

Unboxing the black box of visual expertise in medicine

Halszka Jarodzka^{1,a,b} & Henny P.A. Boshuizen^{a,c}

^aWelten Institute – Research Centre for Learning, Teaching and Technology, Open University of the Netherlands

^bEye tracking Laboratory, Lund University, Sweden

^cSchool of Education, University of Turku, Finland

Abstract

Visual expertise in medicine has been a subject of research since many decades. Interestingly, it has been investigated from two little related fields, namely the field that focused mainly on the visual search aspects whilst ignoring higher-level cognitive processes involved in medical expertise, and the field that mainly focused on these higher-level cognitive processes largely ignoring the relevant visual aspects. Consequently, both research lines have traditionally used different methodologies. Recently, this gap is being increasingly closed and this special issue presents methods to investigate visual expertise in medicine from both research lines, namely those investigating vision (eye tracking, pupillometry, flash preview moving window paradigm), verbalisations, brain activity, and performance measures (ROC analysis, gesture coding, expert performance approach). We discuss the benefits and drawbacks of each method and suggest directions for future research that could help to unbox the black box of visual expertise in medicine.

Keywords: expertise; visual expertise; cognition; medicine

¹ Corresponding author: Halszka Jarodzka, Valkenburgerweg 177, 6419 AT Heerlen, The Netherlands, email: Halszka.Jarodzka@OU.nl. DOI: <http://dx.doi.org/10.14786/flr.v5i3.322>



1. Visual expertise in medicine as a research field

Expertise is known to be highly domain-specific (Chi, 2006). Hence, generalizations from findings cannot be made across different domains and often enough not even across tasks. Thus, the organizers of this special issue made a very important decision to focus on *medical* visual expertise. In this way, we can draw concrete conclusions from each contributing article to unbox the nature of visual expertise in medicine. The topic of visual expertise in medicine has received increasing attention over the past years (Gegenfurtner, Lehtinen, & Säljö, 2011; Kok & Jarodzka, 2016; Kok & Jarodzka, 2017; Norman, Coblenz, Brooks, & Babcock, 1992; Reingold & Sheridan, 2011; Van der Gijp et al., 2016). This comes as no surprise as most medical tasks require some sort of visual input, in the form of a medical image, a tissue sample, or the patient him- or herself. As a result, different theoretical models were constructed that describe different aspects of medical expertise. Holistic models (for recent descriptions, see Kundel, Nodine, Conant, & Weinstein, 2007; Nodine & Mello-Thoms, 2010) describe how experts visually search abnormalities on medical images. Another group of theories focus more on cognitive aspects of expertise and concretely describe the development of expertise (Boshuizen & Schmidt, 1992; Feltovich, Johnson, Moller, & Swanson, 1984; Norman, Young, & Brooks, 2007). Decades of research related to this group of models have provided us with a thorough idea of how expertise is constituted, however, most often only taking the cognitive component into consideration (e.g., ECG studies Gilhooly et al., 1997). Lesgold et al. (1988) tried decades ago to marry these two research lines into one unified model. Unfortunately, this model was not further developed ever since. Recently we have combined the model of Lesgold et al. (1988) with more recent cognitive expertise models (Boshuizen & Schmidt, 2008a) into one model presented in Figure 1 (Jarodzka, Boshuizen, & Kirschner, 2012). The fact that these types of models emerged from rather independent research fields, resulted in different types of research methodologies they use. For instance, while visual search models were often studied with eye tracking data and ROC abnormality detection rates, cognitive models were often studied with different forms of verbal data. However, over the past few years these differentiations do not hold any more and both research lines learn from each other. The current special issue presents these methodologies, irrespective in which research line they were originally used. This provides the ground towards a more unified theoretical model of visual expertise in medicine which will ultimately unbox the black box of visual expertise in medicine.

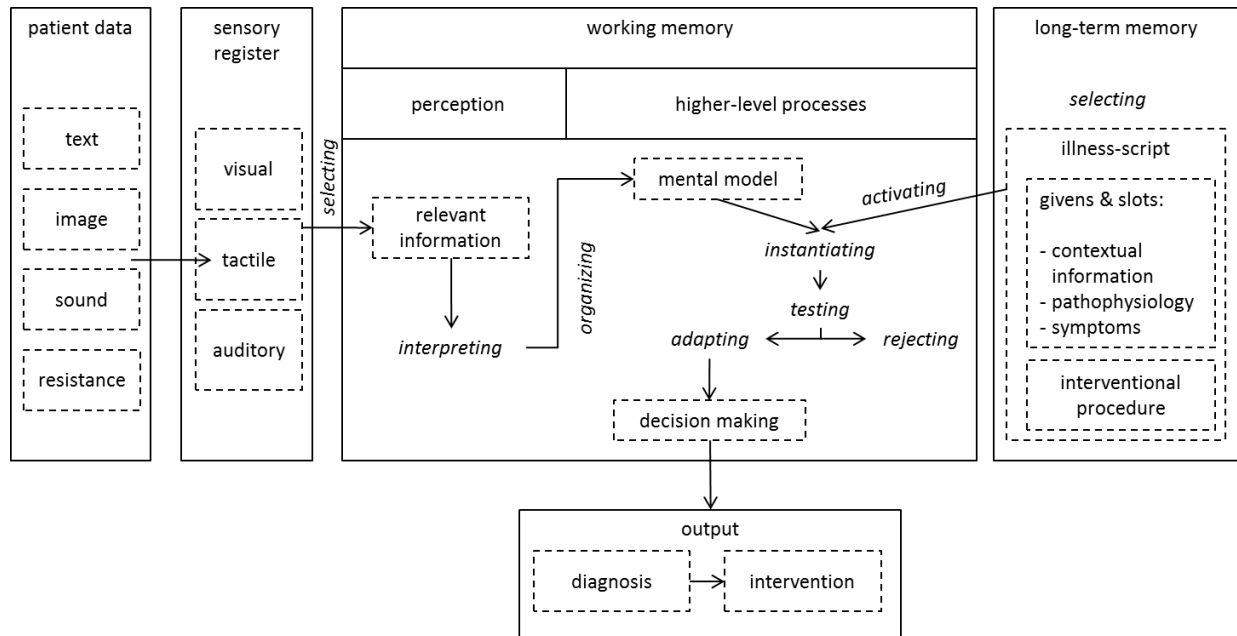
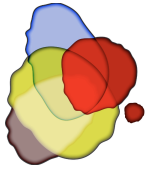


Figure 1. Theoretical model combining the classical model of Lesgold et al (1988) with recent cognitive models of medical expertise (Boshuizen & Schmidt, 2008) as published in Jarodzka, Boshuizen, & Kirschner (2012).

2. Methods presented in the current special issue and the concepts they address

This special issue brings together diverse methods that all aim at “unboxing the black box” of medical expertise from different angles. We chose to structure them according to the concepts they are mainly focusing at (Figure 2). Our structure corresponds to the structuring as Gegenfurtner and Van Merriënboer (2017) use it in their introduction to the current special issue into activation (brain activity), detection (vision), inference (verbalisation), and practice (observations).

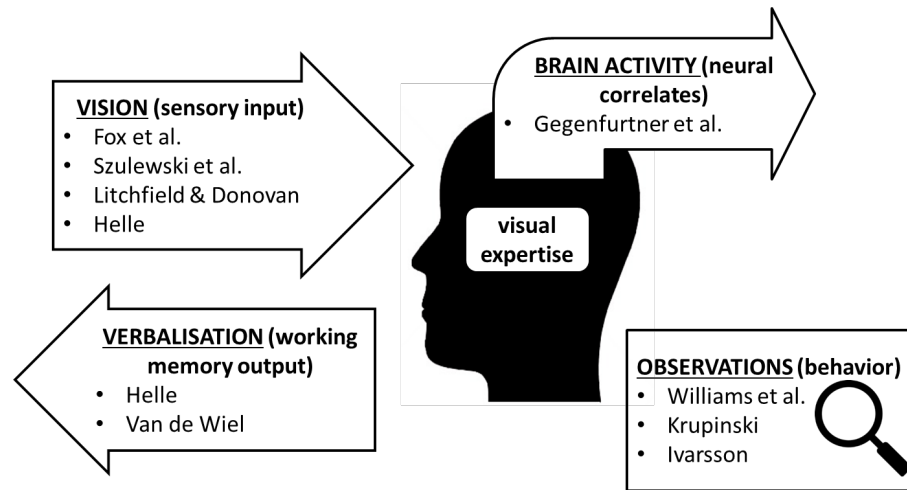
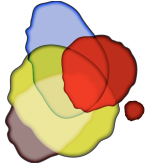


Figure 2. Methods to capture diverse concepts related to visual expertise presented in the current special issue.

2.1 Vision – the sensory input

These articles present methods to measure the sensory input of the medical specialist. How can measuring sensory input help us understanding (visual) expertise? Efficient information-processing is a central part of expertise and its development (Reingold & Sheridan, 2011). On the one hand, an expert is able to detect subtle cues and interpret them within a certain context, but s/he also can detect patterns within seeming random elements. On the other hand, our information-processing system is not a passive recipient, but rather in active search for meaningful information. The theory of Boshuizen and Schmidt (2008b) gives indications on how this process unfolds: information elements that enter the cognitive systems of novices and intermediates activate nodes within large knowledge networks – depending on their experience, this happens more or less efficient. Experts, on the other hand, also begin with a passive reception of information elements. These, however, instantly activate one or several illness-scripts, which in turn, guide the further active search for information (Jarodzka, Boshuizen, et al., 2012). Hence, the passive intake or the active search of (visual) information elements has the potential to reveal crucial aspects of a person’s expertise. The current special issue, provides three manuscripts that discuss methodology to do address this aspect of expertise.

2.1.1 Eye tracking

The idea that medical experts “see more” than untrained individuals do, appears immediately, when you see a medical expert “reading” from an X-ray or a mammogram. Hence, obviously medical expertise was already very early investigated with eye tracking (for a comprehensive overview, see Reingold & Sheridan, 2011). Eye tracking (Holmqvist et al., 2011) entails (1) the apparatus that measures the motion of the eye balls in relation to a stimulus, (2) the software that allows to relate parameters derived from the eye movements to certain parts of the stimulus (in time or space), and (3) the researchers, who interpret these findings within a theoretical framework. The clear benefits of this methodology are that it unobtrusively captures unconscious processes directly; it captures all relevant visual input to working memory. On the



other hand, eye tracking data are ambiguous, task-dependent, idiosyncractic, and last but not least challenging in data collection and analysis.

The article by Fox and Faulkner-Jones (2017) provides a brief historical overview of eye tracking. For a broader view, we would like to refer the reader to the informative and entertaining book by Wade and Tatler (2005) on the history of eye tracking. We applaud Fox and Faulkner-Jones for their excellent analysis of the different medical tasks and how they should be differently approached by means of eye tracking. We fully agree with them, in particular, as it is well-known how task-dependent expertise is and how broad the field of medicine is, at the same time. We hope, that this detailed analysis will forgo overly generalized statements, such as one these authors surprisingly made themselves “eye-tracking studies across medical specialties have suggested that more experienced physicians require fewer fixations, and less time spent on areas of interest, [...] than novices.” (p.3). Such statements are not only too reductionistic to reveal interesting insights into the nature of expertise, but even worse: they are often enough simply wrong. In many medical areas, we find exactly the opposite to be true, namely that experts are looking longer at relevant areas of interest (e.g., Balslev et al., 2012). That does not mean that studies finding the one or the other were wrong; it means that these findings cannot be generalized, but depend on the exact task and the stimuli that were used.

What many of the here reported eye tracking studies in medical expertise “suffer” from, is, that they report eye tracking measures, that are too basic to allow drawing conclusions on the nature of expertise. One example are the findings reported from Fox, Law, and Faulkner-Jones (2016) that trainees make more eye movements than experts. By itself, this statement comes down to a simple time-on-task difference, that does not make use of the potentials that eye tracking as a methodology offers. One example of how to make more concrete statements from eye tracking more concrete is a study by Kok, De Bruin, Robben, and Van Merriënboer (2012). These authors have investigated how experts, in comparison to medical residents and students, visually explored focal and diffused diseases on chest X-rays. Amongst others, they calculated a measure that captured, how broadly an image was scrutinized, by calculating the global/local ratio of saccades. In this way, the authors showed that images containing a diffuse disease (i.e., a disease that is spread all over the lungs and cannot be brought down to one location) were visually examined in a broader way, by inducing a higher global/local ration. Similarly, in one of our own studies (Jaarsma, Jarodzka, Nap, Van Merriënboer, & Boshuizen, 2014), we have investigated how expert pathologists, pathology residents, and medical students visually examine pathological slides. Amongst other, we found that experts and residents diagnosed the slides equally well. However, the way they processed the slides differed severely: while experts looked immediately to the relevant location and scrutinized it with long fixations, they had afterwards time to explore the rest of the slide – with short fixations – for other potential abnormalities. Residents on the other hand, took quite some time to detect the relevant location, and they examined it up to the end of the trial to verify their diagnosis. Hence, experts had capacities left over in this task, while residents were at their maximum. This could have made a difference for more complex cases, for instance, with several different diseases in one case. Hence, the potentials of eye tracking can be explored much more by going beyond the ‘standard’ eye tracking measures and looking more concretely into the characteristics of the task and the stimulus at hand.

Another important aspect that Fox and Faulkner-Jones (2017) point out is the lack in research on 3D and dynamic medical images. We fully agree with that, but would like to point towards research not



mentioned by these authors, i.e., by Bertram, Helle, Kaakinen, and Svedström (2013) on CT images and our own research on interactive digital pathology slides (Jaarsma et al., 2016; Jaarsma, Jarodzka, Nap, Van Merriënboer, & Boshuizen, 2015; Jaarsma et al., 2014) and on patient-video cases (Balslev et al., 2012). Moreover, the authors mention on several occasions the potential eye tracking has for medical *education*. We agree. On that note, we would like to point towards the idea of using eye movements of experts in instructional videos (Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009) and its successful application to medicine (e.g., Jarodzka, Balslev, et al., 2012). On a final note, it is important to mention that Fox and Faulkner-Jones (2017) have discussed many visual search models, but did not make the connection to current models on medical expertise and its development (Jarodzka, Boshuizen, et al., 2012). However, as we already argue elsewhere (Kok & Jarodzka, 2016), this connection is crucial for theoretical development.

2.1.2 Pupillometry

Szulewski, Kelton, and Howes (2017) describe pupillometry as a method to capture cognitive load in relation to visual expertise in medicine. This idea is very persuasive as it would allow to unobtrusively measure online cognitive processes during medical task performance. Pupillometry is actually the use of one very specific measure derived from eye tracking equipment, namely the size of the pupil and how it changes over time. We have to keep in mind though, that the main purpose of changes in the size of the pupil, is the adaptation of the eyes' photoreceptors to the lighting conditions to allow for optimal vision (Davson, 2012). This is a reflex that everyone can easily observe: look close to a mirror in a brightly lighted room. Now close one of your eyes off with your hand. Open this eye after a moment and observe how your two eyes differ: the eye that was open all the time – and thus exposed to light – has a small pupil. The other eye, the one that was exposed to relative darkness has a larger pupil. This, however, changes very quickly and you can observe how your pupil shrinks to adapt your sight to the changed lightning conditions.

Research has shown that the size of the pupil may also change if lighting conditions are stable. It may vary according to the interest a participant shows in a stimulus (Hess & Polt, 1960), their emotional state (Vanderhasselt, Remue, Ng, & De Raedt, 2014), the musical chill they experience (Laeng, Eidet, Sulutvedt, & Panksepp, 2016) and many other exciting concepts (Loewenfeld, 1999). This measure can even be a valid indicator for certain diseases, such as Parkinson's (Wang, McInnis, Brien, Pari, & Munoz, 2016). On the other hand, pupillometry is a rather coarse measure that often cannot provide specific predictions (failure in predicting sexual orientation: Savin-Williams, Cash, McCormack, & Rieger, 2016). In any case, it is important to keep in mind, that all these changes in pupil size are subtle and can be easily overruled by a ray of light falling onto the eye. It is therefore important, to keep meticulously equal lightning conditions for both eyes over the entire experiment. This not only holds for the laboratory room, but even the stimulus presentation screen luminosity has to be kept stable. These are conditions that can be realized in fundamental laboratory research, but are difficult to realize in applied research, such as medical education. Hence, we will often have to wonder whether the pupil size changes were due to the emotional state of the participants, their mental effort etc. or rather the inevitable changes in the lightning falling onto their eyes in this particular stimulus or their position in the recording room.

Szulewski et al. (2017) also discuss these and other severe drawbacks of using pupillometry in real-life settings, still they come to an optimistic conclusion that this method would have “a particularly promising role in the field of medicine and in the study of physician expertise development”. We would rather suggest to address these methodological problems by triangulating pupillometry with other mental



effort measures that are less vulnerable in real-life settings, for instance, questionnaires (e.g., Paas, 1992), dual-task paradigms (Brünken, Steinbacher, Plass, & Leutner, 2002) or other less vulnerable physiological data, such as skin resistance (e.g., Nourbakhsh, Wang, Chen, & Calvo, 2012).

2.1.3 *The flash-preview moving-window (FPMW) paradigm*

The manuscript by Litchfield and Donovan (2017) presents the ‘moving window paradigm’ to investigate visual expertise. McConkie and Rayner (1975) introduced this paradigm to investigate the so-called perceptual span in reading. Based on the current fixation within a word, a few characters to the left and to the right are masked to investigate to which extent this influences reading. The underlying idea is, that we move our eyes in one direction when reading a text (often from left to right or vice versa, depending on the language), and that the amount of information we can take in towards this direction, without fixating it, increases with increasing expertise in reading. As the visual processing of a written text is so clearly predefined, we know exactly how our eyes will move on a line. Unsurprisingly, McConkie and Rayner (and many more afterwards confirmed and specified this) showed that our perceptual span in reading is skewed to the right and indeed depends on our expertise. This is a method widely used and well-established in reading research. When looking at medical pictures, however, there is no such clearly predefined gazing direction as in reading (e.g., line by line, from left to right). Hence, the methodological set-up is more complicated. Typically, in such ‘scene perception’ settings, researchers use a Gaussian blurring technique to capture the size of the perceptual span (i.e., the image is blurred apart from a certain area around the current fixation point). The “flash preview moving window paradigm” (Litchfield & Donovan, 2017) shows an alternative solution. It combines a method of providing participants only with a short glimpse of an image (‘flash preview’: Kundel & Nodine, 1975) and the above described ‘moving window paradigm’.

The really clever aspect about using this method to investigate visual expertise in medicine, is, that it allows estimating (1) to which extent the pre-activation of a schema based on visual input influences, how a consecutive visual search is carried out, and (2) the exact size of the perceptual span in relation to expertise level. The latter shows that – at least in other domains – the size of the visual span increases with higher levels of expertise. More interesting is the first point, though: to which extent, does an initial schema activation guide the actual search for information relevant to this schema? Litchfield and Donovan could not find strong empirical evidence for this idea (based on the model of Kundel, Nodine, & Toto, 1991). This comes as no surprise, when consulting cognitive theories on medical expertise (as summarized in Jarodzka, Boshuizen, et al., 2012). These theories assume that medical experts activate *a set* of schemata that are – partly – instantiated, tested and often discarded. Hence, from these theories we would assume that this is rather an ongoing process than a strictly serial one. What would be very interesting for future research, is to take these cognitive theories on medical expertise to inform FPMW research in medicine. Considering how a medical expert would pursue in the real world aligns well with these cognitive theories (Jarodzka, Boshuizen, et al., 2012) and would be very interesting in at least two ways. First, in a real-life situation the expert would first receive several background information of the patient, which would activate an illness-script (more complex than a schema, see also Figure 1). This activated illness-script would already guide the expert in his or her subsequent visual search on the medical image. Interestingly, Litchfield and Donovan (2017) did something in this direction within their third experiment, by showing the target word to the participant right before the flash-preview. It would be very informative for further theoretical developments to extend this approach, by using realistic patient data. Second, in real-life, the task of the medical expert is



not to simply state the presence or location of a target, but goes far beyond: providing a diagnosis, requesting further examinations of the patient, and finally suggesting a treatment. Including these aspects into FPMW experiments, would allow seeing not only a potential influence on the visual search itself, but also on its accompanying cognitive processes.

2.2 Verbalisations – the working memory output

From a cognitivist perspective, medical expertise development has been mainly investigated using verbal methods; even for those domains, that heavily draw on visual skill (e.g., radiology Lesgold et al., 1988 and ECG interpretation Gilhooly et al., 1997). At the same time, the visual challenges of those fields were not much in focus within this paradigm, as it was felt, that verbal methods lacked the acuity to discern and untangle perceptual processes. Van de Wiel's article (2017) does a great job in showing how verbal methods can be used in showing reasoning lines and knowledge application by medical experts, intermediates, and novices. It also shows that the validity of verbal methods depends a lot on the vocabulary mastered by the participants, the skill of the researcher to identify references to visual qualities (Helle (2017) refers to Ericsson's (2006) "non-verbal thoughts", captured by briefs labels and referents), and on the successfulness of separating perception of features of the image (if necessary by means of additional methods such as pointing or drawing) from the interpretation of patterns of features in the protocols. The usefulness of verbal methods, stand-alone or in combination with eye-tracking, also depends on a couple of other things: the kind of visual information involved, and – not surprisingly – the research question.

Even when we constrain ourselves to (stacks of) pictures resulting from medical imaging techniques (e.g., EEG, ECG, X-ray images, microscopic pathological slides), the differences in the amount of information embedded in such pictures are huge. A basic ECG consists of one wiggly line that should show a repetitious pattern with several features such as a P, R and T-top and associated Q and S-inflexions with their specific amplitude and latency. Absence of these features, and irregularities of the pattern may have clinical meaning. Importantly, students learn the interpretation of these visual presentations as a combination of features of the visual appearance, the associated vocabulary, and the biomedical and clinical interpretation. Though visualising complex phenomena, an ECG is a simple line, though more advanced equipment can visualise several measurements simultaneously (up to twelve for complex diagnoses). A similar analysis applies to EEG, though the number of channels recorded is much higher. The potential amount of information in X-rays pictures, fMRI-s, PET-scan and microscopic images is much higher than in line graphs. They are (stacks of) 2D, multi-coloured or grayscale pictures that may maximally vary pixel by pixel, independent of the colour or the greyness of an adjacent pixels. A final difference is related to the nature of certain diseases that may be focal or diffuse. Images of local conditions may show isolated, discernible lesions that can be pointed at, but other disease processes only show themselves in the qualities of the image (e.g., 'cloudy' or 'milky'; see Kok et al., 2012). These differences in visual qualities of the domain under investigation have implications for vocabulary building, and thus for the usefulness of verbal reports generated by participants of different expertise levels (see Van de Wiel, 2017), and for foveal detection, and thus for the usefulness of eye-tracking (see Helle, 2017).

It is almost a platitude to state that the research question affects the investigative data collection and analysis methods to be used. Yet, the articles by van de Wiel and by Helle forget to problematize the assumption that feature extraction should be differentiated from pattern recognition and interpretation (van



de Wiel), and whether deep (meaning) coding of think-aloud protocols is better than superficial coding that stays close to the exact verbalisations. There are strong indications that visually detecting and interpreting relevant visual features among irrelevant ones, goes hand in hand with interpretation, and is guided by a person's expectations. Many interpretative choices in visual information processing seem to take place in the early and non-analytic phase (Norman et al., 2007). On the other hand, recognition of even obvious features (for instance, a clear-cut jaundice) is a gradual process that is interlaced with the developing hypotheses about what might be wrong. Manipulation of these hypotheses can lead to the non-recognition of such features (Brooks, LeBlanc, & Norman, 2000; LeBlanc, Brooks, & Norman, 2002; LeBlanc, Dore, Norman, & Brooks, 2004; LeBlanc, Norman, & Brooks, 2001). The alleged superficiality of our own analyses of verbal protocols turned out to reveal aspects of perception and cognition (Jaarsma et al., 2014) that we have never become aware of in earlier research that made use of deep, semantic analyses (see for instance, Boshuizen & Schmidt, 1992). Lack of a professional vocabulary in novice groups seems to be associated with a lack of a repertoire of perceptual features in that domain thus hampering perception and interpretation. Stated differently, what might be interpreted as 'poor' protocols, may be a veridical representation of the perceptual and cognitive skills of the participant. A decision pro or con one or the other interpretation cannot be made just on procedural features of the way the research method was applied, but requires the theory-based assessment of expertise level, image quality and research question.

2.3 Brain activity – neural activity and blood flow

Measuring the neural activity of the brain, gives us the most 'pure' look inside the brain. It is very persuasive to believe that one day we will be able to observe the brain of experts while they are performing a task of their domain and observe how their brilliance unfolds under our very eyes. However, we are not quite there, yet, and the question is whether we ever will – or ever *need* to be. As fascinating as fancy new visualizations of activities in the brain might be, we have to keep in mind what they represent: increased blood flow in some regions of the brain in comparison to others (fMRI: Huettel, Song, & McCarthy, 2008) or increased neural activity *somewhere* in the brain (EEG: Niedermeyer & Da Silva, 2005). Thus, we can observe and record some physiological processes either specific over time (EEG) or space (fMRI). What we cannot tell, though, is which thoughts these activities represent. We need to keep that in mind, when estimating the insights we can derive from such techniques.

Imagine lying on a narrow stretcher, while your head is tied within a small cage-like object. You are asked not to move and are left alone in the room. Then the stretcher begins to move into a tube which makes strange noises. First you might feel very scared (many people do!) and later on coming close to fall asleep (also quite common). This is what it feels like when you participate in an fMRI scan. From this we can immediately tell, that this is an extremely fundamental laboratory study. The reason for this extremely restricted position for the participant is that even the slightest movements (even blinks and eye movements) can cause neural activity that was not induced by the experimental intervention. That is also the reason for the very many repetitive trials the experimenter has to run on one participant (to filter this noise out). The situations are similar for other measures of neural activity.

When talking about visual *expertise* research, this seems to be an extremely reductionistic approach. This has to do with two issues: first, expertise is a result of decades of deliberate practice and only measures under representative circumstances (e.g., Ericsson & Lehmann, 1996). Obviously, the scenario described



above does not represent any form of visual expertise in nowadays medical expertise. A serious problem is that expertise is extremely domain- and task-specific and thus evolves only under the very specific circumstances of the very task. Drawing conclusions on expertise from pseudo-expertise tasks (note that studying artificial objects even for several sessions does not qualify for the definition of real expertise), will never be possible. Second, when measuring neural activity or blood flow, do we really do justice to the nature of expertise in its complexity? For these and other more pragmatical reasons (costs of conducting such research), most of the research presented in the article by Gegenfurtner, Kok, Van Geel, De Bruin, and Sorger (2017) is actually not related to medical expertise. Actually, only three of the presented studies investigated medical expertise as they involved real, medical tasks (Fiorio, Cesari, Bresciani, & Tinazzi, 2010; Melo et al., 2011; Ribas, Rocha, Siqueira Ortega, Freitas de Rocha, & Massad, 2013). Interestingly, first studies show activation differences in relation to medical expertise – at least for certain stimuli (Hruska et al., 2016). The challenge remains to understand what these found differences actually mean in terms of expertise development. Hence, as interesting as these studies are, it is difficult to draw concrete conclusions from them already and clearly far more research is needed.

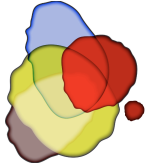
So, do we argue that such medical expertise research is pointless? On the contrary! But we need to be very careful, what it can be used for. We argue that such research on neurological processes can inform fundamental research on memory and attention, which in turn can inform cognitive science, which forms a basis for (medical) expertise research. On that route, it could even eventually reach educational research. What is now urgently needed to make this information flow possible are solid theoretical models that allow for these connections between these research fields.

2.4 Observations – behavioural performance

One key step in expertise research is to estimate whether participants are ‘real’ experts by checking whether their performance systematically exceeds the one of individuals with less expertise (Ericsson & Lehmann, 1996; Ericsson & Smith, 1991). This is easier for some domains than it is for others (e.g., chess expertise can be clearly defined by the ELO system). For visual expertise in medicine performance estimates are not trivial as Krupinski (2017) describes in her article. In this section, we review three articles that capture very different aspects of performance related to visual expertise in medicine.

2.4.1 ROC analysis

Krupinski (2017) describes the well-established method of Receiver Operating Characteristics (ROC) analysis to tackle this issue. This analysis method allows scrutinizing the ability of medical specialists or lay persons to detect one abnormality in a medical image in a very detailed manner. This detailed statistical analysis allows for clear interpretations of the findings. This article provides a very comprehensive and concrete ‘hands-on’ on how to conduct this specific methodology. Such an article is of extreme practical value for other researchers. Unfortunately, such publications are still rare, even though more of these would be needed. This holds even true for the current special issue; the article by Krupinski (2017) has the highest practical value for other researchers interested in the topic of visual expertise in medicine – and not only! Many other domains of visual expertise could benefit from this approach, too. As long as the task can be boiled down to a binary, exclusive decision.



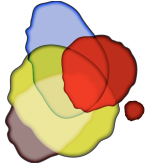
This is also where the drawbacks of the ROC method begin. Even though, it is very well-established as a sophisticated statistical method, it is only applicable for a very limited type of task, namely the binary, exclusive decision. However, visual expertise in medicine goes far beyond that, as the author mentions already in the very beginning of her article. Already the detection itself goes beyond this simplified present vs. not-present decision: the medical specialist needs to know WHERE the abnormality is located, whether other ones are present as well, etc. More recent types of ROC analyses take these issues into account (LROC, JAFROC, etc.) and have been used for many years already. Although it is important to understand ROC first, similar articles to this one on more up-to-date versions of ROC would be very important. But still, visual expertise in medicine goes beyond the mere detection of abnormalities. It is only the very first step and does not capture the entirety of medical expertise performance (e.g., the following diagnostic or treatment decision). The next question is thus, whether it is possible to extend these methods to other forms of performance as well.

On a final note, the author explains how the specific ROC curve and its interpretation depend on the observer's background and experience. However, she does not draw the connection to existing and well-established theories on medical expertise and its development (for a summarized model of them, see Jarodzka, Boshuizen, et al., 2012). For this method to be able to further contribute to more general (medical) expertise research, this connection to existing cognitive theories on medical expertise (development) is urgently needed and should be the next step in research line.

2.4.2 *Gesturing*

The study by Ivarsson (2017) investigates a different observable aspect of medical (visual) expertise, namely gestures. As described above, verbalising visual process is a rather difficult endeavour for which gestures can be really helpful. The article by Ivarsson shows very concrete examples that depict exactly this. What is important to know is that this study – in contrast to all other articles in this special issue – does not deal with the diagnosis of an individual expert. Instead, it acknowledges that a lot of medical practice is carried out in groups. This fact already makes this article a unique contribution to the current special issue. Ivarsson investigates a communication situation between several medical professionals. In such a scenario, communication comes into play, of which non-verbal aspects are a crucial part of, especially in the medical profession as this article shows. It might be interesting to know that even though the author mainly addresses movements of the limbs, there seems to be another body part involved as well: examples [21] and [22] indicate that the professional was not only gesturing with her hands, but also guiding the listeners' attention with her gaze. This is a well-studied and important phenomenon, and following the gazes of others strongly guides our attention (Anstis, Mayhew, & Morley, 1969) in particular in conversations (Argyle & Cook, 1976; Mansfield, Farroni, & Johnson, 2003). Hence, gaze guidance could be explicitly included in such an analysis.

An important question within this research line is what the exact function of the gesture is. In the case described in this article, the main function of the gestures used is to establish a common ground between the professionals in a discourse. But is this really the sole purpose of these gestures? Are they merely communicative or are they an inherent part of a schema? Research in eye tracking has shown, that people often make eye movements even when these are not providing any information, such as when looking at a blank screen or being in the dark (Foulsham et al., 2012; Johansson & Johansson, 2014). In these cases, participants automatically move their eyes when remembering a prior encoded scene. If they are forced not



to move their eyes, their recoding performance significantly drops. Hence, these eye movements (also a form of body motion) have a functional role in restoring long term memory content. Couldn't the same be true – at least in part – for gestures?

Ivarsson (2017) chose a rather atypical task for these experts. The situation is very specific and the episode rather short. So, what can we learn from that? The purpose of this endeavour can only be hypothesis building and further research needs to follow to *test* these hypotheses. These might, for instance, investigate whether the here found gestures are typically used by these professionals? For instance, the ones indicating the representations of digital manipulations [7-9]. Is there a gesture 'language' that professional use? A different way to apply such methodology would be to investigate individual professionals. For instance, future research could investigate gestures that are part of the clinical routine (e.g., surgery, but also a radiologist turning the X-ray upside down or holding it in a particular angle) or are used as preparations for the clinical routine (a phenomenon that often can be observed in sports). These analyses could be triangulated with other data. In our own research, we have also investigated the interplay between hand movements as navigations within a pathological digital slide, eye movements on this slide as well as the verbalization about this examination (Jaarsma et al., 2016; Jaarsma et al., 2015; Jaarsma et al., 2014). This approach was also already investigated in a more natural setting with mobile eye tracking, although in the non-medical task of tea making (Tatler et al., 2013). Such a triangulated analysis of eye-hand coordination could be a very meaningful addition to the analysis of gestures.

2.4.3 *The expert performance approach*

Williams, Fawver, and Hodges (2017) provide an excellent overview of methodologies of research on expertise. They describe three steps, namely (1) developing representative tasks that elicit systematic performance differences in individuals of different levels of expertise, (2) process-tracing techniques to study processes underlying expert performance and (3) identifying individual characteristics in training or learning that lead to the expert level. We would like to add to this list the importance of a thorough definition and description of different expertise levels, which is described very concretely in the medical domain by Boshuizen and Schmidt (2008b).

What Williams and colleagues (2017) then mainly focus on, is studying the learning towards expertise. In principle, this could be done from two perspectives. On the one hand, one can identify successful methods that improve performance towards visual expertise. One example of an instructional method to train aspects of visual expertise are eye movement modelling examples (Van Gog et al., 2009). These are instructional videos showing how an expert model approaches a task. Therefore, the model verbally explains the steps taken in this task. Moreover, the attentional focus of the expert (based on his or her eye movements) is overlaid on this video. We have already successfully applied this method in the medical domain (Jarodzka, Balslev, et al., 2012). We showed that diagnosing patient-video cases improved after such a training not only in terms of performance, but also on the visual processes. Williams and colleagues (2017) favour another approach: the investigation of individual learning trajectories to identify 'good' and 'poor' learners. Based on such an analysis they argue, not only instruction could be improved, but also the identification of future experts might be possible. A long-standing research line of self-regulated learning in non-medical professions could be very informative for such a research (Kicken, Brand-Gruwel, Van Merriënboer, & Slot, 2009). In any case, Williams et al. – and we fully agree with them – call for two important issues: (a) more longitudinal studies to truly understand the development of visual expertise in



medicine, as well as (b) a process-tracing approach to identify relevant (and probably heterogeneous) processes underlying this development.

3. Lessons learned and future avenues

This special issue presented different methods to investigate different aspects of visual expertise in medicine emerging from different research lines within this field. To reach the aim of unboxing the black box of visual expertise in medicine, we argue that future research in this line, should consider the following points:

- Methodological triangulation: we saw that each of the methods presented in this special issue have unique potentials, but also severe drawbacks. To counterbalance these drawbacks, we need to use more methodological triangulation of different methods when conducting studies within visual expertise in medicine (see also: Gegenfurtner et al., 2016; Kok & Jarodzka, 2016).
- We need to systematically discuss more on the challenges we face with new methodologies and detailed process measures. Unfortunately, traditional empirical articles leave hardly room to do this. We thus plead for a forum where such issues could be identified, evaluated and solutions to them agreed upon.
- Several articles in this special issue have shown the importance of more interdisciplinary research that combine different research fields, such as medical image perception and scene processing.
- Finally, as uttered many times throughout this discussion, we need more solid theoretical models that allow to form bridges between these different methodologies presented in this special issue (see also: Gegenfurtner et al., 2016; Kok & Jarodzka, 2016).

References

- Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television portrait is looking. *American Journal of Psychology*, 82, 474-489.
- Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Oxford, England: Cambridge University Press.
- 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
- Bertram, R., Helle, L., Kaakinen, J. K., & Svedström, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PLoS ONE*, 8(6), 1-15. doi: 10.1371/journal.pone.0066169
- 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
- Boshuizen, H. P. A., & Schmidt, H. G. (2008a). The development of clinical reasoning expertise. In J. Higgs, M. Jones, S. Loftus & N. Christensen (Eds.), *Clinical reasoning in the health professions* (pp. 113-122). Oxford: Butterworth Heinemann / Elsevier.
- Boshuizen, H. P. A., & Schmidt, H. G. (2008b). The development of clinical reasoning expertise: Implications for teaching. In J. Higgs, M. Jones, S. Loftus & N. Christensen (Eds.), *Clinical reasoning in the health professions* (Vol. 3rd rev. ed. ed., pp. 131-121). Oxford, UK: Butterworth-Heinemann/Elsevier.



- Brooks, L. R., LeBlanc, V. R., & Norman, G. R. (2000). On the difficulty of noticing obvious features in patient appearance. *Psychological Science, 11*, 112-117. doi: 10.1111/1467-9280.00225
- Brünken, R., Steinbacher, S., Plass, J., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology, 49*, 109-119. doi: 10.1027//1618-3169.49.2.109
- Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, R. R. Hoffman & P. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21-30). Cambridge: Cambridge University Press.
- Davson, H. (2012). *Physiology of the Eye*: Elsevier.
- Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 223-241). Cambridge: Cambridge University Press.
- Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Reviews in Psychology, 47*(1), 273-305. doi: 10.1146/annurev.psych.47.1.273
- 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.
- Feltovich, P. J., Johnson, P. E., Moller, J. H., & Swanson, D. B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. In W. J. Clancey & E. H. Shortliffe (Eds.), *Readings in medical artificial intelligence: the first decade* (pp. 275-319). Reading, MA: Addison-Wesley Publishing Company.
- Fiorio, M., Cesari, P., Bresciani, M. C., & Tinazzi, M. (2010). Expertise with pathological actions modulates a viewer's motor system. *Neuroscience, 167*, 691-699. doi: 10.1016/j.neuroscience.2010.02.010
- Foulsham, T., Dewhurst, R., Nyström, M., Jarodzka, H., Johansson, R., Underwood, G., & Holmqvist, K. (2012). Comparing scanpaths during scene encoding and recognition: A multidimensional approach. *Journal of Eye Movement Research, 5*(5), 1-12. doi: 10.16910/jemr.5.4.3
- Fox, S. E., & Faulkner-Jones, B. E. (2017). Eye-tracking in the study of visual expertise: methodology and approaches in medicine. *Frontline Learning Research, 5*(3), 43 - 54. doi: 10.14786/flr.v5i3.258
- Fox, S. E., Law, C. C., & Faulkner-Jones, B. E. (2016). Quantitative gaze assessment of a dual "side by side" viewer versus a single whole slide image viewer for pathology education. *Journal of Pathology Informatics*.
- Gegenfurtner, A., Kok, E., Van Geel, K., De Bruin, A. B. H., Jarodzka, H., Szulewski, A., & V., V. M. J. J. G. (2016). The challenges of studying visual expertise in medical image diagnosis. *Medical Education, 51*(1), 97-104. doi: 10.1111/medu.13205
- Gegenfurtner, A., Kok, E., Van Geel, K., De Bruin, A. B. H., & Sorger, B. (2017). Neural correlates of visual perceptual expertise: Evidence from cognitive neuroscience using functional neuroimaging. *Frontline Learning Research, 5*(3), 14 - 30. doi: 10.14786/flr.v5i3.259
- Gegenfurtner, A., Lehtinen, E., & 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. doi: 10.1007/s10648-011-9174-7
- Gegenfurtner, A., & Van Merriënboer, J. (2017). Methodologies for studying visual expertise. *Frontline Learning Research, 5*(3), 1-13. doi: 10.14786/flr.v5i3.361
- Gilhooly, K. J., McGeorge, P., Hunter, J., Rawles, J. M., Kirby, I. K., Green, C., & Wynn, V. (1997). Biomedical knowledge in diagnostic thinking: The case of electrocardiogram (ECG) interpretation. *European Journal of Cognitive Psychology, 9*(2), 199-223. doi: 10.1080/713752555
- Helle, L. (2017). Prospects and pitfalls in combining eye-tracking data and verbal reports. *Frontline Learning Research, 5*(3), 81 - 93. doi: 10.14786/flr.v5i3.254



- Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science*, *132*(3423), 349-350. doi: 10.1126/science.132.3423.349
- 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.
- Hruska, P., Hecker, K. G., Coderre, S., McLaughlin, K., Cortese, F., Doig, C., . . . Krigolson, O. (2016). Hemispheric activation differences in novice and expert clinicians during clinical decision making. *Advances in Health Science Education*, *21*(5), 921–933. doi: 10.1007/s10459-017-9787-9
- Huettel, S., Song, A., & McCarthy, G. (2008). *Functional Magnetic Resonance Imaging*. Sunderland, MA: Sinauer Associates.
- Ivarsson, J. (2017). Visual expertise as embodied practice. *Frontline Learning Research*, *5*(3), 123 - 138. doi: 10.14786/flr.v5i3.253
- 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. (2015). Expertise in clinical pathology: bridging the gap. *Advances in Health Sciences Education*, *20*(4), 1089-1106. doi: 10.1007/s10459-015-9589-x
- 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
- 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.
- Johansson, R., & Johansson, M. (2014). Look here: Eye movements play a functional role in memory retrieval. *Psychological Science*, *25*, 236-242. doi: 10.1177/0956797613498260
- Kicken, W., Brand-Gruwel, S., Van Merriënboer, J. J. G., & Slot, W. (2009). The effects of portfolio-based advice on the development of self-directed learning skills in secondary vocational education. *Educational Technology, Research and Development*, *57*, 439-460. doi: 10.1007/s11423-009-9111-3
- Kok, E. M., De Bruin, A. B. H., Robben, S., G. F., & Van Merriënboer, J. (2012). Looking in the same manner but seeing it differently: Bottom-up and expertise effects in radiology. *Applied Cognitive Psychology*, *26*(6), 854-862. doi: DOI: 10.1002/acp.2886
- Kok, E. M., & Jarodzka, H. (2016). Before your very eyes: The value and limitations of eye tracking in medical education. *Medical Education*. doi: 10.1111/medu.13066
- Kok, E. M., & Jarodzka, H. (2017). Beyond your very eyes: eye movements are necessary, not sufficient. *Medical Education*. doi: 10.1111/medu.13384
- Krupinski, E. (2017). Receiver Operating Characteristic (ROC) analysis. *Frontline Learning Research*, *5*(3), 31 - 42. doi: 10.14786/flr.v5i3.250
- Kundel, H. L., & Nodine, C. (1975). Interpreting chest radiographs without visual search. *Radiology*, *116*, 527-532.
- Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic Component of Image Perception in Mammogram Interpretation: Gaze-tracking Study1. *Radiology*, *242*(2), 396-402. doi: 10.1148/radiol.2422051997
- Kundel, H. L., Nodine, C. F., & Toto, L. (1991). Searching for lung nodules. The guidance of visual scanning. *Investigative radiology*, *26*(9), 777-781.
- Laeng, B., Eidet, L. M., Sulutvedt, U., & Panksepp, J. (2016). Music chills: The eye pupil as a mirror to music's soul. *Consciousness and Cognition*, *44*, 161-178. doi: 10.1016/j.concog.2016.07.009



- LeBlanc, V. R., Brooks, L. R., & Norman, G. R. (2002). Believing is seeing: The influence of a diagnostic hypothesis on the interpretation of clinical features. *Academic Medicine*, 77(10), 67-69.
- LeBlanc, V. R., Dore, K., Norman, G. R., & Brooks, L. R. (2004). Limiting the playing field: does restricting the number of possible diagnoses reduce errors due to diagnosis-specific feature identification? *Medical Education*, 38, 17 - 24. doi: 10.1111/j.1365-2923.2004.01730.x
- LeBlanc, V. R., Norman, G. R., & Brooks, L. R. (2001). Effect of a diagnostic suggestion on diagnostic accuracy and identification of clinical features. *Academic Medicine*, 76(10), 18-20.
- Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser & M. Farr (Eds.), *The nature of expertise* (pp. 311-342). Hillsdale, NJ: Erlbaum.
- Litchfield, D., & Donovan, T. (2017). The flash-preview moving window paradigm: Unpacking visual expertise one glimpse at a time. *Frontline Learning Research*, 5(3), 66 - 80. doi: 10.14786/flr.v5i3.269
- Loewenfeld, I. (1999). *The pupil: Anatomy, physiology, and clinical applications*. Oxford: Butterworth-Heinemann.
- Mansfield, E., Farroni, T., & Johnson, M. (2003). Does gaze perception facilitate overt orienting? *Visual Cognition*, 10(1), 7-14. doi: 10.1080/713756671
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17, 578-586.
- Melo, M., Scarpin, D. J., Amaro, E., Passos, R. B., Sato, J. R., & Friston, K. J. (2011). How doctors generate diagnostic hypotheses: A study of radiological diagnosis with functional magnetic resonance imaging. *PLoS ONE*, 6, e28752. doi: 10.1371/journal.pone.0028752
- Niedermeyer, E., & Da Silva, F. (2005). *Electroencephalography*. Lippincott Williams & Wilkins.
- Nodine, C. F., & Mello-Thoms, C. (2010). The role of expertise in radiologic image perception. In E. Samei & E. Krupinski (Eds.), *Medical image perception and techniques* (pp. 139-156). Cambridge: Cambridge University Press.
- Norman, G. R., Coblenz, C. L., Brooks, L. R., & Babcock, C. J. (1992). Expertise in visual diagnosis: a review of the literature. *Academic Medicine*, 67(10), 78-83.
- Norman, G. R., Young, M., & Brooks, L. (2007). Non-analytical models of clinical reasoning: the role of experience. *Medical Education*, 41, 1140-1145. doi: 10.1111/j.1365-2923.2007.02914.x
- Nourbakhsh, N., Wang, Y., Chen, F., & Calvo, R. A. (2012). *Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks*. Paper presented at the Proceedings of the 24th Australian Computer-Human Interaction Conference.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429-434. doi: 10.1037//0022-0663.84.4.429
- 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.
- Ribas, L. M., Rocha, F. T., Siqueira Ortega, N. R., Freitas de Rocha, A., & Massad, E. (2013). Brain activity and medical diagnosis: an EEG study. *BMC Neuroscience*, 14, 109. doi: 10.1186/1471-2202-14-109
- Savin-Williams, R. C., Cash, B. M., McCormack, M., & Rieger, G. (2016). Gay, mostly gay, or bisexual leaning gay? An exploratory study distinguishing gay sexual orientations among young men. *Archives of Sexual Behavior*, 1-8. doi: 10.1007/s10508-016-0848-6
- Szulewski, A., Kelton, D., & Howes, D. (2017). Pupillometry as a tool to study expertise in medicine. *Frontline Learning Research*, 5(3), 55 - 65. doi: 10.14786/flr.v5i3.256
- Tatler, B. W., Hirose, Y., Finnegan, S. K., Pievilainen, R., Kirtley, C., & Kennedy, A. (2013). Priorities for selection and representation in natural tasks. *Philosophical Transactions of the Royal Society B*, 368(1628), 1-9. doi: 10.1098/rstb.2013.0066



- Van de Wiel, M. (2017). Examining expertise using interviews and verbal protocols. *Frontline Learning Research*, 5(3), 94 - 122. doi: 10.14786/flr.v5i3.257
- 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., 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
- Vanderhasselt, M.-A., Remue, J., Ng, K. K., & De Raedt, R. (2014). The interplay between the anticipation and subsequent online processing of emotional stimuli as measured by pupillary dilatation: the role of cognitive reappraisal. *Frontiers in psychology*, 5, 207. doi: 10.3389/fpsyg.2014.00207
- Wade, N. J., & Tatler, B. (2005). *'The Moving Tablet of the Eye': The Origins of Modern Eye Movement Research*. Oxford: Oxford University Press.
- Wang, C.-A., McInnis, H., Brien, D. C., Pari, G., & Munoz, D. P. (2016). Disruption of pupil size modulation correlates with voluntary motor preparation deficits in Parkinson's disease. *Neuropsychologia*, 80, 176-184. doi: 10.1016/j.neuropsychologia.2015.11.019
- Williams, A. M., Fawver, B., & Hodges, N. J. (2017). Using the 'expert performance approach' as a framework for improving understanding of expert learning. *Frontline Learning Research*, 5(3), 139 - 154. doi: 10.14786/flr.v5i3.267