Frontline Learning Research
An official journal of EARLI

Frontline Learning Research Vol. 5 No. 3 Special Issue (2017) 1-13
ISSN 2295-3159

Methodologies for Studying Visual Expertise

Andreas Gegenfurtner[1] & Jeroen J. G. van Merriënboer[2]

[1] Technische Hochschule Deggendorf, Germany

[2] Maastricht University, the Netherlands

Article received 14 July / revised 23 July / accepted 23 July / available online 31 July

Abstract

Visual expertise can be defined as maximal adaptation to the requirements of a vision-intensive task. The process of developing a "good eye" in vision-intensive tasks is proposed, indicated, and elaborated by various measures contingent on diverse methodological arenas, all of which attempt to advance our understanding of what constitutes visual expertise. The aim of this special issue is to provide a reflection on this methodological pluralism and to offer a discussion of the affordances and constraints of some of these methodological approaches. Specifically, grounded on the medical domain, this special issue brings together a selection of nine articles that discuss cognitive-neurosciences, receiver operating characteristics (ROC) analysis, eye tracking, pupillometry, the flash-preview moving window paradigm, the combination of eye tracking data and verbal report data, the use of interviews and verbal protocols, ethnomethodology, and the expert performance approach. Two commentaries conclude the special issue. As an introduction, this article presents a comparative metaphorical mapping of visual expertise research. Metaphors are a useful tool for mirroring in simple terms the often complex paradigms underlying theory and applied research practice. We first identify four metaphors used in the analysis of visual cognition: activation, detection, inference, and practice. These metaphors are described with an empirical example and discussed to elicit (partly tacit) assumptions associated with prototypical method decisions. We then link the proposed metaphorical mapping to the contributions in this special issue.

Keywords: visual expertise; expert performance; visual cognition; professional vision; methods in learning research.

[1] Corresponding author: Andreas Gegenfurtner, Technische Hochschule Deggendorf, Institut für Qualität und Weiterbildung, Dieter-Görlitz-Platz 1, 94469 Deggendorf, Germany, Email: andreas.gegenfurtner@th-deg.de DOI: http://dx.doi.org/10.14786/flr.v5i3.316

1. Introduction

This special issue is devoted to research on visual expertise. Visual expertise can be defined as maximal adaptations to the requirements of a vision-intensive task. Examples of vision-intensive tasks include the identification of different types of fish (Boucheix & Lowe, in press; Jarodzka, Scheiter, Gerjets, & Van Gog, 2010) or the detection of abnormalities in microscopic specimen (Helle, Nivala, Kronqvist, Gegenfurtner, Björk, & Säljö, 2011; Krupinski, Graham, & Weinstein, 2013). In many professions, visual material constitutes an important part of the epistemic resources used for conducting professional work (Gegenfurtner, Nivala, Säljö, & Lehtinen, 2009; Goodwin, 1994; Gruber & Degner, 2016; Palonen, Boshuizen, & Lehtinen, 2014; Säljö, 2012). Consequently, newcomers need to learn, appropriate, and master the skills associated with domain-specific visual tasks (Gegenfurtner, Lehtinen, Jarodzka, & Säljö, 2017a; Kok, Van Geel, Robben, & Van Merriënboer, 2017; Seppänen & Gegenfurtner, 2012; Szulewski, Gegenfurtner, Sivilotti, Howes, & Van Merriënboer, in press). Past research has employed very different strategies to examine these learning processes of novices as well as the processes and practices underlying the superior performance of domain experts.

Different research methodologies are associated with different underlying epistemologies (Damsa et al., in press). Therefore, the phenomenon of visual expertise is approached from different perspectives that correspond with different assumptions about what constitutes the allegorical "good eye" of an expert. Because expertise and expert performance are largely domain-specific (Gegenfurtner & Seppänen, 2013; Gruber & Degner, 2016), this special issue focuses on one example domain: medicine. Medicine has been chosen here because many medical specialties rely on the analysis of visual material, such as the human skin, X-rays, pathological slides, or electrocardiograms. Although the focus is on medicine, the lessons learned from a reflection on methodologies for studying visual expertise can inform other vision-intensive domains to advance our understanding of an expert's professional vision (Goodwin, 1994). Some investigators in medicine, particularly Kundel, Nodine, and Carmody (1978), in a now classic study, suggest that the highly perceptual nature of image comprehension requires intensive processing of visual data through oculomotor activity that guides signal detection and decision-making. Others, particularly Lesgold and colleagues (1988), in what is now also a classic study, put less emphasis on the perceptual aspect; rather, they suggest that visual expertise is mainly the function of cognitive inference that aligns schemata from episodic memory consistent with the perceptual features detected. Much medical research done in both traditions has been reviewed elsewhere (Boshuizen & Schmidt, 2008; Ericsson, 2004; Gegenfurtner, Kok, Van Geel, De Bruin, Jarodzka, Szulewski, & Van Merriënboer, 2017; Gegenfurtner, Siewiorek, Lehtinen, & Säljö, 2013; Patel, Arocha, & Zhang, 2005). Recently, alternative approaches have been formulated; these suggest that a good eye is indicated by neurophysiologic events in certain brain areas (Bilalić, 2017; Gegenfurtner, Kok, Van Geel, De Bruin, & Sorger, 2017b; Haller & Radue, 2005), and is accomplished through situated social discourse (Ivarsson, 2017; Johansson, Lindwall, & Rystedt, 2017; Koschmann, LeBaron, Goodwin, Zemel, & Dunnington, 2007). In short, the allegory of having a good eye in medicine is proposed, indicated, and elaborated by various measures contingent on diverse methodological arenas that all attempt to advance our understanding of what constitutes visual expertise.

The main aim of this special issue is to provide a reflection on this methodological pluralism. For this purpose, we have invited scholars from various professional backgrounds to contribute an article that introduces a particular methodological approach used to study visual expertise. Each article is devoted to one approach (or to a combination of approaches) and discusses its affordances and constraints for empirically analyzing visual expertise. These approaches are: cognitive-neurosciences (Gegenfurtner et al., 2017b), receiver operating characteristics analysis (Krupsinki, 2017), eye tracking (Fox & Faulkner-Jones, 2017), pupillometry (Szulewski, Kelton, & Howes, 2017), the flash-preview moving window paradigm (Litchfield & Donovan, 2017), the combination of eye tracking data and verbal report data (Helle, 2017), the use of interviews and verbal protocols (Van de Wiel, 2017), ethnomethodology (Ivarsson, 2017), and the expert performance approach (Williams, Fawyer, & Hodges, 2017). The special issue closes with two commentaries (Boucheix, 2017; Jarodzka & Boshuizen, 2017). Before we introduce each contribution, we would like to present a framework that aims to structure the pluralism of methodologies in visual expertise research.

2. Four Metaphors in Visual Expertise Research

Our goal is to present a framework in which the different metaphors of having a "good eye" can be considered as mutually constituting the richness we have on ideas, concepts, and theories relating to visual expertise in medicine and beyond. We have no intent to judge some methodological traditions as being more valuable than others, nor do we intend to unify them in some abstract way. Although the idea of what was termed *interactive complexification* (Alexander, Schallert, & Reynolds, 2009) is considered meaningful for highlighting the confluence of factors that determines "any given aspect" of the product and the process of learning to diagnose medical images, we believe there is, at times, also a place for simple answers. We hope that the framework put forward in this article will provide such a simple answer that can be used as a glass through which we look at the phenomenon of visual expertise. The framework addresses the questions of how to analyze visual expertise and how to elicit tacit assumptions underlying common research practice. By discussing four examples drawn from radiology, we identify four metaphors that constitute our framework on visual expertise research. Despite its groundwork in medical literature, the framework also has implications beyond medicine, due to its common interest in learning and comprehension in technology-rich vision-intensive contexts.

The framework is built around four metaphors. Metaphors are a useful tool for mirroring in simple terms the often complex paradigms underlying theory and applied research practice. An example of how metaphors are able to elicit complex, different, yet partly tacit, assumptions on a commonly studied phenomenon can be found in "learning as acquisition", "learning as participation", and "learning as knowledge creation" (Paavola, Lipponen, & Hakkarainen, 2002; Sfard, 1998; see also Hakkarainen, Palonen, Paavola, & Lehtinen, 2004). As Sfard (1998, p. 4) notes, "metaphors are the most primitive, most elusive, and yet amazingly informative objects of analysis". We believe that their value and power stems from the fact that metaphors converge and portray, in a snapshot format, what took years of scientific discourse to develop; this allows frank presentation of positions and their entailments in a condensed way and invites a critical (re)consideration of accepted and perhaps unreflected practice. Of course, metaphors are simple and simplistic; there is no claim that they attempt to depict all of the breadth and depth of what often is a complex epistemology.

Table 1

A comparative metaphorical mapping of visual expertise as activation, detection, inference, and practice

	Activation	**Detection**	**Inference**	**Practice**
Indicators of visual expertise	Neurophysiologic activity	Eye movements	Verbal reports	Representational practices
Unit of analysis	Individual	Individual (social)	Individual and social	Sociotechnical
Place of visual cognition	Neural network system	Optic system	(Distributed) memory system	Activity system
Analytic time span	Milliseconds to seconds	Seconds	Minutes to few hours	Minutes to decades
Associated methodology	Cognitive neuroscience	ROC analysis; eye tracking methodology	Protocol analysis; interviews	Ethnomethodology, ethnography

Below, we identify, exemplify, and discuss four metaphors often used when analyzing visual expertise; these metaphors are seen as four of the many dimensions possessed by one with “a good eye”. We reflect on these metaphors in terms of how they contribute to research devoted to examining visual expertise. It is hoped that the value and significance of this methodological reflection will be that it helps to map a scattered and fragmented field of inquiry. Table 1 serves as the guiding framework for the comparison of the four metaphors, including their methodological entailments: activation, detection, inference, and practice.

2.1. Visual expertise as activation metaphor

Research adhering to the activation metaphor uses measures of neurophysiologic activity as an indication of visual expertise. In the activation metaphor, there is a strong emphasis on the neurological and biological basis of our humanness (Alexander et al., 2009; see also Meltzoff, Kuhl, Movellan, & Sejnowski, 2009). This emphasis might originate from the widely held belief that “information is stored in neural networks in the brain, and that human behavior arises from extremely complex communication between neurons in these networks and also between separate networks or assemblies” (Sauseng & Klimesch, 2008, p. 1003). This neural network system is seen as the place where visual cognition and expertise “occurs”. Hence, visual expertise is indicated by neural activity. Specifically, this activity can be measured by an electroencephalograph (EEG) as the electric current in axons; by a magnetoencephalograph (MEG) as the magnetic field induced by those electric currents; by a positron emission tomography (PET) as the blood flow distribution in the cells; or by functional magnetic resonance imaging (fMRI) scanning as differences in cellular oxygen consumption. Whenever information stored in neural networks is used for cognitive processes, neural activation can be measured with one of those tools. For example, if a radiologist formulates a diagnosis based on a patient’s medical image, an fMRI scanner could be used to indicate the processes of this radiologist’s visual cognition. These processes are extremely fast; the best conventional apparatus currently available—the EEG scanner—is able to trace this activity with a temporal resolution in the range of milliseconds (Sauseng & Klimesch, 2008). An empirical example prototypical for the activation metaphor can further illustrate epistemological and methodological premises.

Haller and Radue (2005) investigated differences in neuronal activations of radiologists and laypersons in reading radiologic and non-radiologic images. Using functional magnetic resonance (fMRI) imaging, the brain scans showed that radiologic images evoked stronger activations in the brains of radiologists than in those of laypersons, with the bilateral middle and inferior temporal gyrus, bilateral medial and middle frontal gyrus, and left superior and inferior frontal gyrus being particularly affected. These regions are generally assumed to be linked to the encoding and storing of memory of visual objects and events. Hence, this finding seems to imply that what is seen on the presented image is automatically referenced to memorized images, indicating an unconscious, stimulus-driven indexical relation between the pictorial representation and the corresponding mental representation. Being prototypical for research in neuroscience, Haller and Radue (2005) used technological images as stimuli in a series: Stimuli were shown for 2.5 seconds followed by a fixation cross for 8.5 seconds to compensate for blood oxygenation level-dependent signal delay. Subjects gazed at the stimuli series with an immobilized head in darkened and (electrical and auditory) noise-protected rooms. Settings of this kind are highly controlled. These types of controls are necessary because neural measures are highly sensitive: activation should be shown in response to the stimulus only. Strong controls therefore aim to guarantee bias-free recordings.

At this point in time, the activation metaphor for visual expertise has rarely been used in medical diagnosis studies (Gegenfurtner et al., 2017b). However, the coming together of learning research and neuroscience is beginning to form an exciting new field (Ansari & Coch, 2006; De Jong, Van Gog, Jenks, Manlove, Van Hell, Jolles, et al., 2009). Neuroscience has the potential to trace implicit and experiential learning before it can be observed in behavior. This can help us understand when, how, and why learning occurs. In particular, the how and why can be tackled within these highly controlled settings. Certainly, it is not a novel statement that behaviors, such as diagnosing a medical image, that appear similar on the surface may involve very different cognitive/perceptual mechanisms underlying this behavior. Neuroscience, in combination with the learning sciences, now provides a new avenue for tackling these issues, to further understand visual expertise (Bilalić, 2017; Gegenfurtner et al., 2017b).

2.2. Visual expertise as detection metaphor

Detection can be defined as “determining whether a simple, featurally defined stimulus is present in, or absent from, the visual field” (Smith & Ratcliff, 2009, p. 283). A central premise of research using the detection metaphor is uncertainty; that is, the degree to which a subject is able to discriminate between signal (the stimulus of interest) and noise (background stimuli distracting visual attention, thus causing decision-making under conditions of uncertainty). In medical image diagnosis, the signal would be a tumor, while noise would be (healthy) organic material surrounding the tumor. Clearly, in pictures as complex as radiographs, with an abundance of structures, forms, and elements displayed in manifold shadings of grey, and with the typical presence of technical artefacts, detection of a tumor is a challenging task. Tasks of this kind are frequently used to quantify the ability of discerning between signal and noise. Two approaches are prevalent: eye-tracking methodology and receiver operating characteristic (ROC) analysis. Below, an empirical example by Kundel, Nodine, Conant, and Weinstein (2007), which combined eye-tracking with ROC analysis, can illuminate prototypical premises of each approach.

Kundel et al. (2007) investigated rapid initial fixations (detections) on abnormalities on mammograms. Briefly, they found that more experienced radiologists showed global perceptual processes that helped them detect the abnormality (malignant breast lesions) in less than a second. In contrast, less experienced radiologists showed search-to-find strategies that took considerably longer to first fixate the abnormality.

Expertise differences in these two groups were indicated by eye-tracking and by ROC analysis. With respect to eye-tracking, the recording of eye movements is usually used to visualize the scan paths of observers. In Kundel's study, radiologists with more experience had longer saccades and fewer fixations than did less experienced radiologists. With respect to ROC analysis, detectability was significantly higher for observers with more experience than for those with less experience. Detectability is a measure that quantifies the sum of true positives and true negatives, divided by the sum of all positives and negatives in a detectability value, d_a.

A conclusion can be made that the detection metaphor indicates that novice diagnosticians develop a good eye through medical education in terms of their ability to discriminate a potential signal from background noise. This ability can be quantified and expressed mathematically in a formula that allows comparison of observers at the individual or the group level. Studies in the detection metaphor, which usually employ eye-tracking methodology and/or ROC analysis, indicate that superior visual cognition can be characterized as a high decision-speed accuracy relation: Visual perception changes with a rise in experience, from a relatively slow search-to-find mode to a global holistic mode. This change then increases sensitivity (i.e., proportion of correctly identified abnormalities), specificity (i.e., proportion of correctly identified healthy tissue), and thus accuracy of the detection performance. Usually, the analytic time span is somewhat longer than the time span of cognitive neuroscience studies. The work of Kundel et al. (2007), which can be seen as a prototypical example, reported an average search time of 26.90 seconds, and a median time to first fixate the abnormality (the signal) of 1.13 seconds. An exciting direction for further research adopting the detection metaphor is a focus on the transfer of expertise. Transfer is a concept to describe how skills in one field are applicable in a second field (Gegenfurtner et al., 2010; Quesada-Pallarès & Gegenfurtner, 201). In the context of expertise research, eye tracking (and other methods as well) can be used to study how visual expertise transfers from a domain-specific, typical, or routine task to a domain-general, atypical, or novel task (Gegenfurtner & Seppänen, 2013; Gegenfurtner et al., 2017c). Another exciting research direction is collaborative gaze. Traditionally, eye-tracking studies have focused on individual observers as the unit of analysis (Seppänen & Gegenfurtner, 2012); however, developments of eye-tracking technology and analytic algorithms now allow collaborative gaze studies (e.g., Sangin, Molinari, Nüssli, & Dillenbourg, 2008). It will be fascinating, from an epistemological point of view, to follow the coping of tension between attentional detection as an individual quantifiable performance, notable in mathematical functions (i.e., Smith & Ratcliff, 2009), on one hand, and detection as collaborative achievement and intersubjective meaning-making (much in line with Koschmann & Zemel's, 2009, notion of *discovery as occasioned production*), on the other hand.

2.3. Visual expertise as inference metaphor

Lesgold and colleagues (1988, p. 336) speculated that radiological diagnosis "is largely a matter of cognitive inference. That is, given a set of findings (perceptual features), one has to determine which diseases are consistent with those findings. If more than one disease is consistent, then one either looks further, (…) or suggests additional medical tests to discriminate among the possibilities". Two issues are striking in this initial quote. First, Lesgold emphasizes cognition and memory processes in diagnosing medical images. Back in the 1980s, this was not customary in the medical literature. Although there have been pioneering studies on cognitive processes (e.g., Patel & Groen, 1986; Boshuizen, 1989), most focused on perceptual processes (based on Arnheim, 1969, see also the section on the detection metaphor). What Lesgold indicated and empirically tested was thinking as an essential function in medical diagnosis. Second, in this quote and elsewhere in his chapter, Lesgold emphasized how vision and cognition correlated in the formation and evaluation of diagnostic decisions: Experienced radiologists build mental representations that guide perception. The literature now knows a variety of rhetorical functions built from verbal protocols to describe those mental representations, among them Lesgold's schemata, encapsulated scripts (Boshuizen & Schmidt, 2008), E-MOPs (Kolodner, 1983), or SUSTAINs (Love, Medin, & Gureckis, 2004). In the next paragraph, we describe one example on the correlation of vision and cognition in medical image diagnosis that is still informed by Lesgold's (one might tend to say: seminal) speculation of cognitive inference.

Morita and colleagues (2008) investigated how perceptual and conceptual processing interrelates in the diagnosis of computer tomograms (CT). Shortly, they found that expert compared to novice CT readers verbalized more findings, more hypotheses, and more perceptual activities. Importantly, experts verbalized many perceptual features during conceptual activities, and verbalized conceptual words during perceptual processing. Put differently, this indicates that experts retrieved and used knowledge from memory based on information that they saw on the CT image, which iteratively stimulated looking at the image based on knowledge coded in memory (be it in the form of encapsulated scripts, E-MOPs, or schemata). From a methodological point of view, it would be tempting to criticize the neglect to use eye movement recordings; this would have allowed highly specific, quantifiable measures of perceptual activity. Yet, verbal protocols can also be used as indicators for visual expertise (Helle, 2017; Van de Wiel, 2017). Usually, as prototypically shown in this example, protocols are collected for a duration of up to few hours and then are analyzed with a focus on cognitive mechanisms. From this perspective, the inference metaphor on visual expertise clearly emphasizes the cognitive parts of the interrelated process. Morita and colleagues (2008) decided on individual CT readers as a unit of analysis. However, protocols can also be used at a group level (Greeno, 2006; see Simpson & Gilhooly, 1997, for an example in cardiology) to indicate collaborative negotiations and intersubjective meaning-making.

The inference metaphor in visual expertise research can answer, in two respects, the question regarding what develops in novice diagnosticians that moves them toward higher accuracy. First, knowledge develops; an extensive knowledge base is the foundation for expert performance and for rapid inference of coded memory to detected visible features. Second, the perceptual-conceptual processing linkage develops. Morita and colleagues have demonstrated that protocol measures are a valid tool for eliciting cognitive mechanisms underlying CT diagnoses that guide, and are guided by, perceptual detection. Epistemologically, the inference metaphor appears to account for both signal detection and the alignment of knowledge from memory (inference) that is consistent to what is detected. Nevertheless, methodologically, it lacks the precise and time-sensitive measures such as eye-tracking gaze recordings or cortical oscillation EEGs. This is because researchers rely on explicit, conscious think-aloud utterances from participants and these cannot account for their underlying implicit, non-conscious processing. Hence, the sole use of protocols—be it at an individual or at an agglomerated group level—risks the resemblance of linguistic descriptions that play a rhetorical function in describing and illustrating phenomena; examples of these fancy rhetorical functions, which simply cannot be fully validated by protocol analysis alone, are provided above (SUSTAINs and the like).

2.4. Visual expertise as practice metaphor

Finally, the last metaphor we identify as being frequently used in visual expertise research is the practice metaphor. Sociocultural practice generates semantic structures of information that shape and are shaped by sequentially unfolding activity in relevant manners for a domain of scrutiny, such as laparoscopy or sports (Gegenfurtner & Szulewski, 2016; Koschmann et al., 2007). As a starting point, we present the following quote from Carsetti (2004, p. 307) that we found interesting enough to use to introduce our reflection on the practice metaphor: "A percept is something that lives and becomes, it possesses a biological complexity which is not be explained simply in terms of the computations by a neural network classifying on the basis of very simple mechanisms". This quote has two interesting elements. First, it emphasizes the lived nature of visual cognition, or what Livingston (1986) referred to as the *lived work of reading*. We will present an empirical example in the next paragraph that elaborates on this notion. Livingston, in a series of ethnographic field descriptions, highlighted the sociability of practices that constitute intersubjective thinking and acting. As such, the author provided a look that differed from looks "behind the skull" (Garfinkel, 1967) or from "computations by a neural network" (Carsetti, ibid). The second interesting element in this quote is that it seems to align neuroscientific work with labels such as "simply" and "very simple". We lack authority and motivation to judge such a judgment about the simpleness of neuroscience as being itself simplistic. Nevertheless, it illustrates the position of this author that something that focuses only on neural activation is unable to account for the full complexity of visual expertise (interestingly, compare the quote of Sauseng & Klimesch, 2008, starting the activation metaphor section, where the complexity of neural network communication is emphasized). Certainly, it is a matter of definition what "complex" is or what shall be allowed—based on which methodological and epistemological considerations—to have "complexity". Making such (maybe tacit, maybe deliberate) assumptions explicit is one of the purposes of the framework in Table 1.

To further illustrate the methodological entailments of the practice metaphor, we describe an example elaborating on the lived work of mammography. Slack, Hartwood, Procter, and Rouncefield (2007) highlight how diagnosing a mammogram is reflexively contingent on artful practices, in which multiple readers interact and intersubjectively constitute breast biographies. Central in their analyses are practices. Goodwin (1994, 2000) indicated that seeing and interpreting what is seen are not exclusively cognitive processes located in the individual brain (cf. activation and detection metaphors); rather, seeing is a socially situated activity accomplished through the deployment of a range of historically matured discursive practices. These practices constitute visual expertise in Goodwin's terms, and they are negotiated around a common object of disciplined perception (Ivarsson, 2017; Lindwall & Lymer, 2008; Stevens & Hall, 1998), in the study of Slack and colleagues (2007): pictorial representations of the breast produced by an X-ray apparatus. Slack et al. identified practices such as arranging mammograms, manipulating, annotating, gesturing, and pointing that contribute to the lived work of doing a radiologic diagnosis. These representative practices (Greeno, 2006) unfold within an activity system, in many cases temporally over the course of minutes, but their sociogenesis stretches over the course of decades (such as the material resources used; i.e., pictures produced by X-ray technology). Hence, analysis of visual expertise using the practice metaphor adopts a different analytic time span than does, for example, analysis using the activation metaphor; and it adopts a sociotechnical system as the unit of analysis that explicitly accounts for the mediating role of technology (Burri & Dumit, 2008; Gegenfurtner, 2013; Säljö, 2012; Siewiorek & Gegenfurtner, 2010). It is essentially the focus on embodied talk-in-interaction—talk between people (Knogler et al., 2013) and between humans and non-human objects (Gibson, 1979; Suchman, 2007)—that makes the practice metaphor a useful tool to analyze visual expertise and to generate a practice-based theorizing about the work of experts.

3. Structure of the Special Issue

Devoted to research on visual expertise, this special issue consists of 12 contributions in three sections: (a) This Introduction sets the stage for the unfolding reflection, (b) 9 articles reflect on diverse methodological approaches, and (c) 2 commentaries by Boucheix (2017) as well as Jarodzka and Boshuizen (2017) close the special issue. These two concluding commentaries offer a detailed discussion of each of the nine articles and synthesize their lessons in novel ways. Hence, an in-depth summary of the nine articles is redundant here. However, we would like to link the methodological approaches outlined and discussed in the nine articles with the comparative metaphorical mapping presented in Table 1.

First, the *activation metaphor* corresponds with cognitive-neurosciences. Prominent methods within cognitive-neurosciences are functional magnetic resonance imaging and electroencephalography. Gegenfurtner et al. (2017b) reflect on these methods, their potential advantages, and their risks for studying visual perceptual expertise in medicine. Second, the *detection metaphor* corresponds with ROC analysis and eye tracking methodology. Krupinski (2017) discusses ROC analysis. Several articles cover eye tracking. Fox and Faulkner-Jones (2017) offer a general discussion. Szulewski et al. (2017) focus on pupillometry. Litchfield and Donovan (2017) reflect on the benefits of the flash-preview moving window paradigm. And Helle (2017) explores how eye tracking and verbal report data can be combined. This is already a bridge toward the third metaphor: inference. Most prominently, the *inference metaphor* corresponds with verbal report data and interviews. The latter is introduced and discussed by Van de Wiel (2017). Finally, the *practice metaphor* corresponds with ethnomethodology. Ivarsson (2017) offers an ethnomethodological reflection of visual expertise as embodied practice. As a contribution that overarches the four metaphors, Williams and colleagues (2017) describe how the expert performance approach can inform research designs intending to study visual expertise.

Different methods lead to different answers based on different indicators (Damşa et al., in press). It is a reflection on these indicators, and more generally on the methodological entailments behind seemingly different metaphors, that can help raise awareness of each metaphor's (epistemological and pragmatic) potentiality and contingency, and that can thus advance our research practice in medical education and the learning sciences. Sfard (1998) noted that a combination of learning metaphors yields to more robust findings than does a non-combination. The detection metaphor seems to be currently dominant in the field of visual expertise research (which is also reflected in the number of articles in this special issue). Further developments in the field could profit from combining this particular metaphor with alternative metaphors. The special issue is thus a first step towards reaching this ambitious goal (see also the commentaries by Boucheix, 2017, and Jarodzka and Boshuizen, 2017, as well as Gegenfurtner et al., 2013; Lehtinen, 2012; Säljö, 2009). Probably, research aimed at method triangulation will advance the field more than discussions about the superiority of a particular method (Carsetti, 2004; Sauseng & Klimesch, 2008). This is in line with research in times that have been labeled the *decade of synergy* (Bransford et al., 2006). Yet, combining methods is neither trivial nor simple. It is an essential task for future research on visual expertise to explore the synergies between metaphors to avoid the dangers associated with choosing just one metaphor.

Keypoints

- The special issue discusses the methodological pluralism of research on visual expertise.
- This introduction offers a comparative metaphorical mapping of visual expertise research.
- The mapping includes four metaphors: activation, detection, inference, and practice.
- Combining metaphors will help the further development of visual expertise research.

References

Alexander, P. A., Schallert, D. L., & Reynolds, R. E. (2009). What is learning anyway? A topographical perspective reconsidered. *Educational Psychologist*, *44*, 176-192. doi:10.1080/00461520903029006

Ansari, D., & Coch, D. (2006) Bridge over troubled waters: education and cognitive neuroscience. *Trends in Cognitive Sciences*, *10*, 146-151. doi:10.1016/j.tics.2006.02.007

Arnheim, R. (1969). *Visual thinking*. Berkeley: University of California Press.

Bilalić, M. (2017). *The neuroscience of expertise*. Cambridge: Cambridge University Press.

Boshuizen, H. P. A. (1989). *De ontwikkeling van medische expertise: Een cognitief-psychologische benadering*. Maastricht: Maastricht University.

Boshuizen, H. P. A., & Schmidt, H. G. (2008). 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* (pp. 113-121). Oxford: Butterworth-Heinemann.

Boucheix, J.-M. (2017). The interplay between methodologies, tasks and visualisation formats in the study of visual expertise. *Frontline Learning Research*.

Boucheix, J.-M., Bonnetain, E., Avena, C., & Freysz, M. (2010). Benefits of learning technologies in medical training, from full-scale simulators to virtual reality and multimedia presentations. In J. P. Didier, E. Bigand, & A. Vinter (Eds.), *Rethinking physical and rehabilitation medicine* (pp. 171-191). New York: Springer.

Boucheix, J.-M., & Lowe, R. (in press). Generative processing of animated partial depictions fosters fish identification skills: Eye tracking evidence. *Le Travail Humain*.

Bransford, J., Stevens, R., Schwartz, D., Meltzoff, A., Pea, R., Roschelle, J., et al. (2006). Learning theories and education. Toward a decade of synergy. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 209-244). Mahwah, NJ: Erlbaum.

Burri, R. V., & Dumit, J. (2008). Social studies of scientific imaging and visualization. In E. J. Hackett, O. Amsterdamska, M. Lynch, & J. Wajcman (Eds.), *Handbook of science and technology studies* (pp. 297-317). Cambridge, MA: MIT Press.

Carsetti, A. (2004). The embodied meaning. Self-organization and symbolic dynamics in visual cognition. In A. Carsetti (Ed.), *Seeing, thinking, and knowing. Meaning and self-organization in visual cognition and thought* (pp. 307-327). Dordrecht: Kluwer.

Damsa, C. I., Froehlich, D. E., & Gegenfurtner, A. (in press). Reflections on empirical and methodological accounts of agency at work. In M. Goller & S. Paloniemi (Eds.), *Agency at work: An agentic perspective on professional learning and development*. New York: Springer.

De Jong, T., Van Gog, T., Jenks, K., Manlove, S., Van Hell, J., Jolles, J., et al. (2009). Explorations in *learning and the brain. On the potential of cognitive neuroscience for educational science*. Berlin: Springer.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*, S70-S81. doi:10.1097/00001888-200410001-00022

Fox, S. E., & Faulkner-Jones, B. E. (2017). Eye-tracking in the study of visual expertise: Methodology and approaches in medicine. *Frontline Learning Research*. doi: doi:10.14786/flr.v5i3.258

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.

Gegenfurtner, A. (2013). Transitions of expertise. In J. Seifried & E. Wuttke (Eds.), *Transitions in vocational education* (pp. 305-319). Opladen: Budrich.

Gegenfurtner, A., Kok, E., Van Geel, K., De Bruin, A., Jarodzka, H., Szulewski, A., & Van Merriënboer, J. J. G. (2017a). The challenges of studying visual expertise in medical image diagnosis. *Medical Education*, 51, 97-104. doi:10.1111/medu.13205

Gegenfurtner, A., Kok, E. M., Van Geel, K., De Bruin, A. B. H., & Sorger, B. (2017b). Neural correlates of visual perceptual expertise: Evidence from cognitive neuroscience using functional neuroimaging. *Frontline Learning Research.* doi:10.14786/flr.v5i3.259

Gegenfurtner, A., Lehtinen, E., Jarodzka, H., & Säljö, R. (2017c). Effects of eye movement modeling examples on adaptive expertise in medical image diagnosis. *Computers & Education*, *113*, 212-225. doi:10.1016/j.compedu.2017.06.001

Gegenfurtner, A., Nivala, M., Säljö, R., & Lehtinen, E. (2009). Capturing individual and institutional change: Exploring horizontal versus vertical transitions in technology-rich environments. In U. Cress, V. Dimitrova, & M. Specht (Eds.), *Learning in the synergy of multiple disciplines. Lecture Notes in Computer Science* (pp. 676-681). Berlin: Springer. doi:10.1007/978-3-642-04636-0_67

Gegenfurtner, A., & Seppänen M. (2013). Transfer of expertise: An eye-tracking and think-aloud study using dynamic medical visualizations. *Computers & Education*, *63*, 393-403. doi:10.1016/j.compedu.2012.12.021

Gegenfurtner, A., Siewiorek, A., Lehtinen, E., & Säljö, R. (2013). Assessing the quality of expertise differences in the comprehension of medical visualizations. *Vocations and Learning*, *6*, 37-54. doi:10.1007/s12186-012-9088-7

Gegenfurtner, A., & Szulewski, A. (2016). Visual expertise and the Quiet Eye in sports - comment on Vickers. *Current Issues in Sport Science*, *1*, 108. doi:10.15203/CISS_2016.108

Gegenfurtner, A., Vauras, M., Gruber, H., & Festner, D. (2010). Motivation to transfer revisited. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the disciplines: ICLS2010 proceedings* (Vol. 1, pp. 452-459). Chicago, IL: International Society of the Learning Sciences.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton.

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*, 606-633. doi:10.1525/aa.1994.96.3.02a00100

Goodwin, C. (2000). Practices of seeing: Visual analysis: An ethnomethodological approach. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 157-182). London: Sage.

Greeno, J. A. (2006). Learning in activity. In R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 79-96). Cambridge, MA: Cambridge University Press.

Gruber, H., & Degner, S. (2016). Expertise und Kompetenz [Expertise and competence]. In M. Dick, W. Marotzki, & H. Mieg (Eds.), *Handbuch Professionsentwicklung* (pp. 173-180). Bad Heilbrunn: Klinkhardt.

Hakkarainen, K., Palonen, T., Paavola, S., & Lehtinen, E. (2004). *Communities of networked expertise: Professional and educational perspectives.* Amsterdam: Elsevier.

Haller, S., & Radue, E. W. (2005). What is different about a radiologist's brain? *Radiology*, *236*, 983-989. doi:10.1148/radiol.2363041370

Helle, L. (2017). Prospects and pitfalls in combining eye-tracking data and verbal reports. *Frontline Learning Research.* doi:10.14786/flr.v5i3.254

Helle, L., Nivala, M., Kronqvist, P., Gegenfurtner, A., Björk, P., & Säljö, R. (2011). Traditional microscopy instruction versus process-oriented virtual microscopy instruction: A naturalistic experiment with control group. *Diagnostic Pathology*, *6*, S81-S89. doi:10.1186/1746-1596-6-S1-S8

Ivarsson, J. (2017). Visual expertise as embodied practice. *Frontline Learning Research*. doi:10.14786/flr.v5i3.253

Jarodzka, H., & Boshuizen, H. P. A. (2017). Unboxing the black box of visual expertise in medicine. *Frontline Learning Research*.

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*, 146-154. doi:10.1016/j.learninstruc.2009.02.019

Johansson, E., Lindwall, O., & Rystedt, H. (2017). Experiences, appearances, and interprofessional training: The instructional use of video in post-simulation defbriefings. *International Journal of Computer-Supported Collaborative Learning*, *12*, 91-112. doi:10.1007/s11412-017-9252-z

Knogler, M., Gegenfurtner, A., & Quesada Pallarès, C. (2013). Social design in digital simulations: Effects of single versus multi-player simulations on efficacy beliefs and transfer. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *To see the world and a grain of sand: Learning across levels of space, time, and scale* (Vol. 2, pp. 293-294). Madison, WI: International Society of the Learning Sciences.

Kok, E. M., Van Geel, K., Van Merriënboer, J. J. G., & Robben, S. G. F. (2017). What we do and do not know about teaching medical image interpretation. *Frontiers in Psychology, 8*, 309. doi: 10.3389/fpsyg.2017.00309
Kolodner, J. L. (1983). Towards an understanding of the role of experience in the evolution from novice to expert. *International Journal of Man-Machine Systems*, *19*, 497-518. doi:S0020-7373(83)80068-6
Koschmann, T., LeBaron, C., Goodwin, C., Zemel, A., & Dunnington, G. (2007). Formulating the triangle of doom. *Gesture*, *7*, 97-118. doi:10.1075/gest.7.1.06kos
Koschmann, T., & Zemel, A. (2009). Optical pulsars and black arrows: Discoveries as occasioned productions. *Journal of the Learning Sciences*, 18, 200-246. doi:10.1080/10508400902797966
Krupinski, E. A. (2017). Receiver operating characteristics. *Frontline Learning Research*. doi: doi:10.14786/flr.v5i3.250
Krupinski, E. A., Graham, A., R., & Weinstein, R. S. (2013). Characterizing the development of visual search expertise in pathology residents viewing whole slide images. *Human Pathology*, *44*, 357-364. doi:10.1016/j.humpath.2012.05.024
Kundel, H. L., Nodine, C. F., & Carmody, D. (1978). Visual scanning, pattern recognition, and decision-making in pulmonary nodule detection. *Investigative Radiology*, *13*, 175-181.
Kundel, H. L., Nodine, C. F., Conant, E. F., Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: Gaze-tracking study. *Radiology*, *242*, 396-402. Doi:10.1148/radiol.2422051997
Lehtinen, E. (2012). Learning of complex competences: On the need to coordinate multiple theoretical perspectives. In A. Koskensalo, J. Smeds, R. de Cillia, & Á. Huguet (Eds.), *Language: Competencies - change - contact* (pp. 13-27). Berlin: LIT.
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. J. Farr (Eds.), *The nature of expertise* (pp. 311-342). Hillsdale, NJ: Erlbaum.
Lindwall, O., & Lymer, G. (2008). The dark matter of lab work: Illuminating the negotiation of disciplined perception in mechanics. *Journal of the Learning Sciences*, *17*, 180-224. doi:10.1080/10508400801986082
Litchfield, D., & Donovan, T. (2017). The flash-preview moving window paradigm: Unpacking visual expertise one glimpse at a time. *Frontline Learning Research*. doi:10.14786/flr.v5i3.269
Livingston, E. (1986). *The ethnomethodological foundations of mathematics*. London: Kegan Paul.
Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*, 309-332. doi:10.1037/0033-295X.111.2.309
Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science*, *325*, 284-288. doi:10.1126/science.1175626
Morita, J., Miwa, K., Kitasaka, T., Mori, K., Suenaga, Y., Iwano, S., et al. (2008). Interactions of perceptual and conceptual processing: Expertise in medical image diagnosing. *International Journal of Human-Computer Studies*, *66*, 370-390. doi:10.1016/j.ijhcs.2007.11.004
Paavola, S., Lipponen, L., & Hakkarainen, K. (2002). Epistemological foundations for CSCL: A comparison of three models of innovative knowledge communities. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. CSCL 2002 Proceedings* (pp. 24-32). Hillsdale, NJ: Erlbaum.
Palonen, T., Boshuizen, H. P. A, & Lehtinen, E. (2014). How expertise is created in emerging professional fields. In S. Billett, T. Halttunen, & M. Koivisto (Eds.), *Promoting, assessing, recognizing and certifying lifelong learning: International perspectives and practices* (pp. 131-150). New York: Springer.
Patel, V. L., Arocha, J. F., & Zhang, J. (2005). Thinking and reasoning in medicine. In K. J. Holyoak (Ed.), *The Cambridge handbook of thinking and reasoning* (pp. 727-750). Cambridge, MA: Cambridge University Press.
Patel, V. L., & Groen, G. J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, *10*, 91-116. doi:10.1207/s15516709cog1001_4

Quesada-Pallarès, C., & Gegenfurtner, A. (2015). Toward a unified model of motivation for training transfer: A phase perspective. *Zeitschrift für Erziehungswissenschaft*, *18* (Suppl. 1), 107-121. doi:10.1007/s11618-014-0604-4

Säljö, R. (2009). Learning, theories of learning, and units of analysis in research. *Educational Psychologist*, *44*, 202-208. doi:10.1080/00461520903029030

Säljö, R. (2012). Literacy, digital literacy and epistemic practices: The co-evolution of hybrid minds and external memory systems. *Nordic Journal of Digital Literacy*, *7*, 5-19.

Sangin, M., Molinari, G., Nüssli, M., & Dillenbourg, P. (2008). How learners use awareness cues about their peer knowledge? Insights from synchronized eye-tracking data. *International Conference of the Learning Sciences Proceedings.*

Sauseng, P., & Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience and Biobehavioral Reviews*, *32*, 1001-1013. doi:10.1016/j.neubiorev.2008.03.014

Seppänen, M., & Gegenfurtner, A. (2012). Seeing through a teacher's eyes improves students' imaging interpretation. *Medical Education*, *46*, 1113-1114. doi:10.1111/medu.12041

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, *27*, 4-13. doi:10.3102/0013189X027002004

Siewiorek, A., & Gegenfurtner, A. (2010). Leading to win: The influence of leadership style on team performance during a computer game training. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the disciplines: ICLS2010 proceedings* (Vol. 1, pp. 524-531). Chicago, IL: International Society of the Learning Sciences.

Simpson, S. A., & Gilhooly, K. J. (1997). Diagnostic thinking processes: Evidence from a constructive interaction study of electrocardiogram (ECG) interpretation. *Applied Cognitive Psychology*, *11*, 543-554. doi:10.1002/(SICI)1099-0720(199712)11:6<543::AID-ACP486>3.0.CO;2-C

Slack, R., Hartswood, M., Procter, R., & Rouncefield, M. (2007). Cultures of reading: On professional vision and the lived work of mammography. In S. Hester & D. Francis (Eds.), *Orders of ordinary action. Respecifying sociological knowledge* (pp. 175-193). Aldershot: Ashgate.

Smith, P. L., & Ratcliff, R. (2009). An integrative theory of attention and decision-making in visual signal detection. *Psychological Review*, *116*, 283-317. doi:10.1037/a0015156

Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. In M. Lampert & M. L. Bunk (Eds.), *Talking mathematics in school: Studies of teaching and learning* (pp. 107-149). New York: Cambridge University Press.

Suchman, L. (2007). *Human-machine reconfigurations. Plans and situated actions* (2nd ed.) Cambridge, MA: Cambridge University Press.

Szulewski, A., Gegenfurtner, A., Howes, D., Sivilotti, M., & Van Merriënboer, J. J. G. (in press). Measuring physician cognitive load: Validity evidence for a physiologic and a psychometric tool. *Advances in Health Sciences Education.* doi:10.1007/s10459-016-9725-2

Szulewski, A., Kelton, D., & Howes, D. (2017). Pupillometry as a tool to study expertise in medicine. *Frontline Learning Research*. doi:10.14786/flr.v5i3.256

Van de Wiel, M. (2017). Examining expertise using interviews and verbal protocols. *Frontline Learning Research*. doi:10.14786/flr.v5i3.257

Williams, M. A., Fawver, B., & Hodges, N. J. (2017). Using the 'expert performance approach' as a framework for examining and enhancing skill learning: Improving understanding of expert learning. *Frontline Learning Research*. doi:10.14786/flr.v5i3.267