 2. The computer-based testing environment in a split (left) and in an integrated design (right). Adapted from Jarodzka, Janssen, et al. (2015), pp. 808 & 809.

Surprisingly, pupils achieved better test scores in the split-version of the test (50% correct, vs. 44% correct in the integrated format). Eye tracking data showed that pupils largely neglected the additional information in the split-design (32 sec fixation time). Contrary to the predictions of the CTMML (Mayer, 2009), pupils did not put a lot of effort to integrate the related information that would have consumed up cognitive capacity (5 points on a 9-point score for both conditions). Actually, these pupils were 'lazy' (or clever!) and ignored everything that they figured was not mandatory to solve the task. This was indeed the better strategy as it turned out that this additional information was not crucial to solve these tasks correctly. So was the integrated design pointless? On the contrary! Eye tracking results showed that exactly the same pupils processed all information in the integrated design (44 sec fixation time). Hence, they might have built a richer mental model in these cases. Probably, the test items were just not appropriately designed to tackle this richness of the mental model. Either way, learners might not always be as eager to actively process all given information as multimedia theories assume them to be.

In another example, we investigated the **multimedia effect** (Ögren, Nyström, & Jarodzka, 2016). In this study, we used multimedia material on the topic of vector calcu-

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

lus. Again, for our participants this was relevant educational material, as they were university physic students. These students solved eight tasks. Each task was composed of a text describing the problem including a formula, and a statement about this formula that the students had to confirm or reject (i.e., task performance). Additionally, half of the problems included a graph that presented one exemplary instance of the formula (Figure 3).
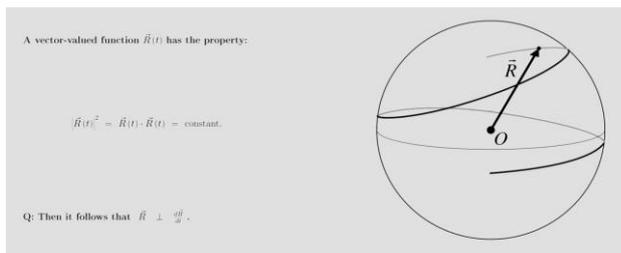

Figure 3. Exemplary task from the multimedia condition. Adapted from Ögren et al. (2016).

CTMML would predict that such an additional visualization should enrich the mental model the students are building and thus, lead to better performance. This was not what we found (56% correct with graphs vs. 52% correct without graphs). Instead, we found a bias in students to confirm the statement, if a graph was present (65% confirmation vs. 47% rejection). This is in line with findings that scientific pictures make text appear more credible (McCabe & Castel, 2008). Hence, our students probably saw the graph, judged it as being correct and concluded the same for the statement. Eye tracking data revealed that in the multimedia tasks, students paid less attention to the task description (50% vs. 40%) and to the statement (45% vs. 40%) – obviously, as they also looked at the graph (20%). The amount of looking at the graph was not related to task performance (dwelling on graph when answering correctly 20% vs. incorrectly 19%). Looking at the statement, however, was positively related to task performance (dwelling on statement when answering correctly 43% vs. incorrectly 38%). Also, many transitions between the statement and the graph were positively related to task performance (correct answer: 9 transitions vs. 6 transitions for incorrect answers). Lin and Lin (2014) received similar findings when investigating geometrical problem solving with eye tracking: while looking at the graph was an indicator for perceived diffi-

culty, looking at the area where the task performance actively takes place (here: calculation area) was positively correlated with task performance. Consequently, we must specify the CTMML based on our findings: it is not enough that the learners process a graph; they must process it in the context of the main task question. Only then graphs are beneficial, otherwise they might even pursue learners to be uncritical. Moreover, a recent study by Krejtz, Duchowski, Krejtz, Kopacz, and Chrzastowski-Wachtel (2016) has shown that the type of graph that is presented plays a role: interactive graphs evoke most systematic text-graph integrative saccades than static or dynamic graphs. Future research should investigate, whether this has also a positive effect on learning outcomes.

Research agenda for Instructional Design theories

We can conclude from these two examples already that eye tracking can help to explain unexpected findings, as it allows unique insights into processes underlying learning outcomes. One possible reason for the unexpected findings might be that the perceptual processes assumed by multimedia theories (CTML, CLT) were not directly tested with eye tracking when these theories were developed. These theories have been very helpful heuristics to design instructional material. However, now we must unravel new evidence to further develop, specify, correct, and form these theories. The following issues should be considered in future eye tracking research to achieve this:

- In the latter example presented above (Ögren et al., 2016), we saw that the guidelines given, might need to be specified. Hence, it is crucial to test also the **other guidelines for Instructional Design** with eye tracking, but also under ecologically valid circumstances (i.e., actual learning material with real students).

- In the first example above (Jarodzka, Janssen, et al., 2015), we saw that even if assuming that those guidelines are appropriate, some **basic pre-assumptions of these theories** might not be (e.g., that learners do their best to actively integrate material). Hence, it is crucial to test also these. In particular, the many assumptions about perceptual processes must be tested directly with eye tracking.

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

- The research discussed thus far considered cognitive processes. However, **metacognitive processes** are also crucial for learning (i.e., monitoring what I already can do what I still need to practice). However, too little research has been conducted on this important topic until now (Van Gog & Jarodzka, 2013).

- Finally, eye tacking research is conducted in laboratories where one participant at a time is tested under minimal disturbance. This has, however, nothing to do with educational practice. From social psychology research, we know that performing a task in the presence of others might be inhibiting, but also facilitating (Bond & Titus, 1983). Eye tracking research also shows effects of social presence on attention (Oliva, Niehorster, Jarodzka, & Holmqvist, in press; Richardson et al., 2012) Hence, future eye tracking research should investigate **social effects** on processes of learning, for instance within so-called digital classrooms.

It has to be noted that eye tracking – in particular in methodological triangulation with other process data – cannot only be used to derive instructional guidelines, but also to concretely usability test concrete computer-based multimedia learning environments. For a comprehensive description on how to proceed in such a case, see Groner and Siegenthaler (2009).

Expertise development – scaffolding the learner to the environment

Theories of human learning – the long-term memory perspective

So far, we have looked into initial learning processes. The more a person knows about a task or a domain, the more we must take the long-term memory into account as well. In the long-term memory all knowledge is stored and with increasing experience in a task it is re-organized. This **knowledge organization**, in turn, changes the deal for the working memory. It changes it to this extent that Ericsson and Kintsch (1995) suggested the concept of long-term working memory. For instance, with increasing numerical skills, children do not have to memorize six digits separately, but can form two *chunks* of three digits each and thus increase their working memory capacity (Miller, 1956). With ongoing mathematical education, children can even solve mathematical

problems described in text form. They quickly see the crucial cues that indicate which type of formula should be used. Based on this info, they know which other information they have to search for in the text and which they can ignore to fill in the formula. Next they solve the formula and formulate a solution to the problem. This procedure describes an exemplary use of a *schema* (Van Lehn, 1996). Similar to a chunk, a schema is not only an efficient way to store information in long-term memory, but it also expands working memory: one entire schema functions as only one entity. Thus, plenty capacity is left over to collect new information to fill in the schema's empty slots. If a schema includes a specific temporal order, such as visiting a restaurant (enter a restaurant, look for a table, order from menu, …), it is called a *script* (Schank & Abelson, 2013). Another form of knowledge organization is forming *short-cuts* within long chains of reasoning by encapsulating parts of it into entities that are only unfolded into its pieces if necessary (Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1992). The more knowledge a person has in a task and the more efficient it is organized, the faster and more correct this person can execute this task. Until he or she eventually becomes an expert (Ericsson, Charness, Feltovich, & Hoffman, 2006). For certain professions, such as medicine, we already know so much from research that these short-cuts and organizations of knowledge can be described very specific (Jarodzka, Boshuizen, & Kirschner, 2012). In the current section, we specifically focus on visual expertise and what we know so far about its knowledge and skill organization.

The specific case of visual expertise.

Expertise is defined as a *consistently superior* performance on a specified set of **representative tasks** for a domain (Ericsson & Lehmann, 1996; Ericsson & Smith, 1991). This superiority is due to the above described efficient organization of large amounts of knowledge and skills in a domain. This efficient knowledge organization reflects in different aspects, depending on the task itself. One example is the above mentioned well documented cognitive *chunking* in chess (Chase & Simon, 1973; De Groot, 1946/2008). Typically, expert and novice chess players are asked to build chess formation from memory; a task in which experts excel largely (Freyhof, Gruber, & Ziegler, 1992; Gruber, 1991). Eye tracking research revealed that this chunking is also reflected in perceptual

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

processes: experts look rather in between chess figures, while novices look at each single figure (Reingold, Charness, Pomplun, & Stampe, 2001; Reingold & Sheridan, 2011). We see that the concept of chunking in chess is reflected in two aspects: a cognitive recall performance and perceptual processes. Similar findings occur also in other domains of expertise, such as playing music (Lehmann & Gruber, 2006). In most cases, reading from notes is an important part of playing music and thus, it is one aspect of musical expertise that is investigated with eye tracking (Arthur, Blom, & Khuu, 2016; Penttinen & Huovinen, 2011; Penttinen, Huovinen, & Ylitalo, 2013, 2015).

Reingold and Sheridan (2011) provide a comprehensive overview of eye tracking research on **visual expertise**. The authors draw two main conclusions from their review. First, experts are able to *encode domain related patterns* in a superior way, which is due to their larger visual span. Second, eye tracking data of experts often entails information that they were not aware of. This is a clear indicator of experts' *tacit knowledge*. The increased visual span is a reflection of the above described chunking in perceptual processes. The tacit knowledge could be linked to encapsulated knowledge and its automated use. When reading this review, you will quickly realize that most research was conducted on the traditional expertise domains of chess and medicine. These studies used static and perceptually simple stimuli, such as chess boards or X-rays of the chest.

However, a lot of visual expertise plays a role in perceptually much more complex environments, such as air traffic control (Beck, Trenchard, Van Lamsweerde, Goldstein, & Lohrenz, 2012), new medical imaging techniques (Bertram et al., 2016), meteorology (Stofer & Che, 2014), etc. These environments are difficult for cognitive processing for two reasons (Atkinson & Shiffrin, 1968; Chandler & Sweller, 1991; Mayer, 2009). First, they are **information-rich** (Dwyer, 1976; Schnotz & Lowe, 2008). Hence, they entail large amount of information; and a lot of it is irrelevant. On top of that, the relation of thematic relevance and visual saliency is often not optimal. Hence, it is challenging to *select* the relevant information. Moreover, these environments are **dynamic** (Hegarty, 1992; Lowe, 2003). Thus, information may be transient. Also, several information elements may appear (and disappear) simultaneously (cf. split-attention effect). Consequently, it is challenging to keep information active

so that it can be *integrated*. Consequently, the stimuli used in most visual expertise research so far are not representative for most expertise domains. Thus, we cannot simply generalize these findings to information-rich or even dynamic domains. Research in this field, is in focus of the following section.

Research on visual expertise in information-rich environments

The concept of visual expertise is difficult to tackle as it entails so many different aspects (as already described above). In most cases, it is thus necessary to approach this concept from different angles by means of **methodological triangulation** (Denzin, 2012; Thurmond, 2001). One the one hand, eye tracking can tackle the perceptual aspects of visual expertise, while other data sources complete the picture on the more cognitive side, such as performance data, verbal data, and even drawings of what a person thinks where he or she looked at. Due to the nature of this concept and the research tradition, verbal data are most often used to investigate expertise (Ericsson & Simon, 1993). They can take the form of interviews, self-explanation, retrospective reports or thinking aloud (for an overview of different forms of verbal data and how to combine them with eye tracking see Chapters 3.4.8 and 4.7.3 in Holmqvist et al., 2011). If implemented carefully, verbal reports will not disturb the actual task performance. Instead, they will give us more information on the reason *why* a person looked at a certain area. In the following, we present examples from own research using this methodological triangulation for investigating visual expertise and its knowledge organization in information-rich environments.

One reason that this field is still so little investigated (Reingold & Sheridan, 2011) besides its obvious relevance as described above, are software issues. In 2010, we published the very first article investigating visual expertise with eye tracking using video material and an AOI analysis (Jarodzka, Scheiter, Gerjets, & Van Gog, 2010). This study investigated expertise in the domain of **marine zoology**. In other words, seven professors and PhD students, and 14 biology students classified the swimming modes of reef fish. In reality, marine zoologists often execute their profession under water (either snorkeling or diving). To get as close as possible to this situation, we asked participants to watch four *videos* of

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

single fish swimming in a colorful reef for as long as they wanted to. In this way, we created representative, but at the same time experimentally controllable tasks. Afterwards, they watched their own eye tracking recordings and reported what they were thinking while approaching this task (Van Gog, Paas, Van Merriënboer, & Witte, 2005). As we wanted to compare where experts and where novices looked at, we used a cumbersome manual procedure to define AIOs on videos, which delivered interesting findings: Experts clearly outperformed novices (experts: 4/4 points, novices: 3/4 points; $\eta_p^2 =.18$), which meant that they were indeed true experts in this task (not a trivial finding in expertise research!). Also, we compared the sequences in which participants inspected the different body parts of the fish. Experts were more diverse than novices (similarity of experts: 67%; novices: 72%; $\eta_p^2 =.08$). Probably, novices just followed the most visually salient features, which resulted in a rather similar scanpath. Experts, on the other hand, seem to have had different scripts to approach this task, which resulted in different scanpaths. These different scripts might be due to different forms of experience (i.e., when diving you see the fish from the side, while when snorkeling you see it from above; consequently, you rely on different features when classifying its motion). Indeed, dwell time analyses of AOI data taken together with participants' verbal reports, showed that part of the experts took a short-cut: they first classified the fish and deduced from this, how it must swim (dwell time on according AOIs of experts: 375 ms; novices: 160 ms; $\eta_p^2 =.29$; according verbal utterances of experts: 57; novices: 26; $\eta_p^2 =.56$). In sum, we found that visual expertise in marine zoology (a) leads to different types of scripts, probably depending on the concrete experience in that task, and (b) which form these scripts can take.

In a following step, we moved towards an interactive task stimulus, namely **digital pathology** (Jaarsma et al., 2016; Jaarsma, Jarodzka, Nap, Van Merriënboer, & Boshuizen, 2014, 2015). In the first study (Jaarsma et al., 2014) we compared how participants of three expertise levels diagnosed 10 pathological slides based on a two seconds inspection. They were eye tracked during this inspection and reported afterwards how they came about their diagnosis. Obviously, novices were incorrect (38% correct diagnoses), incomplete and inconclusive in their diagnosis (hardly conclusive terms or diagnostic specifications mentioned) and looked little at relevant areas (3 fixations). Experts (85%) and intermediates (87% correct

diagnoses), on the other hand, diagnosed these slides equally well. However, they differed in how they processed the slides. Experts relied on their first inspection of the relevant area (fixation dispersion 1st trial part: 135) and then further checked the slide for other potentially relevant information (2nd trial part: 167). In their explanations they mainly focused on the typicality of the slide (e.g., high usage of comparative terms). Intermediates kept inspecting the relevant area throughout the entire trial duration (fixation dispersion 1st trial part: 192; 2nd: 165) and considered many potential diagnoses (e.g., a lot of mentioning of pathologies). For their knowledge organization, we may conclude that experts have such consolidated illness-scripts that they can rely on, which leaves them capacity to check for further potential problems. Intermediates, instead, possess already according schemata, however, they still have to check many competing schemata to reach a diagnosis. Even though this study yielded interesting findings, the task we used was not really representative for this profession. Hence, in following studies (Jaarsma et al., 2015, 2016), we used a digital version of a tissue sample that could be operated as under a regular microscope: zooming in and out as well as panning around the slide. Hence, this was a highly representative task. Despite the progress in commercial eye tracking software, using a stimulus that can be individually changed that much (and that is not a website) is still challenging and requires a lot of manual work and programming. We found that experts were more efficient as they used fewer microscopic movements (e.g., opposed zooming movements: $\eta_p^2 = 0.03$; expertise effect for all navigation behavior: $\eta_p^2= 0.11$) and shorter reasoning chains to reach a diagnosis (reasoning terms used by experts: 109; intermediates: 63; novices: 159). This is in line with the findings from the first study that indicated that experts possess consolidated illness-scripts that allow fast decision making. Also, navigation data showed that experts visited fewer diagnostically relevant areas (experts: 3.05; intermediates: 3.98; novices: 4.05). This poses the question whether it is even possible to define areas as being relevant for each expertise group. It might be difficult to grasp the effects, because experts understand the stimuli so quickly. Intermediates also showed processes that are in line with Study 1: they took longer to reach a decision (expert: 86 sec; intermediates: 110 sec; novices: 152 sec) and looked more at relevant areas while basing their diagnosis on many specific abnormalities (novices: 35; intermediates: 96; experts: 94). Thus,

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

intermediates already have established schemata. However, they still need a lot of time to check them. Novices again were simply all over the place and clearly lacked any relevant knowledge (or its organization).

Another expertise domain we have investigated is **air traffic control** (Jarodzka, Gouw, Van Meeuwen, & Brand-Gruwel, 2015; Van Meeuwen et al., 2014). Controlling air traffic is a really challenging task: constantly flying in and departing airplanes need to be coordinated with a high emphasis of safety, but also on environmentally friendly travel. 31 air traffic controllers of three different expertise levels solved nine situations. Each depicted a real radar screen, with airplanes (including type, height, and speed), sectors, and start and landing points. Participants reported the optimal order of arrival of the airplanes while their eye movements were recorded. Individuals with higher levels of expertise clearly outperformed those of lower levels (experts: 4.63, intermediates: 4.30; novices: 3.82; $\eta_p^2 = .49$). Interestingly, the performance of those with higher expertise was more similar than of those of lower expertise (experts: 0.59; intermediates: 0.53; novices: 0.43; $\eta_p^2 = .44$; in contrast to our findings with marine zoologists: Jarodzka et al., 2010). In this profession it seems, thus, that there is one optimal script to solve this task. Eye tracking analyses revealed that individuals with higher expertise looked mainly at the aircrafts and at the background between them (e.g., time to first fixation on aircraft for experts: 41.59 sec; intermediates: 54.6 sec; novices: 65.06; $\eta_p^2 = .37$). This indicates that the script individuals with more expertise establish allows them to better focus on the relevant information and chunk single information entities. Novices, on the other hand, had no appropriate strategy to relay on and fall back on the sub-optimal means-end-strategy as indicated by them looking mainly at the destination of the airplanes (e.g., time to first fixation on destination for experts: 38.38 sec; intermediates: 36.62 sec; novices: 25.37 sec; $\eta_p^2 = .36$). We have to admit, though, that participants only saw *static* screenshots of radar screens. In a recent follow up study, we used a more representative task of this profession (Jarodzka. Gouw, et al., 2015). In that twelve participants with varying expertise levels worked on a simulation of an actual airport. The situation was entirely realistic including communication with other co-workers. Already the first eye tracking analyses reveal a drastic difference to the first study: novices mainly focus on the area of their own responsibility, while individuals with higher expertise

look more outside this area, including the starting and landing points of the planes. This strategy allowed them to plan ahead in this very dynamic environment. Hence, the scripts individuals with higher expertise possess in this task, must be updated dynamically if the task includes more time pressure.

Research agenda for visual expertise research

From the research presented above, but also from other research on visual expertise of teachers (Lachner, Jarodzka, & Nückles, 2016; Wolff, Jarodzka, Van den Bogert, & Boshuizen, 2016), neurological pediatrists (Balslev et al., 2012), or radiology (Kok et al., 2015; Van der Gijp et al., 2016) we have learned already a lot about visual expertise in information-rich environments. Experts use chunks (e.g., air traffic control) and shortcuts (e.g., marine zoology) and this can be also seen in their perceptual processes and measured with eye tracking. Also, we have clearly seen the use of cognitive scripts or schemata and their influence on the visual processing of an environment and vice versa in each profession. Often, even very concrete statements about the form of these schemata or scripts could be made. Still, many open research questions remain.

- To which extent can we generalize these findings? We have seen that sometimes even slight changes in the task can lead to different outcomes (cf. air traffic control), while sometimes the changes go in the same direction (cf. pathology). Also, some findings that are found in one profession (e.g., experts become more similar in air traffic control) are not true for another profession (e.g., experts in marine zoology become more diverse). Hence, future research should consistently vary **task characteristics and professions** to understand, which aspects of visual expertise are generic and which domain-specific.

- A lot of research on visual expertise has been conducted on simplified tasks. This was largely due to **technological restrictions** of the eye tracking apparatuses and software. Research should not be hold back by technological obstacles, but rather feed their development. In particular two issues must be tackled to foster ecologically valid research on visual expertise. First, the detection of smooth pursuit to enable valid analysis of dynamic stimuli. Thereby, it is not enough to detect smooth pursuit with a stand-alone

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

algorithm, but it must be implemented into existing analysis software, so it can be used in applied research as well. Second, more automated analyses for mobile eye tracking. Clearly, the truest way of analyzing visual expertise often requires real-world eye tracking. However, cumbersome manual analyses often hold researchers back.

- The presented research has shown how much we can benefit from methodological triangulation when investigating multifaceted concepts such as visual expertise. In a next step, research should **directly link the analysis of verbal and eye tracking data**. Only in this way it will be possible to make more concrete statements about the cognitive structures underlying these processes.

- Finally, it must become the ultimate aim of this research line to unravel the **organization of knowledge and skills in long-term memory** and how it develops with increasing expertise. Only then it is possible to draw meaningful conclusions from eye tracking data that go beyond superficial statements such as 'experts had longer fixation durations' that have virtually no meaning for professional or educational practice (Kok & Jarodzka, 2016).

Eye movement modeling examples: Bridging Instructional Design and expertise research

Theories of human learning – training visual aspects of expertise

So far, we have discussed how initial learning takes place, how it can be supported by Instructional Design, and which role eye tracking can play in this. Then, we have shown how individuals develop further over time and until they become experts in visual domains. In this section, we try to bring both research areas together to show how this road to visual expertise can be supported by instruction. This is not as trivial as it may sound, as Instructional Design entails the simplification of learning material, while expertise development requires to be faced with the authentic, information-rich tasks.

One very powerful way of learning authentic tasks is imitation. It is so inherent to our system that even two weeks old babies imitate adults (Meltzoff & Moore, 1977). Bandura (1977) has shown in his classic bobo doll experiment that imitation leads indeed to learning. Children watched videos of an adult playing with a 'bobo doll', which is an inflatable, large doll that stands up again once it is tipped over. Depending on the experimental condition this adult was either behaving aggressively (e.g., punching the doll) towards this doll or not. Once these children were confronted with this doll, they treated it in a similar way as the model they saw in the video before (Bandura, Ross, & Ross, 1961).

Consequently, research on teaching and training has picked up this approach. Indeed, decades of research have shown that studying examples of a model successfully executing a task is more efficient for learning than learning by trial-and-error (Kirschner, Sweller, & Clark, 2006). It is not trivial, though, to model a task. Many critical processes are not observable from outside, such as solving a mathematical equation. In such cases the model verbalizes his or her thoughts (cognitive apprenticeship: Collins, Brown, & Newman, 1989; process-oriented modeling examples: Van Gog, Paas, & Van Merriënboer, 2004). But what about perceptual processes in a visual task? We know that simply telling beginners to "look the way experts do" does work, but does not necessarily improve their performance (Kok et al., 2015). These beginners may now know *where* to look, but not *why*.

To address this issue, we developed e**ye movement modeling examples (EMME)**. These are video recordings of a model executing a task and explaining how he or she goes about that. On top of that, the model's eye movements are tracked and replayed on top of the video (Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). However, novices are often already overwhelmed with information-rich material that forms the basis of visual tasks. Adding an eye movement display on top of that, is likely to overwhelm them. An alternative is to display the model's eye movements by reducing existing information on videos (Dorr, Vig, Gegenfurtner, Martinetz, & Barth, 2008; Nyström & Holmqvist, 2008). This results in a spotlight wandering across the video, while the rest of it appears blurred. Figure 4 presents screenshots of both, a traditional and a spotlight display used in EMME.

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

Figure 4. Eye movement modeling examples with a traditional dot display (left) and a spotlight display (right). Material used in Jarodzka et al. (2010).

Research on eye movement modeling examples

Research described in the last section has shown that experts dramatically differ from novices. Hence, there is no point in trying to 'make novices act like experts'. Consequently, in our research, we have always used a systematic way to make the **expert model act more didactical**. On the one hand, the models in our studies were always not only experts in their domains, but also highly experienced in *teaching* this domain. Hence, they knew from experience which difficulties students face in these tasks and how to best explain these tasks to them. On the other hand, we used a specific *recording procedure* to ensure that the EMME videos were of high quality. First, to ensure a close relation of the voice and the eye movements of the models, we first show them the task itself (e.g., a video recording of something they need to classify). Only after they are familiar with this specific task, we begin with the recording. Such recording procedure have resulted tight gaze-voice couplings elsewhere (Richardson & Dale, 2005). Second, to shift the models' focus from the task to the novice recipient, they evaluate their own recordings based on several questions: Will a student know what each term means? Is the task explained in comprehensible enough terms for students? Is it explained in enough detail? Are all information that a student needs contained? Are all contained information really important? Such questions have shown to improve written communication of experts to novices (Jucks, Schulte-Löbbert, & Bromme, 2007). Third, if necessary, the models could revise their recordings.

We have used such EMMEs, for instance, to train the **classification of the locomotion patterns** of reef fish (Jarodzka, Van Gog, Dorr, Scheiter, & Gerjets, 2013). In the learning phase, participants studied four videos with either a dot display EMME, a spotlight EMME or a video with verbal explanations only (Figure 4). In the mean-

time, their eye movements were recorded to study whether they actually did follow the eye movement display of the model on the videos. In the testing phase, participants watched four new videos without any form of guidance or verbal explanation. They had the task to classify these videos accordingly. While watching the testing videos, participants' eye movements were recoded to investigate the efficiency of their visual search of relevant information on the videos. Then, they indicated via a questionnaire how they interpreted this information. Results showed that both EMME videos guided the eye movements of the participants to the spots where the model looked at (measured as coherence between the model's and the learner's scanpath: Spot = 15.10; Dot = 15.11; Control = 12.07; η_p^2 = .39). Moreover, in the spotlight condition, participants showed a more efficient visual search on testing videos (measured e.g., time to first fixation on relevant areas: Spot = 1236 ms; Dot = 1530 ms; Control = 1632 ms; η_p^2 = .11), while participants in the dot group exhibited better interpretation performance in comparison to the control group (measured as % correct: Dot = 74%; Spot = 69%; Control = 67%; η_p^2 = .12).

We have conducted a similar study in the domain of **diagnosing epileptic seizures** in infants (Jarodzka, Balslev, et al., 2012). The experimental procedure was just as in the study described above, except from the task: participants watched videos of infants either suffering a form of epileptic seizure or a differential diagnosis. Even though both tasks sound very different, they had crucial commonalities: participants had to identify relevant body parts (fins that were used to produce propulsion vs. limbs that might be affected by the disease) and to describe how exactly these body parts move. Based on these two steps, a classification or a diagnosis, respectively, can be made. Also, these steps rely on a visual inspection of a video input. A further difference to the fish locomotion study was that display of the eye movements: The traditional display was shown as a circle instead of a dot to not occlude relevant information on the video (e.g., a twitching eye). The spotlight display was far more subtle than in the fish locomotion study. Results showed an overall advantage of the spotlight display on attention guidance in the learning phase (measured as Euclidean distance to model's gaze: Spot = 210; Circle = 238; Control = 237; η_p^2 = .13), visual search (measured as e.g., time until looking at relevant area: Spot = 189 ms; Circle = 274 ms; Control = 289 ms; η_p^2 = .13) and interpretation perfor-

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science: Theoretical frameworks and research agendas

mance (measured as % correct: Spot = 60% Circle = 53%; Control = 50%; $\eta_p^2 = .11$) in the testing phase.

Similar training approaches have been used for visual tasks, which require hardly prior knowledge (Litchfield & Ball, 2011; Mason, Tornatora, & Pluchino, 2013; Skuballa, Fortunski, & Renkl, 2015), for expertise tasks (Leff et al., 2015; McNamara et al., 2012; Sridharan, Bailey, McNamara, & Grimm, 2012), and even for problem solving in dyads (Cherubini, Nüssli, & Dillenbourg, 2010). However, these studies did not test whether the found performance differences could be transferred to similar tasks (as we did on our studies), i.e., whether learning took place. Thus, strictly speaking, these cannot be seen as educational studies.

Research agenda for EMME

EMME as similar gaze-based approaches may be helpful in training visual tasks. Still, we should not become too enthusiastic, as there are also enough examples where these approaches had no (single conditions in the two studies reported above) or even detrimental effects (Skuballa, Schwonke, & Renkl, 2012; Van Gog et al., 2009). Hence, the question is not whether EMME does foster the performance of visual tasks (or even visual expertise), but rather, under which circumstances in does so. We thus recommend the following research questions to be addressed in the future:

- The role of the **task** and the **stimulus characteristics**: The research on EMME covers a diversity of tasks (from insight problem solving, to performance only, to transfer and learning) and a diversity of stimuli (from simple line drawings to complex videos). A systematic variation and concrete description of these factors should shed more light into when EMME are effective. For instance, existing studies already indicate that the visual complexity of the task is crucial: Van Gog et al. (2009) used a task that could be executed without perceptual input and found negative effects of EMME on performance (Van Marlen, Van Wermeskerken, Jarodzka, & Van Gog, in press). Jarodzka et al. (2013) used a fish locomotion classification task where all relevant information was visual salient. EMME was in part helpful in this case. Jarodzka, Balslev, et al. (2012) used a pediatric neurology task, where the relevant information was transient and not salient. This is where EMME were most helpful.

- The role of the **eye movement display design** is an entirely understudied aspect. Apart from two studies (Jarodzka, Balslev, et al., 2012; Jarodzka et al., 2013), none has compared different designs directly even though these studies indicate that this might be a crucial success factor for EMME. Results showed that reducing information on a spotlight manner guides visual attention on EMME videos best. Also, the spotlight facilitates visual search on testing videos most. However, the interpretation of relevant features is only enhanced, if a holistic processing is possible during learning.

- Moreover, the role of **didactizing the expert model**, as we have done in our studies, has not been directly investigated. In fact, most studies provide hardly any description on how the model's eye movements were collected. This is surprising as we know very well from research to which large extent experts and novices differ in their processing and how unlikely it thus is that forcing experts' processes upon novices can hardly work.

- Finally, the EMME methodology could be embedded into well-established methods of **expertise trainings**. For instance, the 4C-ID training (Van Merriënboer & Kirschner, 2007) is an elaborated model to design a curriculum for complex tasks. It includes modeling episodes that might easily be filled in with EMME for specific visual tasks. Another example is deliberate practice (Ericsson, Krampe, & Tesch-Römer, 1993). This method involves a detailed study of own and others performance. If the task includes visual aspects, studying the eye movements of an expert (or one owns) might provide additional benefits.

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

Discussion

In the current paper we have introduced Educational Science as a field of applied eye tracking research. We have structured it along three topics, namely Instructional Design, expertise development, and eye movement modeling examples. The topic of *Instructional Design* investigates how learning of a new skill or knowledge by optimally designing the according learning material. Educational theories on human cognitive processing, in particular in the working memory, resulted in guidelines on how to design such material and which processes learners should devote to efficiently achieve learning gains. Up until now, eye tracking helped us to understand how learners actually process such instructional material, which was not always in line with what theory predicted. Future eye tracking research on this topic can thus help to further corroborate, improve, and enrich these theories. Not only to understand and support processes of initial learning, but also to better understand how we as humans process information in working memory under realistic circumstances.

The topic of *expertise development* investigates the other side of the learning spectrum, namely people, who already have a lot of experience and knowledge on a task. How do they process information? How do they differ from people with slightly less or more experience? A large body of expertise research started already many years ago to expend towards visual processes underlying expertise and thus, eye tracking research. This research showed that, indeed, changes in long-term memory structures that come along the development of expertise influence not only working memory processing, but also visual processing of the environment and vice versa. Future eye tracking research on this promising topic must dive into more real-world scenarios with diverse tasks and information-rich, dynamic environments. Not only will we understand in this way more about the development and characteristics of visual expertise, but we will also better understand how long-term memory structures influence the way we see and interpret our environment, both in every day and in challenging situations.

The third topic we have presented are *eye movement modeling examples*. This is the youngest topic within the field of applied eye tracking research in Educational Science, but nonetheless, a very promising one. It addresses the question, how visual expertise could be trained with the help of instructional videos of real-world tasks that are explained by experts in the field. These videos include an overlay of these experts' visual focuses to support the learner in connecting the verbal explanation of the expert to the real-world complexity of the task. Of course, this research topic gives us practical implications for educational practice. But it also provides interesting research questions apart from education, such as: how to best guide eye movements of people on videos? How to support speech comprehension with displaying the eye movements of the speaker to the listener? Etc.

It is important to keep in mind that the area of applied eye tracking in Educational Science is clearly applied research. This means that the tasks and stimuli used are very diverse and less well controlled in comparison to fundamental experiments in vision science, for instance. However, they are *ecologically valid*. This is crucial for this research to allow drawing actual conclusions for educational practice. Therefore, research questions should always be developed together with stakeholders from educational practice. And the models or frameworks derived in research should always be tested 'in the wild' (aka schools, universities). But this also means that we can learn a lot from this research field on real-world processing, which in turn can be fruitful to establish new research question for fundamental research.

Furthermore, this research area is still relatively new. This means that there are no well-established eye tracking measures, like in reading research, that can be clearly related to concrete processes. This is due to the fact, that there is simply less research conducted as, for instance, in reading. But the ecologically valid nature makes it almost impossible to hope for such simple relations: each learning environment, each expertise domain are so inherently different in terms of tasks and stimuli that the eye tracking measures have to be found each time anew. The process of finding the appropriate measures must not be driven by what is given by the manufacturers. Instead, it is important to work along existing theories and carefully operationalize measures that are clearly related to concrete hypotheses.

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

Acknowledgements

References

Arthur, P., Blom, D., & Khuu, S. (2016). Music sight-reading expertise, visually disrupted score and eye movements. *Journal of Eye Movement Research*, *9*(7), 1-12. doi: 10.16910/jemr.9.7.1

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-192). New York: Academic Press.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63(1)*, 1-29. doi: 10.1146/annurev-psych-120710-100422

Balslev, T., Jarodzka, H., Holmqvist, K., De Grave, W. S., Muijtjens, A., Eika, B., . . . Scherpbier, A. (2012). Visual expertise in paediatric neurology. *European Journal of Paediatric Neurology, 16*(2), 161-166. doi: 10.1016/j.ejpn.2011.07.004

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs: Prentice-Hall.

Bandura, A., Ross, S. A., & Ross, D. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology, 63*(3), 575-582. doi: 10.1037/h0045925

Beck, M. R., Trenchard, M., van Lamsweerde, A., Goldstein, R. R., & Lohrenz, M. (2012, September). Searching in clutter: Visual attention strategies of expert pilots. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 56, No. 1, pp. 1411-1415). Santa Monica, USA: Sage.

Bertram, R., Kaakinen, J., Bensch, F., Helle, L., Lantto, E., Niemi, P., & Lundbom, N. (2016). Eye movements of radiologists reflect expertise in CT study interpretation: A potential tool to measure resident development. *Radiology*, *281*(3), 805-815. doi: 10.1148/radiol.2016151255

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin, 94*(2), 265-292. doi: 10.1037/0033-2909.94.2.265

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates, and novices. *Cognitive Science, 16*(2), 153-184. doi: 10.1207/s15516709cog1602_1

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293-332. doi: 10.1207/s1532690xci0804_2

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.

Cherubini, M., Nüssli, M. A., & Dillenbourg, P. (2010). This is it! Indicating and looking in collaborative work at distance. *Journal of Eye Movement Research*, *3*(5), 1-20. doi: 10.16910/jemr.3.5.3

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Cognition and instruction: Issues and gendas* (pp. 453-494). Mahwah: Erlbaum.

De Groot, A. D. (1946/2008). *Thought and choice in chess*. The Hague: Mouton.

Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research, 6*(2), 80-88. doi: 10.1177/1558689812437186

Dorr, M., Vig, E., Gegenfurtner, K. R., Martinetz, T., & Barth, E. (2008, October). Eye movement modelling and gaze guidance. Paper presented at the Fourth International Workshop on Human-Computer Conversation, Oxford, UK.

Dwyer, F. M. (1976). Adapting media attributes for effective learning. *Educational Technology, 16*(8), 7-13.

Ericsson, K. A., Charness, N., Feltovich, P., & Hoffman, R. R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. Cambridge: Cambridge University Press.

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science: Theoretical frameworks and research agendas

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211-245. doi: 10.1037/0033-295X.102.2.211

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363-406. doi: 10.1037/0033-295X.100.3.363

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaption to task constraints. *Annual Reviews in Psychology, 47*(1), 273-305. doi: 10.1146/annurev.psych.47.1.273

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge: MIT Press.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits in the empirical study of expertise. In K. A. Ericsson & J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits* (pp. 1-38). Cambridge, MA: Cambridge University Press.

Freyhof, H., Gruber, H., & Ziegler, A. (1992). Expertise and hierarchical knowledge representation in chess. *Psychological Research, 54*(1), 32-37. doi: 10.1007/BF01359221

Groner, R., & Siegenthaler, E. (2009). Improving the usability of eLearning tools: The IFeL Multifunctional Analysis and its application in distance teaching. *Proceedings of the ICDE/EADTU Conference in Maastricht*, June 2009. Retrieved from: http://www.ou.nl/Docs/Campagnes/ICDE2009/Papers/Final_Paper_100Groner.pdf

Gruber, H. (1991). *Wissensakquisition und Gedächtnisleistung in Abhängigkeit vom Expertisegrad.* Munich: University of Munich, Chair of Education and Educational Psychology.

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 1084-1102. doi: 10.1037//0278-7393.18.5.1084

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

Jaarsma, T., Boshuizen, H. P. A., Jarodzka, H., Nap, M., Verboon, P., & Van Merriënboer, J. J. G. (2016). Tracks to a medical diagnosis: expertise differences in visual problem solving. *Applied Cognitive Psychology, 30*(3), 314-322. doi: 10.1002/acp.3201

Jaarsma, T., Jarodzka, H., Nap, M., Van Merriënboer, J. J. G., & Boshuizen, H. P. A. (2014). Expertise differences under the microscope: processing histopathological slides. *Medical Education, 48*(3), 292-300. doi: 10.1111/medu.12385

Jaarsma, T., Jarodzka, H., Nap, M., Van Merriënboer, J. J. G., & Boshuizen, H. P. A. (2015). Expertise in clinical pathology: Bridging the gap. *Advances in Health Sciences Education, 20*(4), 1089-1106. doi: 10.1007/s10459-015-9589-x

Jacob, R., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573-605). Oxford: Elsevier.

Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye-movement modelling examples. *Instructional Science, 40*(5), 813-827. doi: 10.1007/s11251-012-9218-5

Jarodzka, H., Boshuizen, H. P. A., & Kirschner, P. A. (2012). Cognitive skills in medical diagnosis and intervention. In P. Lanzer (Ed.), *Catheter-based cardiovascular interventions; Knowledge-based approach* (pp. 69-86). Berlin, Germany: Springer.

Jarodzka, H., Gouw, P., Van Meeuwen, L., & Brand-Gruwel, S. (2015, August). *Air traffic control: Visual expertise in a dynamic problem solving task*. Paper presented at the 18th European Conference on Eye Movements, Vienna, Austria.

Jarodzka, H., Janssen, N., Kirschner, P. A., & Erkens, G. (2015). Avoiding split attention in computer-based testing: Is neglecting additional information facilitative? *British Journal of Educational Technology, 46*(4), 803-817. doi: 10.1111/bjet.12174

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146-154. doi: 10.1016/j.learninstruc.2009.02.019

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

Jarodzka, H., Van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction, 25*, 62-70. doi: 10.1016/j.learninstruc.2012.11.004

Jucks, R., Schulte-Löbbert, P., & Bromme, R. (2007). Supporting experts' written knowledge communication through reflective prompts on the use of specialist concepts. *Journal of Psychology, 215*(4), 237-247. doi: 10.1027/0044-3409.215.4.237

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. doi: 10.1207/s15326985ep4102_1

Kok, E., & Jarodzka, H. (2016). Before your very eyes: The value and limitations of eye tracking in medical education. *Medical Education, 51,* 114-122. doi: 10.1111/medu.13066

Kok, E., Jarodzka, H., De Bruin, A. B. H., BinAmir, H., Robben, S. G. F., & Van Merriënboer, J. J. G. (2015). Systematic viewing in radiology: Seeing more, missing less? *Advances in Health Science Education, 21*(1), 189-205. doi: 10.1007/s10459-015-9624-y

Krejtz, K., Duchowski, A. T., Krejtz, I., Kopacz, A., & Chrząstowski-Wachtel, P. (2016). Gaze transitions when learning with multimedia. *Journal of Eye Movement Research*, *9*(1), 1-17. doi: 10.16910/jemr.7.1.2

Lachner, A., Jarodzka, H., & Nückles, M. (2016). What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instructional Science, 44*(3), 197-203. doi: 10.1007/s11251-016-9376-y

Leff, D. R., James, D. R., Orihuela-Espina, F., Kwok, K.-W., Sun, L. W., Mylonas, G., . . . Yang, G.-Z. (2015). The impact of expert visual guidance on trainee visual search strategy, visual attention and motor skills. *Frontiers in Human Neuroscience, 9*. doi: 10.3389/fnhum.2015.00526

Lehmann, A. C., & Gruber, H. (2006). Music. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 457-470). Cambridge: Cambridge University Press.

Lilienthal, O. (1889). Der Vogelflug als Grundlage der Fliegekunst, Berlin, Germany.Retrieved from: https://commons.wikimedia.org/wiki/File:LilienthalFliegekunst.png

Lin, J. J., & Lin, S. S. (2014). Tracking eye movements when solving geometry problems with handwriting devices. *Journal of Eye Movement Research*, *7*(1), 1-15. doi: 10.16910/jemr.7.1.2

Lindow, S., Fuchs, H. M., Fürstenberg, A., Kleber, J., Schweppe, J., & Rummer, R. (2011). On the robustness of the modality effect: Attempting to replicate a basic finding. *Zeitschrift für Pädagogische Psychologie, 25*(4), 231-243. doi: 10.1024/1010-0652/a000049

Litchfield, D., & Ball, L. J. (2011). Using another's gaze as an explicit aid to insight problem solving. *The Quarterly Journal of Experimental Psychology, 64*(3), 649-656. doi: 10.1037/a0020082

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13*(2), 157-176. doi: 10.1016/S0959-4752(02)00018-X

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education, 60*(1), 95-109. doi: 10.1016/j.compedu.2012.07.011

Mayer, R. E. (2009). *Multimedia learning*. Cambridge: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43-52. doi: 10.1207/S15326985EP3801_6

McCabe, D., & Castel, A. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107*(1), 343-352. doi: 10.1016/j.cognition.2007.07.017

McNamara, A., Booth, T., Sridharan, S., Caffey, S., Grimm, C., & Bailey, R. (2012, August). *Directing gaze in narrative art.* Paper presented at the Proceedings of the ACM Symposium on Applied Perception, Los Angeles, USA.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*(4312), 75-78. doi: 10.1126/science.198.4312.75

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97. doi: 10.1037/h0043158

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

Nyström, M., & Holmqvist, K. (2008). Semantic override of low-level features in image viewing - both initially and overall. *Journal of Eye Movement Research, 2*, 1-11. doi: 10.1007/s11251-016-9397-6

Ögren, M., Nyström, M., & Jarodzka, H. (2016). There's more to the multimedia effect than meets the eye: Is seeing pictures believing? *Instructional Science,* 1-25. doi: 10.1007/s11251-016-9397-6

Oliva, M., Niehorster, D., Jarodzka, H., & Holmqvist, K. (in press). Influence of coactors on saccadic and manual responses. *i-Perception.*

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287. doi: 10.1037/h0084295

Penttinen, M., & Huovinen, E. (2011). The early development of sight-reading skills in adulthood: A study of eye movements. *Journal of Research in Music Education, 59*(2), 196-220. doi: 10.1177/0022429411405339

Penttinen, M., Huovinen, E., & Ylitalo, A.-K. (2013). Silent music reading: Amateur musicians' visual processing and descriptive skill. *Musicae Scientiae, 17*(2), 198-216. doi: 10.1177/1029864912474288

Penttinen, M., Huovinen, E., & Ylitalo, A.-K. (2015). Reading ahead: Adult music students' eye movements in temporally controlled performances of a children's song. *International Journal of Music Education, 33*(1), 36-50. doi: 10.1177/0255761413515813

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422. doi: 10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506. doi: 10.1080/17470210902816461

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science, 12*(1), 48-55. doi: 10.1111/1467-9280.00309

Reingold, E. M., & Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford handbook of eye movements* (pp. 523-550). Oxford: Oxford University Press.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29*(6), 1045-1060. doi: 10.1207/s15516709cog0000_29

Richardson, D. C., Street, C. N. H., Tan, J. Y. M., Kirkham, N. Z., Hoover, M., & Ghane-Cavanaugh, A. (2012). Joint perception: Gaze and social context. *Frontiers in Human Neuroscience, 6*, 194. doi: 10.3389/fnhum.2012.00194

Rummer, R., Schweppe, J., Fürstenberg, A., Scheiter, K., & Zindler, A. (2011). The perceptual basis of the modality effect in multimedia learning. *Journal of Experimental Psychology: Applied, 17*(2), 159-173. doi: 10.1037/a0023588

Schank, R. C., & Abelson, R. P. (2013). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale: Lawrence Erlbaum Associates.

Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice* (pp. 265-282). New York: Springer Verlag.

Schnotz, W., & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications* (pp. 304-356). New York: Cambridge University Press.

Skuballa, I. T., Fortunski, C., & Renkl, A. (2015). An eye movement pre-training fosters the comprehension of processes and functions in technical systems. *Frontiers in Psychology, 6*. doi: 10.3389/fpsyg.2015.00598

Skuballa, I. T., Schwonke, R., & Renkl, A. (2012). Learning from narrated animations with different support procedures: Working memory capacity matters. *Applied Cognitive Psychology, 26*(6), 840-847. doi: 10.1002/acp.2884

Sridharan, S., Bailey, R., McNamara, A., & Grimm, C. (2012, March). *Subtle gaze manipulation for improved mammography training.* Proceedings of the Symposium on Eye Tracking Research and Applications, Santa Barbara, USA.

Stofer, K., & Che, X. (2014). Comparing experts and novices on scaffolded data visualizations using eye-tracking. *Journal of Eye Movement Research, 7*(5), 1-15. doi: 10.16910/jemr.7.5.2

Journal of Eye Movement Research
10(1):3,1-18

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017).
Eye tracking in Educational Science:Theoretical frameworks and research agendas

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychological Review, 10*(3), 251-296. doi: 10.1023/A:1022193728205

Thurmond, V. A. (2001). The point of triangulation. *Journal of Nursing Scholarship, 33*(3), 253-258. doi: 10.1111/j.1547-5069.2001.00253.x

Van der Gijp, A., Ravesloot, C., Jarodzka, H., Van der Schaaf, M., Van der Schaaf, I., Van Schaik, J., & Ten Cate, T. J. (2016). How visual search relates to visual diagnostic performance: a narrative systematic review of eye-tracking research in radiology. *Advances in Health Sciences Education*, 1-23. doi: 10.1007/s10459-016-9698-1

Van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance (meta-)cognitive processes in computer-based learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 143-156). New York: Springer Science+Business Media.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior, 25*(3), 785-791. doi: 10.1016/j.chb.2009.02.007

Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2004). Process-oriented worked examples: Improving transfer performance through enhanced understanding. *Instructional Science, 32*(1/2), 83-98. doi: 10.1023/B:TRUC.0000021810.70784.b0

Van Gog, T., Paas, F., Van Merriënboer, J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*(4), 237-244. doi: 10.1037/1076-898X.11.4.237

Van Lehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47*, 513-539. doi: 10.1146/annurev.psych.47.1.513

Van Marlen, T., Van Wermeskerken, M., Jarodzka, H., & Van Gog, T. (in press). Showing a model's eye movements in examples does not improve performance of problem-solving tasks. *Computers in Human Behavior*. doi: 10.1016/j.chb.2016.08.041

Van Meeuwen, L. W., Jarodzka, H., Brand-Gruwel, S., Kirschner, P. A., De Bock, J. J. P. R., & Van Merriënboer, J. J. G. (2014). Identification of effective visual problem solving strategies in a complex visual domain. *Learning and Instruction, 32*, 10-21. doi: 10.1016/j.learninstruc.2014.01.004

Van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning*. Hillsdale: Erlbaum.

Wolff, C. E., Jarodzka, H., Van den Bogert, N., & Boshuizen, H. P. A. (2016). Teacher vision: Comparing expert and novice teachers' perception of problematic classroom management scenes. *Instructional Science, 44*(3), 243-265. doi: 10.1007/s11251-016-9367-z

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.