The Many Dimensions of Having a Good Eye: A Methodological Reflection of Metaphors in Visual Cognition Analysis

Andreas Gegenfurtner, Anna Siewiorek, Centre for Learning Research, University of Turku, Finland
Email: angege@utu.fi, amsiew@utu.fi

Abstract: There seems to be a stable and widelyheld belief among medical practitioners that good diagnosticians have a good eye that is innate, rather than trained. However, there is no evidence for any identifiable perceptual trait. Nonetheless, the process of developing a good eye in medicine 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 cognition in diagnosing medical images. The purpose of this paper is to provide a reflection on this methodological pluralism. 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 what it implies for current epistemological, methodological, and curricular discussions in medical education.

Introduction

There seems to be a stable and widely held belief among medical practitioners that good diagnosticians are born rather than made (Elstein, Shulman, & Sprafka, 1978; Ericsson, 2004; Norman, Coblentz, Brooks, & Babcook, 1992). Indeed, it is a common observation, both in training programs and in clinical practice, that some residents are quicker and more accurate than others in recognizing subtle abnormalities in medical images. Based on everyday observation of these kinds, it has subsequently been concluded that diagnostic accuracy is due to having a "good eye"—one that is innate, rather than trained. However, there is no evidence for any identifiable perceptual trait. To the contrary, recent research has indicated that what has been labeled as a good eye can in fact be developed through training. For example, Dawes, Vowler, Allen, and Dixon (2004) demonstrated that detection performance of clinical students was significantly higher after a 26-week training period than it had been before; materials included graphics used in diverse medical disciplines such as plain X-ray films, ultrasound, computer tomograms, magnetic resonance images, and positron emission tomograms. Similar training effects have also been reported in cardiology (Issenberg, McGaghie, Gordon, Symes, Petrusa, Hart, et al., 2002), radiology (Brazeau-Lamontagne, Charlin, Gagnon, Samson, & Van der Vleuten, 2004), and microscopic pathology (Crowley, Lecowski, Medvedeva, Tseytlin, Roh, & Jukic, 2006). Overall, we can conclude that training can improve visual accuracy; that recognition of abnormities in perceptual tasks can be learned; and that good diagnosticians are made rather than born

Thus, there is consensus that the development of being a good diagnostician goes in concert with the development of a good eye. What is less clear is what a good eye is. What precisely is it that develops in novice diagnosticians through medical education? What must be appropriated to move toward higher accuracy? Some investigators, 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 diagnosing is mainly a cognitive inference that aligns disease 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; Kundel, 2007; Norman et al., 1992; 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 (Haller & Radue, 2005), and is accomplished through situated social discourse (Koschmann, LeBaron, Goodwin, Zemel, & Dunnington, 2007). In short, the metaphor 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 cognition.

The main task of the present study is to provide a first step in proposing a reflection on this methodological pluralism. Our goal is to present a framework in which the many dimensions of having a good eye can be considered as mutually constituting the richness we have on ideas, concepts, and theories relating to medical image diagnosis. 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 study will provide such a simple answer. The study addresses the questions of how to analyze visual cognition, how to elicit tacit assumptions underlying common research practice; and how to inform curriculum design for medical training in corporate and higher education settings. By discussing four examples drawn from radiology, we identify four metaphors that constitute our framework on visual cognition analysis. Despite its groundwork in medical literature, it could be assumed that the framework also has implications beyond medicine, due to its common interest in learning and comprehension in technology-rich vision-intensive contexts. In the first part of this paper, we introduce four metaphors for visual cognition: activation, detection, inference, and practice. In the second part, we discuss the potential implications of these metaphors in reconsidering standpoints on epistemology, methodology, and curriculum design.

Four Dimensions of Visual Cognition Analysis: A Metaphorical Mapping

Metaphors are a useful tool for mirroring in simple terms the often complex fundamentality 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 is the reflection on learning by use of the metaphors acquisition, participation, and knowledge creation (Paavola, Lipponen, & Hakkarainen, 2002; Sfard, 1998). 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.

Below, we identify, exemplify, and discuss four metaphors often used when analyzing visual cognition; these metaphors are seen as four of the many dimensions possessed by one with "a good eye" in medicine. We reflect on these metaphors in terms of what they contribute as an answer for the two questions stated above: The nature of what develops in novice diagnosticians through medical education is not yet clear. The question of what must be appropriated to move toward higher accuracy remains to be answered. 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.

Table 1: A comparative metaphorical mapping of visual cognition analysis in medicine.

	Activation	Detection	Inference	Practice
Indicators of visual cognition	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	Seconds	Minutes to few hours	Minutes to decades
Associated methodology	Cognitive neuroscience	ROC analysis; eye tracking methodology	Protocol analysis	Ethnomethodology

Activation Metaphor

Research adhering to the activation metaphor uses measures of neurophysiologic activity as an indication of visual cognition. 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 occurs. Hence, visual cognition 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 cellular oxygen consumption. Whenever information stored in neural networks is used for cognitive processes, neural activation occurs that can be measured with one of those tools. For example, if a radiologist formulates a diagnosis based on a patient's fMRI scan, 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 cognition has rarely been used in medical diagnosis studies. However, the coming together of the learning sciences 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). An answer comes into play regarding the question of what develops in novice diagnosticians through medical education; namely, that neuroscience has the potential to trace implicit and experiential learning before it can be observed in behavior. This will allow us to 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 expertise development in novice diagnosticians. Results derived from the activation metaphor therefore have the potential to inform the design of medical education and training; more on this issue will be offered at a later point in this paper.

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, 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. Visual cognition 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. This empirical evidence is consistent with many studies that adhere to the detection metaphor in radiology (Mello-Thoms, 2003), cardiology (Augustyniak & Tadeusiewicz, 2006), and microscopic pathology (Krupinski et al., 2006). 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, specificity, and accuracy of the detection performance. Usually, the analytic time span is 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. Traditionally, eye-tracking studies have focused on individual observers as the unit of analysis; however, recent developments of eye-tracking technology and of new 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.

The 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), 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), SUSTAINs (Love, Medin, & Gureckis, 2004), or PANDEMONIUM processors (Selfridge, 1959). In the next paragraph, we describe a more recent example on the correlation of vision and cognition in medical image diagnosis that nonetheless 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 cognition. 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 cognition 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 on visual cognition 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 subjects and these cannot account for the underlying implicit, non-conscious processing of these subjects. 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). Due to their intrinsic method constraints, it thus seems reasonable to aim toward combining protocols with other measures, and hence to avoid the risk of using one single metaphor only. More on this issue will be elaborated below.

The Practice Metaphor

Finally, the last metaphor we identify as being frequently used in visual cognition analysis in medicine is the practice metaphor. 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 (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. Livingstone, 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 cognition (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 this contribution.

To further illustrate the methodological entailments of the practice metaphor, we describe a recent 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 cognition in Goodwin's terms, and they are negotiated around a common object of disciplined perception (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 cognition 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 role of technology in medical image diagnosis (Burri & Dumit, 2008; Gegenfurtner, 2009). It is essentially the focus on embodied talk-in-interaction—talk between people and between humans and non-human objects (Gibson, 1979; Suchman, 2007)—that makes the practice metaphor a useful tool to analyze the fourth perspective; namely, what constitutes visual cognition, what medical students have to appropriate to move toward diagnostic accuracy, and how medical educators act to help develop this higher level of accuracy (see, e.g., Koschmann et al., 2007; Slack et al, 2007).

Implications and Conclusions

This paper has presented and discussed a mapping of visual cognition analysis in medical education centered on four metaphors: activation, detection, inference, and practice. Reflection on these metaphors can give significant answers in three areas, namely: epistemology, methodology, and curriculum design. First, it provides an epistemological answer on the ongoing debate regarding the vision-cognition divide. Are perceptual and cognitive processes independent of each other? Or is the separation between vision and cognition artificial, a mere analytical distinction? Specifically, many studies in medical research adhering to the detection metaphor include no additional measures that are more prevalent in the inference metaphor, and vice versa. This neglect to combine measures for less fallibly accounting for visual cognition indicates the (maybe tacit) assumption in the medical literature that vision and cognition are independent. An assumption of this kind results from a way of thinking informed by an information processing view of cognition. This view stands in contrast to an ecological view on visual perception and cognition (Gibson, 1979), in which seeing, thinking, and acting are mutually interactive and nondualistic. It is questionable which approach is better: the information processing approach or the ecological approach. We argue here, as we have at the beginning of this paper, that it is not a question of truth or value judgment. Different methods lead to different answers based on different indicators. 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.

The second answer that is informed by the reflection on different metaphors on visual cognition concerns the dangers of choosing just one. Sfard (1998) indicated that a combination of learning metaphors yields to more robust findings than does a non-combination. 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. As a first step, we briefly list three initial approaches for method synergy. First, eye-tracking and verbal reports can be combined to trace the interaction of perceptual and conceptual processing of visual information. Second, eye-tracking and EEG/fMRI recording can be combined to assess the first seconds of scanning an image—a period that is crucial, especially for higher levels of expertise. Third, video-recording and eye-tracking can be combined to measure the relationship between (1) mutual gaze patterns and (2) collaboration patterns as indicators of distributed visual cognition (Sangin et al, 2008; Stahl, 2006). In sum, synergies between metaphors need to be explored to avoid the dangers associated with choosing just one metaphor.

Finally, the third answer relates to curriculum design. It is thanks to the variety of metaphors available that we can illuminate ways to remake curricula in medical education. Specifically, the activation metaphor and its related methodologies inform the high speed with which visual information is processed, in the range of milliseconds; this cannot be captured with other methods. It also indicates the automaticity of information processing. Together with eye-tracking research, these findings indicate the role of implicit learning. In addition, the practice metaphor highlights the role of hands-on activities in interaction with the technological tool that is not merely used as a stimulus, but rather as a mediating object in the learning process. Learning to handle the tool is considered as important for medical education as is learning to discriminate features in the image; hence, curricula in medical schools can aim (or maybe have aimed) to also develop epistemic practices and professional identity. In summary medical curricula aimed at fostering implicit learning, identity formation, and participation can combine classroom-based direct instruction on disease etiology, learning in the lab with authentic tools, and work-based learning in real-life clinical workplaces, in ways that best support the development of "a good eye" in medial image diagnosis.

The ICLS 2010 conference theme invited an exploration of the ways disciplinary perspectives can inform the study of learning in educational or workplace settings. This theme—elaborated in the call for proposals—suggests an awareness of the multispectral image of learning caused by different, discipline-specific thinking traditions and their associated research custom. This paper has attempted to take up this theme; it did so by bringing together tradition and custom that have evolved over the years around a common topic: the analysis of visual cognition in medical image diagnosis. It can be argued that this disciplinary view, a medical education view, has implications for the study of learning that go beyond medicine. Specifically, in mathematics education, similar research practices can be identified that have as underlying inspiration the activation metaphor (e.g., when children focus spontaneously on number aspects during encoding of photos; Hannula, Grabner, & Lehtinen, 2009); the detection metaphor (e.g., when elementary students solve addition and subtraction word problems; De Corte, Verschaffel, & Pauwels, 1990); the inference metaphor (when schema-based instruction facilitates algebraic thinking for students with learning disabilities; Xin, 2008); or the practice metaphor (when participating in disciplined mathematical practices around coordinate systems; Stevens & Hall, 1998). These snapshots from mathematics education illustrate that the four metaphors identified in medical education literature might also have traces in other domains; if a more in-depth analysis of mathematics education research corroborates this initial snapshot finding, then this would answer one of the questions posed in the ICLS 2010 call for proposals—that is, what seems to be constant in learning processes (more specifically: in visual learning processes) beyond intrinsic disciplinary variations can be captured by different metaphors that highlight the multispectral image of our field: learning 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.

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

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

Augustyniak, P., & Tadeusiewicz, R. (2006). Assessment of electrocardiogram visual interpretation strategy based on scanpath analysis. *Physiological Measurement*, *27*, 597-608.

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.

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.

Brazeau-Lamontagne, L., Charlin, B., Gagnon, R., Samson, L., & Van der Vleuten, C. (2004). Measurement of perception and interpretation skills during radiology training: Utility of the script concordance approach. *Medical Teacher*, *26*, 326-332.

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.

Crowley, R. S., Lecowski, E., Medvedeva, O., Tseytlin, E., Roh, E., & Jukic, D. (2006). Evaluation of an intelligent tutoring system in pathology: Effects of external representations on performance gains, metacognition, and acceptance. *Journal of the American Medical Informatics Association*, *14*, 182-190.

Dawes, T. J. W., Vowler, S. L., Allen, M. C., & Dixon, A.K. (2004). Training improves medical student performance in image interpretation. *British Journal of Radiology*, *77*, 775-776.

De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the semantic structure of word problems on second graders' eye movements. *Journal of Educational Psychology*, *82*, 359-365.

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.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Cambridge University Press.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenane of expert performance in medicine and related domains. *Academic Medicine*, *79*, S70-S81.

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

Gegenfurtner, A. (2009). What is seen on the screen? Exploring collaborative interpretation, representational tools, and disciplined perception in medicine. In A. Dimitracopoulou, C. O'Malley, D. Suthers, & P. Reimann (Eds.), *Computer supported collaborative learning practices: CSCL2009 community events proceedings* (pp. 71-72). Rhodes: 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.

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.

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

Hannula, M. M., Grabner, R., & Lehtinen, E. (2009, March). *Neural correlates of spontaneous focusing on numerosity (SFON) in a 9-year-longitudinal study of children's mathematical skills*. Paper presented at the EARLI Advanced Study Colloquium on Cognitive Neuroscience meets mathematics education, Bruges, Belgium.

Issenberg, S. B., McGaghie, W. C., Gordon, D. L., Symes, S., Petrusa, E. R., Hart, I. R., et al. (2002). Effectiveness of a cardiology review course for internal medicine residents using simulation technology and deliberate practice. *Teaching and Learning in Medicine*, *14*, 223-228.

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.

Koschmann, T., LeBaron, C., Goodwin, C., Zemel, A., & Dunnington, G. (2007). Formulating the triangle of doom. *Gesture*, *7*, 97-118.

Koschmann, T., & Zemel, A. (2009). Optical pulsars and black arrows: Discoveries as occasioned productions. *Journal of the Learning Sciences*, 18, 200-246.

Krupinski, E. A., Tillack, A. A., Richter, L., Henderson, J. T., Bhattacharyya, A. K., Scott, K. M., et al. (2006). Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. *Human Pathology*, *37*, 1543-1556.

Kundel, H. L. (2007). History of research in medical image perception. *Journal of the American College of Radiology*, *3*, 402-408.

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.

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.

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.

Mello-Thoms, C. (2003). Perception of breast cancer: Eye-position analysis of mammogram interpretation. *Academic Radiology*, *10*, 4-12.

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science*, *325*, 284-288.

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.

Norman, G. R., Coblentz, C. L., Brooks, L. R., & Babcook, C. J. (1992). Expertise in visual diagnosis: A review of the literature. *Academic Medicine*, *67*, S78-S83.

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.

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.

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.

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In *The mechanisms of thought processes*. London: Stationery.

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

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.

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.

Stahl, G. (2006). Group *cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.

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.

Xie, Y. P. (2008). The effect of schema-based instruction in solving mathematics word problems: An emphasis on prealgrebraic conceptualization of multiplicative relations. *Journal for Research in Mathematics Education*, 39, 526-551.

Acknowledgments

This work was supported by Grant No. 128766 of the Academy of Finland; it is part of an interdisciplinary research project at the confluence of medical sciences and learning sciences investigating "Technologies for seeing and technologies for knowing: Learning and medical imaging (LearnMedImage)". In this project, research seldom is an individual endeavor, but more often is a collaborative achievement. Thus, we would like to thank Roger Säljö, Erno Lehtinen, Hans Rystedt, Laura Helle, Markus Nivala, and Minna M. Hannula for developing with us a good eye for metaphors with which to analyze visual medical cognition.