This is a pre-print of:

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

Halszka Jarodzka[1,*], Thomas Balslev[2], Kenneth Holmqvist[3], Marcus Nyström[3], Katharina Scheiter[4], Peter Gerjets[4], & Berit Eika[5]

Conveying clinical reasoning based on visual observation via eye-movement modelling examples

[1]Open University of the Netherlands, Center of Learning Sciences and Technologies, the Netherlands; [2]Viborg Regional Hospital, Department of Pediatrics, Denmark; [3]Lund University, Humanities Lab, Sweden; [4]Knowledge Media Research Center, Germany; [5]Aarhus University, Center of Medical Education, Denmark

*Correspondence concerning this paper should be addressed to Halszka Jarodzka, Centre for Learning Sciences and Technologies, Open University of the Netherlands, P.O. Box 2960, 6401 DL, Heerlen, the Netherlands, T: +45-5762410, F: +45-5762800, E: Halszka.Jarodzka@OU.nl

Running head: Eye-movement modelling of clinical reasoning

This manuscript is part of the Special Issue: Collaborating with digital tools and peers in medical education. Cases and simulations as interventions in learning. Edited by Helle and Säljö

Abstract

Complex perceptual tasks, like clinical reasoning based on visual observations of patients, require not only conceptual knowledge about diagnostic classes but also the skills to visually search for symptoms and interpret these observations. However, medical education so far has focused very little on how visual observation skills can be efficiently conveyed to novices. The current study applied a novel instructional method to teach these skills by showing the learners how an expert model visually searches and interprets symptoms (i.e., eye-movement modelling examples; EMMEs). Case videos of patients were verbally explained by a model (control condition) and presented to students. In the experimental conditions, the participants received a recording of the model's eye movements superimposed on the case videos. The eye movements were displayed by either highlighting the features the model focused on with a circle (the circle condition) or by blurring the features the model did not focus on (the spotlight condition). Compared to the other two conditions, results show that a spotlight on the case videos better guides the students' attention towards the relevant features. Moreover, when testing the students' clinical reasoning skills with videos of new patient cases without any guidance, participants studying EMMEs with a spotlight showed improved their visual search and enhanced interpretation performance of the symptoms in contrast to participants in either the circle or the control condition. These findings show that a spotlight EMME can successfully convey clinical reasoning based on visual observations.

Keywords: example-based learning, eye tracking, expertise, attention, medical education

In clinical practice, physicians often need to detect and diagnose disorders based upon visual observations of patients' symptomatic behaviour, such as occasionally occurring ictal (i.e., seizure induced) movement patterns of infants' epileptic seizures. In such cases, it is crucial to carefully observe which parts of the infant body are involved and how and when these body parts move, and to assess the level of consciousness of the infant (i.e., diagnosis based on the semiology of ictal events, cf. International League Against Epilepsy 2010; Lüders et al. 1998). According to Nordli et al. (1997), infantile seizures are often so subtle that even skilled observers may fail to identify the seizures correctly. Moreover, some of these movement patterns occur only occasionally, are short-term and subtle, and therefore not salient compared to other movements or characteristics of the infant (Hansen and Balslev 2009). As a result, an infant's epileptic seizures may easily be mistaken for normal behaviour (Egger et al. 2003; Hansen and Balslev 2009) or attributed to other disorders (Nordli 2002).

Clinical reasoning based on visually observing a patient can be seen as a knowledge-rich (Van Lehn 1989) and highly perceptual task (Chi 2006) that requires thorough and intensive processing of the visual input. Hence, such a task requires prior knowledge that needs to be acquired through training. Importantly, this task requires not only conceptual knowledge about diagnostic classes but also skills related to visually searching for symptoms and to correctly interpreting these observations (cf. Krupinski 2010). These skills that are the focus of this paper will be termed visual observation skills hereafter.

Role of visual observations in clinical reasoning

Many studies have been conducted to determine what skills are required for accomplishing complex visual tasks which distinguish experts from novices. These studies showed that experts possess sophisticated visual observation skills which enable them to visually search relevant features of a stimulus within irrelevant features and to interpret these observations correctly (e.g., Antes and Kristjanson 1991; Charness et al. 2001; Jarodzka et al. 2010a; Underwood et al. 2003). Novices, however, get easily distracted by salient, but potentially irrelevant, elements of a visual stimulus, whilst missing relevant information (Jarodzka et al. 2010a; Lowe 1999).

These effects have also been demonstrated in the medical domain, for instance, in clinical reasoning and diagnostic classification based on static and rather abstract stimuli, such as X-rays, microscopic slides, and mammograms (e.g., Krupinski 2005; Krupinski et al. 2006; Kundel et al. 2008; Lesgold et al. 1988) as well as based on realistic static stimuli, such as photographs of patients (Brooks et al. 2000). Although no studies have yet investigated which

differences in visual observation skills exist between experts and novices when processing dynamic stimuli in the medical domain such as patient video cases, expertise differences in visual observation skills have been shown for dynamic, realistic stimuli in other domains, such as biological classification of motion patterns (Jarodzka et al. 2010a) or car driving (Underwood et al. 2003).

Novices experience difficulties in tasks such as clinical reasoning based on visual observation of static medical images as well as in other domains using dynamic stimuli, whereas experts seem to possess rather sophisticated visual observation skills that enable them to perform these tasks. Thus, it can be assumed that these expertise differences will also hold true for a medical task based on dynamic stimuli, namely, diagnostic classification of infants' epileptic seizures. Thus, it is proposed that novices will require training in clinical reasoning skills based on visual observations to appropriately perform this diagnostic task. The next section addresses the question of how to convey these skills.

Teaching clinical reasoning based on visual observations

Although clinical reasoning based on visual observations is likely to require not only knowledge about diagnostic classes but also the skills to visually search for symptoms and interpret these observations (Krupinski 2010), the focus in medical education does not specifically lie in teaching such skills (for first exceptions in microscopy, see Helle et al. in press). Instead, motion patterns of epilepsy are often described in text accompanied by static pictures (e.g., Schmidt and Schachter 2000). However, written text is not adequate for teaching the skills of clinical reasoning based on visual observation (Balslev et al. 2009). To overcome the drawbacks of static representations, medical education often makes use of so-called patient video cases (PVCs; e.g., Dequeker and Jaspaert 1998). PVCs show patients who are video-recorded whilst displaying movement which is potentially abnormal. Using PVCs for educational purposes has been shown to improve clinical reasoning compared to traditional written descriptions of patient cases (Balslev et al. 2009; De Leng et al. 2007; Kamin et al. 2003). However, based on the research described in the last section, one may have serious doubts whether novices would be able to perceive relevant features and interpret them correctly when inspecting PVCs without any further guidance.

Hence, in this study we wanted to explore whether visually cueing relevant information in PVCs in different modes with *eye-movement modelling examples* could improve clinical reasoning based on visual observations.

Teaching visual observation skills via eye-movement modelling examples

Eye-movement modelling examples (EMMEs) have been developed to teach performance on perceptual tasks (Jarodzka et al. 2010b; Van Gog et al. 2009). EMMEs are worked-out examples of how an expert model performs a perceptual task, where a video case (e.g., an infant with ictal movements) is augmented with two types of procedural information. This information is gained from asking an expert to explain verbally how he or she performs the task based on watching the video case as well as from recording his or her eye movements whilst performing the task. The actual instructional examples shown to the learners consist of a video augmented with the expert model's verbal explanation (i.e., modelling of cognitive processes) as well as the model's eye movements superimposed onto the video (i.e., modelling of perceptual processes). The potential of EMMEs is based on three features, namely, the fact that they (a) are grounded in research on learning from examples, (b) allow covert cognitive and perceptual processes to be made accessible for learners, and (c) provide guidance in how to process the visual input. We take up each of these aspects in turn in the following.

The rationale for developing EMMEs is derived from research on learning from worked-out examples showing that example-based learning is a powerful method for early skill acquisition. Examples demonstrate a problem solution to the learner, either by presenting a written, worked-out solution procedure of a problem (i.e., worked-out examples; see, Atkinson et al. 2000; Sweller et al. 1998) or by having the learner directly observe an expert model performing the task live or on video (i.e., modelling examples; Bandura 1977; Collins et al. 1989). For cognitive aspects of a task, where covert cognitive processes are crucial, the model verbalises his or her cognitive processes by explaining why and how certain solution steps were chosen, i.e., cognitive modelling (Collins et al. 1989) and process-oriented worked-out examples (Van Gog et al. 2006).

For perceptual aspects of a task which require the inspection of a visual stimulus (Chi 2006), it is crucial that the learner can see the stimulus the expert is looking at. To actually learn from such examples, however, learners have to attend to the relevant features of the modelled behaviour (Bandura 1977). Otherwise, the learners may miss those visual features relevant for understanding and learning from the example. However, as indicated by the research presented above, novices experience difficulties when dealing with complex visual stimuli. Thus, when learning from examples which involve visual stimuli, novices might need attention guidance to those features that the model is focusing on.

One common form of attention guidance in instructional design is *cueing*. Although many studies on cueing exist, these studies have often failed to find positive effects on learning (for a review, De Koning et al. 2009). One reason might be that the cues used are usually based on what area an instructional designer or domain expert *thinks* a novice learner at a particular moment should be looking at. Experts, however, may not necessarily know what the most important feature for novices is, since experts often have difficulties in correctly assessing the knowledge state of novices (Hinds 1999; Nückles et al. 2006). Moreover, experts may not be able to report what they are looking at whilst performing a task, because their task performance has become highly automatised (Chi 2006). Thus, simply asking experts which information should be cued might not be the appropriate approach. One cueing technique which did prove successful is the one used by Grant and Spivey (2003), who showed that choosing a cue based on eye movements of successful problem-solvers enhanced the probability of correctly solving an insight problem. However, the task used by Grant and Spivey (2003) consisted of a very simple stimulus (a circle with a larger dot in it to represent Duncker's radiation problem). Hence, it could be easily detected which feature the successful problem-solvers focused on in order to cue this clearly distinguishable feature (i.e., the circle). In the medical domain, however, the visual stimuli are far more complex. Visual features are not easily distinguishable, and more importantly, *several* visual features need to be inspected and compared. In this case, it might be important to cue an entire *process* and not only a single visual feature. In line with this reasoning, Litchfield et al. (2010) conveyed information regarding an expert model's perceptual processes by displaying the model's eye movements superimposed onto a complex medical image, namely X-rays. Results showed that this type of display improved other participants' diagnostic performance based on these X-rays. Hence, it seems advisable to externalise a model's covert perceptual processes as a means of improving task performance.

The question remains, however, whether the effectiveness of attention guidance based on a model's eye movements extends beyond improving the performance of the task at hand, by also enhancing *learning*. Here, learning refers to the robust change in a person's skills that enable him or her to independently perform that task after practice with novel stimuli and without any instructional support (Simon 1983). EMMEs directly aim at investigating this research question by combining insights of example-based instruction for cognitive and perceptual aspects of the task. EMMEs do so by comprising three components: the problem (i.e., the to-be-inspected visual stimulus), the externalisations of cognitive processes (i.e.,

verbal explanations), and the perceptual processes (i.e., eye movements) of the model, where in particular the latter provide attention guidance for how to process the visual stimulus.

An unresolved issue so far is how to display the eye movements in modelling examples. Eye movements can be presented either by displaying them as *additional* information superimposed onto the stimulus (e.g., as a dot on the features the model focused on) or by *reducing* the existing visual information of the stimulus. In the latter case, the eye movements act like a spotlight, whereas the features that the expert model did NOT focus are blurred, thereby automatically highlighting the focused, non-blurred information. The first studies with EMMEs in relatively simple tasks (an isomorph of the Tower-of-Hanoi: Van Gog et al. 2009; and a biological classification task: Jarodzka et al. 2010b) have provided interesting insights that may inform the design of EMMEs in that respect. However, displaying eye movements directly as additional information has been shown to impose such high cognitive demands on learners that the demands may even hamper task performance (Van Gog et al. 2009). In contrast, a dot-like representation of eye movements has also been shown to increase skills related to the interpretation of relevant visual features, but without having a positive effect on visual search (Jarodzka et al. 2010b). Reducing existing information by blurring non-focused information instead has been shown to guide attention and to foster visual search, whereas no improvements in interpretation performance could be observed (Dorr et al. 2010; Jarodzka et al. 2010b). Moreover, blurring videos has been shown to guide the observers' attention, without them even noticing it (Nyström 2008). Thus, no display type has yet proven to be superior with regard to all of the aforementioned performance aspects.

Research questions – Hypotheses

Against the backdrop of the reported literature, the current study aimed to answer two questions, namely, can displaying an expert's eye movements in modelling examples be beneficial for learning, and if so, which type of display is more effective in terms of attentional guidance and learning? To investigate these questions, a control condition, consisting of video examples augmented with verbal explanations of an expert model only, was compared to two experimental conditions, which additionally included the replay of an expert model's eye movements superimposed on the video. Two different versions of EMMEs were implemented: a display that adds information (circle display) and one that reduces information (blurring display). Instead of the solid dot used in prior studies (Jarodzka et al. 2010b; Van Gog et al. 2009), the expert model's eye movements were displayed as a circle so that the cue would not cover relevant information. For the spotlight display, we

decided to use a very subtle blurring so that an impression of the overall scene could be gained. The current study used a diagnostic classification task based on video cases of infant patients with a tentative diagnosis of an epileptic seizure as a domain.

The following hypotheses were tested in this study:

We assumed that a prerequisite for learning from EMMEs is to successfully guide learners' attention to where the expert model focused on. Hence, Hypothesis 1 was as follows:

a) Participants' learning with both types of EMMEs will be guided by the eye-movement display, resulting in a scanpath that is highly similar to the expert model's scanpath, whereas the scanpaths of participants in the control condition will be more diverse.

b) It is not clear which type of EMME should guide learners' visual attention better to where the expert model focused on. Since the spotlight display is hardly intrusive (cf. Nyström 2008), and the circle display does not obscure any elements, both should strongly guide attention in similar ways.

Moreover, we assumed that EMMEs would foster clinical reasoning based on visual observation as measured by the efficiency of visual search during a test (by means of eye tracking), and the interpretation of the motion pattern observed (by means of a multiple-choice questionnaire). Hence, Hypothesis 2 addressing the visual search variables was as follows:

a) The EMMEs conditions will learn better what to focus on, thereby yielding a more efficient visual search than the control condition. Efficient visual search is indicated by focusing faster and longer on relevant areas, whilst ignoring irrelevant areas on novel test videos.

b) Since both displays should guide learners' attention equally well to the relevant features, it is unclear whether the two displays will differ with respect to visual search.

Hypothesis 3 addressing the interpretation of test video cases was as follows:

a) EMMEs will result in an enhanced interpretation performance of the infant's motion patterns compared to the control condition.

b) Since the spotlight display had been improved compared to a prior study, the spotlight display is assumed to be at least equally well suited as the circle display for fostering interpretation of the performance.

Method

Participants and design

Participants were 60 medical students in their final year at the University of Aarhus (41 women, 19 men, M_{age} = 26.57 years, *SD* = 2.03). All participants had no prior knowledge of the task and had normal or corrected-to-normal vision. They were randomly assigned to one of three conditions (*n* = 20 each), a control condition, an EMME condition with the spotlight display, and an EMME condition with the circle display. Participants were paid €10 for their participation.

Apparatus and materials

Eye-tracking equipment

The model's eye movements were recorded with an SMI High Speed eye-tracking system with a temporal resolution of 240 Hz on a 17" monitor and iView X 2.2 software. This eye-tracking system has a high accuracy of less than 0.5 degrees of visual angle and a high precision of less than 0.01 degrees. The model's eye-tracking data were edited with BeGaze 2.3 software (www.smivision.com) and self-programmed MatLab algorithms. The same system was used to record participants' eye movements. The participants' eye-tracking data were analysed with BeGaze 2.3 software (www.smivision.com) and self-programmed MatLab algorithms. All video material was presented to the participants via Experiment Center 2.2.

Modelling examples

Diagnosing infants' epileptic seizures may result in many different diagnostic classes. This study, however, does not focus on the final diagnostic step, but rather on the clinical reasoning underlying this step. Hence, it was not the aim to train students to diagnose all possible seizure types, but rather to train them in the procedure of collecting observations of symptoms crucial for a diagnosis. Thus, for the purpose of the current study, two exemplary patient cases of focal seizures were chosen to train clinical reasoning based on visual observations. These modelling examples consisted of two digital PVCs in an audio interleave format (.avi), sized 720 * 576 pixels and presented in full screen on a 1280 * 1024 pixels resolution (corresponding to 17.07 * 13.65 inches). Each PVC depicted a single infant (3 weeks and 7 months old, respectively). Both infants deployed a motion pattern corresponding

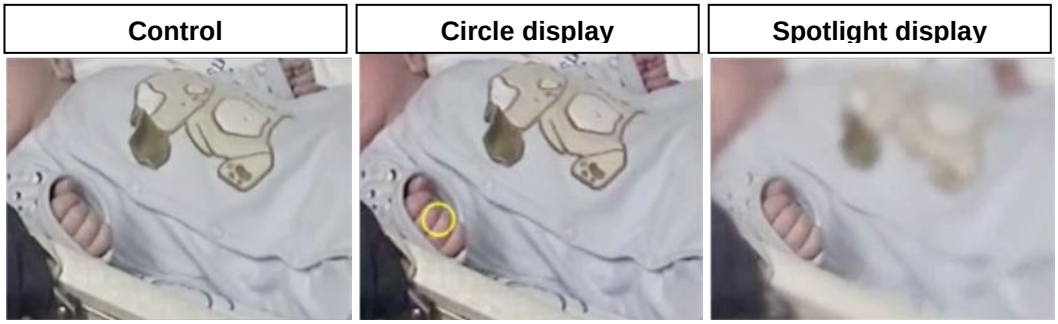


Figure 1: Screenshots from the three conditions used in the study.

Dependent variables

The dependent variables were divided into measures obtained during the learning phase and measures obtained during the subsequent test phase, where participants had to apply their learnt visual observation skills to novel test videos.

Learning phase. In order to estimate whether participants had been able to follow the attention guidance, their eye movements were recorded during watching the instructional videos (i.e., during the learning process). Based on these recordings, the *degree of similarity of the participants' eye movements to the eye movements of the model* displayed on the EMME videos was determined. This was done by calculating the Euclidean distance in pixels between simultaneous gaze points of the model and the participant over time (cf. Rao et al. 2002). Subsequently, the mean Euclidean distance was calculated for each participant. The smaller the distance, the better the learner's eye movements were synchronised with those of the expert.

Test phase. To test learning effects, participants were shown three new PVCs without verbal explanations or eye-movement replays for a mean duration of 29.67 s each (*SD* = 17.79). These PVCs depicted different children each displaying a particular seizure. The duration of the test videos depended on the duration of the seizure. Whilst the participants watched the test videos, participants' eye movements were recorded to determine the participants' visual search behaviour. Afterwards, their interpretation performance was assessed by asking them to answer multiple-choice questions.

Three different measures of learning were obtained, whereby two of them related to the efficiency of visual search, and the third one assessed the ability to interpret areas important for the diagnosis of epileptic seizures.

From the eye-tracking recordings, two different eye-movement measures were obtained as indicators of visual search efficiency. First, *time spent until participants had looked at all relevant areas of interest (AOIs) on the test videos* was recorded. The relevant areas were determined a priori with the help of a domain expert (cf. Antes and Kristjanson 1991; Charness et al. 2001). Each video included several AOIs that had to be looked at to describe the infant's disorder exhaustively. To obtain the time it took participants to have focused on all of the relevant AOIs, only the time that participants spent "outside" of the relevant AOIs until finally each AOI had been looked at was taken into account. Thus, this measure was determined by subtracting the time until all relevant AOIs were looked at for the first time by the time spent looking at other relevant AOIs in the meantime.

Second, the *time spent looking at the relevant areas on the test videos* (total dwell time) was recorded. Since the testing videos were of different durations, both eye-tracking measures were normalised by the video length.

The participants' *interpretation performance* was assessed by asking them to answer each of the three PVCs five multiple-choice questions relevant to the interpretation of the infant's behaviour as being symptomatic or not. The test was developed with the help of domain experts. In particular, the learners had decide (a) which body parts move, (b) how these body parts move, (c) whether touching the infant would change these movements, (d) whether and how the infant's face is involved and what this means in terms of a diagnosis, and (e) what the infant's state of consciousness was. Participants were assigned 1 credit for each question in the case of a correct answer, yielding a maximum of 5 credits for each test video (15 credits in total). The performance of each participant was transformed into a percentage score for ease of interpretation.

Procedure

Students were tested in individual sessions of approximately 30 minutes each. At the beginning, participants filled in a questionnaire on their prior knowledge of paediatric epilepsy and their demographic data. Then the participants received a short introduction to the topic, stating very general information on seizures. And finally the learning phase started. The eye-tracking system was adjusted to the individual features of the participant based on a 13-point calibration. Participants were told that they would subsequently be shown videos of the

to-be-learned disorder, in which a model explains the disorder shown in the video. Depending on the condition, the participants were also told that they would see where the model looked at on the video. Before watching the instructional videos, participants received information on the age and gender and a short problem description of the patient. Whilst the participants watched the corresponding instructional video, their eye movements were tracked.

In the test phase, participants watched three novel PVCs without verbal explanations or eye-movement replay. A fixation cross appeared for two seconds followed by the test video, which was replayed once. Participants watched the video whilst their eye movements were recorded. Afterwards, the video disappeared, followed by a blank screen. Then, the participants had to answer the multiple-choice questions. This procedure was repeated for the remaining two patient video cases in the test.

Results

All dependent variables were tested for normal distribution with the Kolmogorov-Smirnov test. Data for these variables which violated assumptions of normal distribution were logarithmised. Next, all statistical analyses were conducted using ANOVAs. For each dependent variable, an ANOVA was conducted with the condition (control vs. circle display vs. spotlight display) as the independent variable. To analyse differences between each condition, Bonferroni post-hoc tests were conducted.

Eye-movement data during example study

A Kolmogorov-Smirnov test revealed that the data for the Euclidean distance were not normally distributed ($D(60) = .12$, $p = .03$). Hence, these data were logarithmised. An ANOVA was conducted with mean Euclidean distance (logarithmised) as a dependent variable and condition (control vs. circle display vs. spotlight display) as an independent variable. The comparison of participants' eye-movement data with the model's eye-movement data during the example study showed a main effect of condition, $F(2, 57) = 4.40$, $p = .02$, $\eta_p^2 = .13$. Bonferroni post-hoc tests revealed that the spotlight display condition guided the participants' attention significantly more than in the control condition ($p = .03$) and than in the circle display condition ($p = .04$). The circle condition and the control condition did not differ significantly in their potential to guide attention, *ns*. Results are summarised in Figure 2.

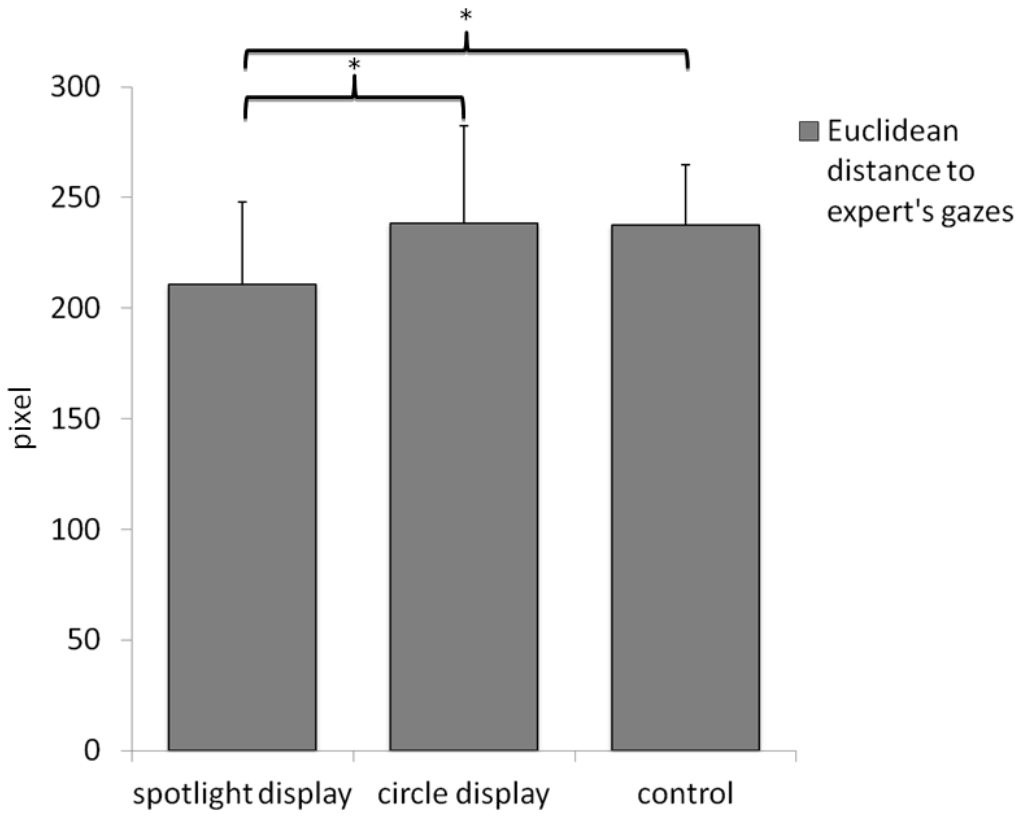
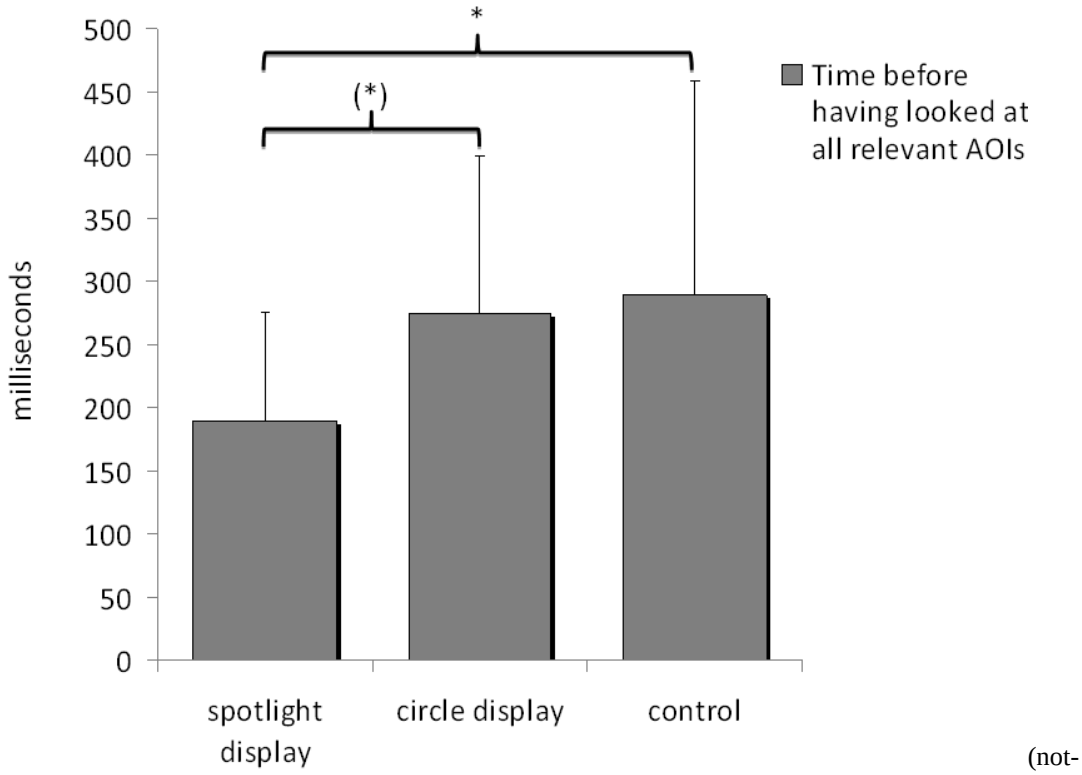


Figure 2: Means and standard deviations for the Euclidean distance (not-logarithmised in pixel) between the model's and participants' gaze points over time during the example study.

Learning outcomes

Eye-movement data during testing

A Kolmogorov-Smirnov test revealed that the data for the time before first having looked at all relevant AOIs were not normally distributed ($D(60) = .13$, $p = .02$). Hence, these data were logarithmised. An ANOVA was conducted with the time before first having looked at all relevant AOIs whilst watching the test videos (logarithmised) as a dependent variable and condition (control vs. circle display vs. spotlight display) as an independent variable. The conditions significantly differed in the time participants spent before having looked at all relevant AOIs once, $F(2, 57) = 4.06$, $p = .02$, $\eta_p^2 = .13$. Bonferroni post-hoc tests indicated that participants in the spotlight condition had looked significantly earlier at all relevant AOIs than in the control condition, $p = .04$ and marginally earlier than the circle condition, $p = .06$. The results for the participants in the circle condition and the control condition did not differ significantly, *ns*. Results are summarised in Figure 3.



(not-logarithmised in ms) during testing.

A Kolmogorov-Smirnov test revealed that the data for dwell time on relevant AOIs were normally distributed ($D(60) = .07$, $p = .20$). An ANOVA with dwell time on relevant AOIs as a dependent variable and condition (control vs. circle display vs. spotlight display) as an independent variable showed significant differences between the conditions in dwell time on relevant AOIs, $F(2, 57) = 3.64$, $p = .03$, $\eta_p^2 = .11$. Bonferroni post-hoc tests indicated that participants in the spotlight condition looked marginally longer at the relevant AOIs than participants in the control condition, $p = .07$, and than participants in the circle condition, $p = .07$. Results for the participants in the circle condition and the control condition did not differ significantly, *ns*. Results are summarised in Figure 4.

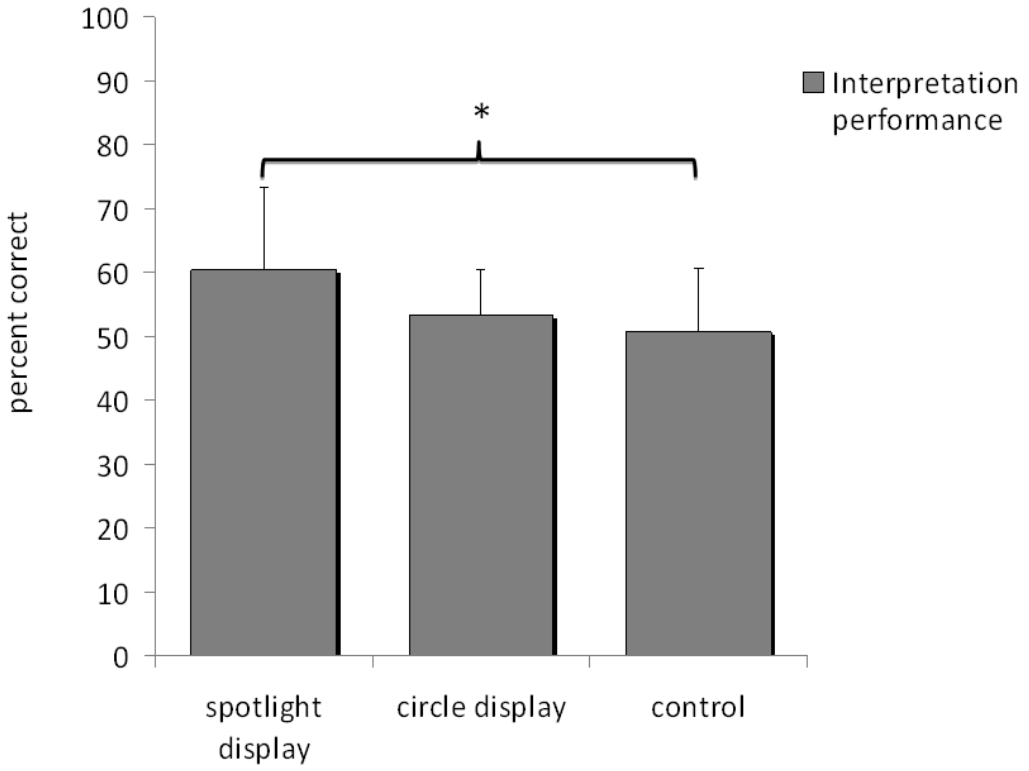


Figure 5: Means and standard deviations for interpretation performance (in percent correct) for the test videos.

Discussion

This study aimed at teaching clinical reasoning based on visual observations necessary for diagnosing infants' epileptic seizures with EMMEs. The EMMEs used in this study were based on recordings of a domain expert model, who was instructed to act as if teaching novices. These EMMEs consisted of two components, firstly, the model's verbal explanation of the symptoms accompanying the epileptic seizures he saw on PVCs (i.e., modelling of cognitive processes). Secondly, the EMMEs consisted of the model's eye movements superimposed onto the PVC (i.e., modelling of perceptual processes). Moreover, two ways of displaying the model's eye movements were implemented: a manufacturer-provided circle display (SMI) and a subtle spotlight display (cf. Nyström and Holmqvist 2008).

Results of this study showed that learners' attention can be guided by displaying the model's eye movements in PVCs. Analyses of the participants' eye movements during an example study indicated that the spotlight display of the model's eye movements guided learners' attention during learning more precisely than the circle display or no attention guidance (Hypothesis 1). Moreover, the results showed that this type of attention guidance fostered effective visual search (as measured with eye tracking; Hypothesis 2) and interpretation performance (as measured with a multiple-choice questionnaire; Hypothesis 3).

Participants in the spotlight condition outperformed participants in the circle display condition and the control condition with regard to both measures. Surprisingly, participants in the circle condition did not differ from the control condition in terms of either attention guidance or learning (i.e., efficiency of visual search and interpretation performance). In sum, this study was the first to show that EMMEs have the potential to promote clinical reasoning based on visual observation. In particular, blurring of irrelevant areas during learning may be the key to fostering learning.

This leads to the question why the two presentation types of EMMEs resulted in such different effects. The reason for the failure of the circle condition might lie in the fact that the circle cue did not help learners in *refraining* from focusing on salient but irrelevant features. Another possible explanation is that the circle display might have tempted the learners to try to "look through'" the small circle instead of focusing on the entire body part that the circle pointed at. This may have hampered learners in getting a holistic overview of the entire body part and therefore may have hindered interpretation of the information, which is at least partially based on relating neighbouring areas to each other (e.g., to see whether the infant's eyes are closed whilst the face is moving). In contrast, the spotlight display in this study did blur very subtly (in contrast to the one in the study by Jarodzka et al. 2010b); the spotlight guided the learners to the non-blurred areas but still allowed for an overview of the motion and shape of the entire object. This might have facilitated the understanding of the motion as explained by the model and, thus, enabled learning.

Irrespective of the promising results of this study, its main limitation has to be acknowledged: the effects were based only on three test videos. This is because analysis of eye-tracking data on videos requires very time-consuming data preparations, thereby limiting the number of stimuli that can be potentially used. However, the fact that these effects could be found even for such a low number of items hints towards the fact that these effects may be quite strong.

Implications for instructional practice

This study underlines the need to effectively guide visual attention when teaching clinical reasoning based on visual observation. EMMEs seem to provide an appropriate instructional format to address this need, provided that the way of displaying eye movements is optimised with regard to the processing demands imposed on the learner.

The question arises how to transfer the findings of this study into practice. What we have clearly seen is that some tasks (here: diagnosing infants' epileptic seizures) require

visual observation skills, which so far are not the focus of teaching. The same is probably true for several other tasks that require substantial visual observation as well (e.g., clinical reasoning based on other types of medical images, sports judging, playing computer games, air traffic control, car driving, etc.). When instructors teach novices to execute these kinds of tasks, the instructors should consider teaching visual observation skills as well. EMMEs might be an appropriate way to do this for many tasks. Of course, it is not always possible to use eye-movement recordings of experts, because eye-trackers are not easily accessible. Still, if teachers are aware of the fact that visual observation skills play a crucial role in the task they are teaching, they could use videos of the task and make their visual observation explicit, that is, point to the relevant areas and explain how they should be interpreted. Moreover, teachers, researchers, and developers of educational software could start exploring the benefits of blurring techniques to guide the attention of the learners. Furthermore, EMMEs could be used in online teaching where pointing is not possible (e.g., during a surgery). An eye-tracking device from which the novices receive the recordings might help them to learn new perceptual tasks quickly. These recordings could later be re-used for offline teaching situations.

In sum, this study showed that EMMEs might be helpful to better convey clinical reasoning based on visual observations in the task of diagnosing infants' epileptic seizures. The effects of EMMEs, however, might strongly depend on the task under investigation and the presentation used of the model's perceptual processes. This study suggests that the best effect – in terms of attention guidance during learning as well as visual search and interpretation performance during testing – might be achieved by an EMME that displays the perceptual processes of the expert model as a spotlight that subtly blurs the contrast of the areas the model did not focus on whilst the focused areas remain unaltered.

References

Antes, J. R., & Kristjanson, A. F. (1991). Discriminating artists from nonartists by their eye-fixation patterns. *Perceptual and Motor Skills, 73,* 893–894.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70,* 181–214.

Balslev, T., De Grave, W. S., Muijtjens, A. M. M., Eika, B., & Scherpbier, A. J. J. A. (2009). The development of shared cognition in paediatric residents analysing a patient video case versus a paper patient case. *Advances in Health Science Education, 14,* 557–565.

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

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

Brooks, L. R., LeBlanc, V. R., & Norman, G. R. (2000). On the difficulty of noticing obvious features in patient appearance. *Psychological Science, 11,* 112–117.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory and Cognition, 29,* 1146–1152.

Chi, M. T. H. (2006). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 167–184). Cambridge: Cambridge University Press.

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

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. Educational *Psychology Review, 21,* 113–140.

De Leng, B. A., Dolmans, D. H. J. M., Van de Wiel, M., Muijtjens, A. M. M., & Van der Vleuten, C. P. M. (2007). How video cases should be used as authentic stimuli in problem-based medical education. *Medical Education, 41,* 181–188.

Dequeker, J., & Jaspaert, R. (1998). Teaching problem-solving and clinical reasoning: 20 years experience with video-supported small-group learning. *Medical Education, 32,* 384–389.

Dorr, M., Jarodzka, H., & Barth, E. (2010). Space-variant spatio-temporal filtering of video for gaze visualization and perceptual learning. In C. Morimoto & H. Instance (Eds.), *Proceedings of Eye Tracking Research & Applications ETRA '10* (pp. 307–314). New York, NY: ACM.

Egger, J., Grossmann, G., & Auchterlonie, I. A. (2003). Benign sleep myoclonus in infancy mistaken for epilepsy. *British Medical Journal, 326,* 975–976.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14,* 462–466.

Hansen, J. K., & Balslev, T. (2009). Hand activities in infantile masturbation: A video analysis of 13 cases. *European Journal of Paediatric Neurology, 13,* 508–510.

Helle, L., Nivala, M., Kronqvist, P., Gegenfurtner, A., Björk, P., & Säljö, R. (in press). Traditional microscopy instruction versus process-oriented virtual microscopy instruction: A naturalistic experiment with control group. *Diagnostic Pathology*.

Hinds, P. I. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied, 5,* 205–221.

International League Against Epilepsy. (2010). Revised terminology and concepts for organization of the epilepsies: Report of the Commission on Classification and Terminology. *Epilepsia, 51,* 676–685.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010a). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20,* 146–154.

Jarodzka, H., Van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2010b). How to convey perceptual skills by displaying experts' gaze data. In N. A. Taatgen, & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2920–2925). Austin, TX: Cognitive Science Society.

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

Kamin, C., O'Sullivan, P., Deterding, R., & Younger, M. (2003). A comparison of critical thinking in groups of third-year medical students in text, video, and virtual PBL case modalities. *Academic Medicine, 78,* 204–211.

Krupinski, E. A. (2005). Visual search of mammographic images: Influence of lesion subtlety. *Academic Radiology, 12,* 965–969.

Krupinski, E. A. (2010). Current perspectives in medical image perception. *Attention, Perception, & Psychophysics, 72,* 1205–1217.

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., Nodine, C., Krupinski, E., & Mello-Thoms, C. (2008). Using gaze-tracking data and mixture distribution analysis to support a holistic model for the detection of cancers on mammograms. *Academic Radiology, 15,* 881–886.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 311–342). Hillsdale, NJ: Erlbaum.

Litchfield, D., Ball, L. J., Donovan, T., Manning, D. J., & Crawford, T. (2010). Viewing another person's eye movements improves identification of pulmonary nodules in chest X-ray inspection. *Journal of Experimental Psychology: Applied, 16,* 251–262.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14,* 225–244.

Lüders, H., Acharya, J., Baumgartner, C., Benbadis, S., Bleasel, A., Burgess, R., et al. (1998). Semiological seizure classification. *Epilepsia, 39,* 1006–1013.

Nordli, D. (2002). Infantile seizures and epilepsy syndromes. *Epilepsia, 43,* 11–16.

Nordli, D., Bazil, C. W., Scheuer, M. L., & Pedley, T. A. (1997). Recognition and classification of seizures in infants. *Epilepsia, 38,* 553–560.

Nückles, M., Winter, A., Wittwer, J., Herbert, M., & Hübner, S. (2006). How do experts adapt their explanations to a layperson's knowledge in asynchronous communication? An experimental study. *User Modeling and User-Adapted Interaction, 16,* 87–127.

Nyström M. (2008). *Off-line foveated compression and scene perception: An eye tracking approach.* Unpublished doctoral dissertation, Lund University.

Nyström M., & Holmqvist, K. (2008). Semantic override of low-level features in image viewing – both initially and overall. *Journal of Eye Movement Research, 2,* 1–11.

Rao, R. P. N., Zielinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research, 42,* 1447–1463.

Schmidt, D., & Schachter, S.C. (2000). *Epilepsy: Problem solving in clinical practice.* London: Martin Dunitz.

Simon, H. A. (1983). Why should machines learn? In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (pp. 25–38). Palo Alto, CA: Tioga.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychological Review, 10,* 251–296.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46,* 629–646.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model‘s eye movements. *Computers in Human Behavior, 25,* 785–791.

Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2006). Effects of process-oriented worked examples on troubleshooting transfer performance. *Learning and Instruction, 16*, 154–164.

Van Lehn, K. (1989). Problem solving and cognitive skill acquisition. In M. Posner (Ed.), *Foundations of cognitive science* (pp. 527–579). Mahwah, NJ: Erlbaum.