Eye Tracking as a Tool to Study and Enhance Cognitive and Metacognitive Processes in Computer-Based Learning Environments

10

Tamara van Gog and Halszka Jarodzka

Abstract

This chapter discusses the use of eye tracking to assess cognitive and metacognitive processes and cognitive load in computer-based learning environments. Benefits of eye tracking for studying such processes are discussed (e.g., the very detailed information it provides on where a participant was looking, in what order, and for how long), but also limitations (e.g., that detailed information does not tell one which processes exactly are occurring; this has to be inferred by the researcher). In addition, this chapter provides examples of how eye tracking can be used to improve the design of instruction in computer-based learning environments, both indirectly and directly. For example, an indirect way would be to use the information on experts' or successful performers' viewing patterns to adapt instructions prior to a task (e.g., emphasizing what should be attended to later on) or to adapt the format of the task (e.g., cueing attention). A more direct way would be to display experts' or successful performers' eye movements overlaid onto the instructional materials. In the discussion, the opportunities provided by eye tracking, but also the technical challenges it poses are addressed.

Eye tracking, that is, tracking the movement of the eye ball(s) and relating these movements to a stimulus, allows researchers to determine to what part(s) of the stimulus a person allocated visual attention, for how long, and in what order (Duchowski, 2003; Holmqvist et al., 2011). Depending on the kind of eye-tracking equipment used, the stimulus can be anything ranging from naturalistic scenes (e.g., walking through a supermarket or driving in a car (see Land & Tatler, 2009)) to materials presented on a computer monitor, which is the main focus of this chapter. Determining visual attention allocation can provide researchers with information about the stimulus itself, because salient environmental

T. van Gog (✉)
Educational Psychology, Institute of Psychology, Erasmus University Rotterdam, P.O. Box 1738, Rotterdam 3000 DR, The Netherlands
e-mail: vangog@fsw.eur.nl

H. Jarodzka
Center for Learning Sciences and Technologies, Open University of The Netherlands, Heerlen, The Netherlands

features will draw attention automatically (e.g., Stelmach, Campsall, & Herdman, 1997), as well as the viewer's cognitive processes, because attention shifts also occur driven by instructions (e.g., Yarbus, 1967) or by knowledge of the task or the environment (e.g., Jarodzka, Scheiter, Gerjets, & Van Gog, 2010; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003). As such, eye tracking may be a useful tool for detailed study of attention allocation during learning in computer-based environments. In this chapter, we will provide a review of research in which eye tracking was used to study, as well as enhance, (meta)cognitive processes in computer-based learning environments.

A Brief History of Eye Tracking

First used in the nineteenth century, eye-tracking technology has undergone dramatic changes in the last decades, making it more widely available and more easy to use. We will provide a brief overview here based on elaborate reviews of the history of eye tracking, for which the reader is referred to Richardson and Spivey (2004) and Wade and Tatler (2005). The very first studies on eye movements consisted of direct observations of the eyes during reading (e.g., using mirrors). This allowed Javal to distinguish two different types of eye movements: short rapid movements and stops (so-called saccades and fixations). However, this procedure did not allow for objective measurements of the eye movements. At the end of the nineteenth century, Delabarre and Huey addressed this issue by developing rather crude and highly intrusive eye-tracking devices using ceramic lenses with a small hole, to which a wire was attached that "drew" the movement of the eye. A major breakthrough in eye-tracking technology came early in the twentieth century when Dodge started using photography to capture the movements of the eyes, which was far less intrusive and not painful for the participants (people still had to be restrained from moving their heads though). Later *video-based eye trackers* allowed for more freedom of movement and for very precise

analysis of the allocation of the eye movements on the stimulus. Most widely used in applied eye-tracking research nowadays is the pupil and corneal reflection method, in which an infrared light source is directed towards the eye, causing a reflection on the cornea captured by an infrared-sensitive video camera. This corneal reflection is the brightest spot on the image, while the pupil is the darkest one. When the eye moves, the pupil does too, but the corneal reflection hardly does. So, by calculating the distance between the pupil and the corneal reflection, the direction of the eye can be calculated, and in combination with parameters of the environment, it can be inferred at which part of the stimulus the eye was directed at different points in time. A wide variety of measures can be obtained by means of eye tracking (see Duchowski, 2003; Holmqvist et al., 2011); we will shortly discuss only a few main measures here that appear in the research discussed in this chapter.

Measures Obtained via Eye Tracking

Two important eye movement measures were already mentioned in the previous section: fixations and saccades. During *fixations*, the eye is (almost completely) still and information can be extracted from a stimulus. As a consequence, the location and duration of fixations provide an indication of what information is attended to and how intensively that information is being processed (relative to other information; cf. eye-mind assumption by Just & Carpenter, 1980). During *saccades*, that is, the rapid eye movements in between fixations, the focus of visual attention is moved to another location, and we are not able to take in visual information—although it seems that under specific circumstances, some information, like motion, can be very roughly processed (Castet & Masson, 2000). Both fixations and saccades occur for all kinds of stimuli. A type of eye movement that occurs only when inspecting dynamic stimuli such as videos or animations is *smooth pursuit*, which occurs when the eye

follows moving objects (Dodge, 1903). While fixations and saccades can be easily detected by contemporary eye-tracking software, there are no adequate algorithms yet to detect smooth pursuit, which therefore requires complex calculations on raw gaze data (Holmqvist et al., 2011).

Other measures that can be obtained through eye tracking and that may be relevant in research on computer-based learning are blinks and pupil dilation. *Blinks* of the eye are quite easy to identify, and the frequency of occurrence depends, for instance, on tiredness (e.g., Barbato et al., 2007) or—as will be discussed later—mind wandering (e.g., Smilek, Carriere, & Cheyne, 2010). The dilation of the pupil can, for example, provide information about cognitive load as we will discuss below (e.g., Hyönä, Tommola, & Alaja, 1995; Kahneman & Beatty, 1966; Klingner, Tversky, & Hanrahan, 2010; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004). It is a difficult measure to use though, as pupil dilation is very sensitive to influences of other factors which need to be carefully controlled (e.g., light changes and changes in the brightness of the stimulus). For more information on these different measures, the reader is referred to Duchowski (2003) and Holmqvist and colleagues (2011).

Studying Cognitive and Metacognitive Processes in Computer-Based Learning Environments

Written text is still a core component of many computer-based learning environments. As mentioned above, early eye-tracking studies focused on reading, and it probably still is one of the most widely studied processes in eye-tracking research. A comprehensive review of eye-tracking research in reading is beyond the scope of this chapter. The reader is referred to Rayner (1998, 2009) for elaborate reviews. Here, we will first discuss some applications of eye tracking for studying cognitive processes in multimedia and hypermedia learning environments.

Then, we will address the use of eye tracking to assess cognitive load. Last but not least, we will discuss what eye-tracking research can reveal about metacognitive processes in computer-based learning environments.

Cognitive Processes: Multimedia and Hypermedia Learning

Presentation of Hypertext

Written text in a computer-based learning environment is usually hypertext, that is, it contains hyperlinks to other information which the reader can immediately access (Conklin, 1987). As a consequence, hypertexts have a nonlinear structure which not only allows but also requires the user to determine their own sequence of reading information and therefore carries a risk of disorientation and overload. However, even though hypertexts are nonlinear, they may be preceded by *concept maps* to guide navigation. Amadieu, Van Gog, Paas, Tricot, and Mariné (2009) investigated the effects of a network concept map structure that provides relational links to a hierarchical structure that provides organizational links and cues (and can be considered somewhat more "linear" than network structures). The latter was hypothesized to guide learners' attention towards the main concepts and their semantic relationships. In the network structure, participants with higher prior knowledge spent more time fixating certain key nodes than participants with lower prior knowledge, whereas no such difference occurred in the hierarchical structure. This suggests that a hierarchical structure, in which attention is guided to main concepts, is especially helpful for low-prior-knowledge learners, whereas learners with more prior knowledge can apply that knowledge in searching for relevant concepts in a network structure.

Next to written or spoken textual information, most computer-based learning environments contain visualizations associated with those texts, such as pictures, drawings, diagrams, animations, and videos. The use of text combined

with visualizations, however, places certain attentional demands on learners that may or may not be helpful for learning depending on the design. Therefore, the use of eye tracking may have added value in discovering the underlying mechanisms of effects on learning (Van Gog, Kester, Nievelstein, Giesbers, & Paas, 2009) as will be shown in the examples that follow.

Effects of Split Attention or Spatial Contiguity

Research has shown that when providing different mutually referring information sources, such as written text and a graphic, a separate presentation format hampers learning compared to an integrated presentation format. This is known as the split-attention effect or spatial contiguity effect (for a review, see Ayres & Sweller, 2005). However, what exactly causes this effect is unclear. For instance, do learners, when presented with a separate format, fail to integrate both information sources and study them separately one after the other? Or do they try to process them simultaneously and switch between both sources, but lose their last position in the graphic or text as a consequence, leading to unnecessary search, rereading, or both?

Because eye movement data reflect attention and shifts in attention, eye tracking may be very helpful in investigating the underlying mechanisms of the split-attention effect. Hegarty and Just (1993) conducted an eye-tracking study on comprehension of text and diagrams in separated format. They found that readers often switched attention from the text to the diagrams, mostly at the end of sentences or clauses, suggesting that integrations of both representations were made at the level of individual components or groups of components. Using illustrated science textbook passages, Hannus and Hÿönä (1999) found that learners spent by far the most time on the text: Only 6% did they spend on illustrations, and this did not differ between high- and low-ability learners. However, although switching attention between text and pictures was also relatively low in general, high-ability learners did switch more

often than low-ability learners. Studying effects of animations with written text, Schmidt-Weigand, Kohert, and Glowalla (2010) also found that learners spent more time reading the text than inspecting the animation and consistently started reading before alternating between text and animation.

Jarodzka, Janssen, Kirschner, and Erkens (submitted) studied this effect in computer-based *testing*. For an authentic arts exam, students completed an electronic version with half of the questions presented in the original separated format and the other half in an integrated format (i.e., within-subject design). Eye tracking was used to estimate the amount of visual search required. Results showed that, in the integrated format, students attended more (indicated by total fixation durations) to additional information provided next to the question text, like pictures and historical background information, and processed this additional information more intensively (indicated by more fixations) than they did when the information was presented in a separated format. By changing the design of such testing environments, students' attention was guided so that they intensively processed *all* given information. Interestingly, however, the integrated format did in this case not lead to higher but to lower test scores. These results suggest that (part of) the additional information given in the tests was redundant, which was useful information for the organization that developed these tests to further improve them.

Under experimental conditions, learners are often "forced" to study material for a certain amount of time. In computer-based learning environments, however, there is usually a large amount of information available (often more than can be studied during the experimental session), and students can decide for themselves which information to consult and for how long. Research on authentic reading behavior suggests that under such circumstances, separate presentation of text and pictures may have even more deleterious effects in that the text may be skipped altogether: In a naturalistic newspaper-reading study, Holsanova, Holmberg, and Holmqvist (2009) found that when text and graphic were presented separately, readers typically read the headline and then switched to the

graphic while mostly ignoring the text, whereas when the graphic was integrated with the text, both were processed together.

In sum, it seems that under experimental (learning) conditions people seem to focus on the main text in a separated format (Hannus & Hÿonä, 1999; Jarodzka et al., submitted; Schmidt-Weigand et al., 2010), while under naturalistic (leisure) conditions they mostly focus on pictures (Holsanova et al., 2009). When information is presented in an integrated format, however, all information seems to be processed (Holsanova et al., 2009; Jarodzka et al., submitted).

Effects of Cueing or Signaling

Another well-known effect established by research on multimedia learning is the cueing or signaling effect (for reviews, see De Koning, Tabbers, Rikers, & Paas, 2009; Mayer, 2005) in which the visual saliency of parts of the stimulus material is manipulated to draw the learner's attention. Ozcelik and colleagues used eye tracking to investigate the effect of cueing by means of temporarily changing the color of labels in an otherwise *static* illustration (Ozcelik, Arslan-Arib, & Cagiltay, 2010) or cueing corresponding information in the text and illustration by giving it the same color (Ozcelik, Karakus, Kursun, & Cagiltay, 2009; see also Folker, Sichelschmidt, & Ritter, 2005). They established that such cues indeed successfully guided visual attention and led to more efficient information processing and better learning outcomes.

Increasingly, visual materials provided in computer-based learning environments are *dynamic*, like videos or animations. Cueing that is effective for static presentation formats is not necessarily effective for dynamic formats, and cueing may be even more necessary in dynamic visualizations because (part of) the information may be transient and hence no longer available for processing if it is not attended to at the right moment.

Using dynamic visualizations, De Koning, Tabbers, Rikers, and Paas (2010) showed that spotlight cues in which the important information is made more salient by reducing the saliency of surrounding information (e.g., through darkening) were effective for guiding attention to the cued parts. Boucheix and Lowe (2010) established that continuous cues in which a colored "ribbon" was spreading were more effective than arrow cues for attention guidance in dynamic visualizations. They also showed the importance of temporal aspects of cueing (i.e., guiding attention to the right place at the right time) for attention guidance and learning.

In sum, by using eye tracking, it can be established whether cues in multimedia learning materials indeed are successful at guiding learners' attention.

Effects of Pedagogical Agents

Animated pedagogical agents are often used in multimedia materials in computer-based learning environments (for a review, see Moreno, 2005). Louwerse, Graesser, McNamara, and Lu (2009) applied eye tracking to investigate how learners interact with embodied conversational agents (ECAs), that is, animated humanoid characters that communicate with the learner. They found that learners interact with those agents much as they would with a real human conversational partner, fixating mostly on the agent, or, when multiple agents were present, fixating on the agent that was speaking. This could perhaps explain why the presence of such agents does not always foster learning; when the learner is attending to the agent, she/he may not be attending to the learning content on the screen that the agent is referring to.

Cognitive Load

Eye-tracking data can provide information not only about the processes evoked by different types of materials but also about the *demands* on working memory imposed by those processes (i.e., cognitive load; e.g., Hÿonä, Tommola, & Alaja, 1995; Kahneman & Beatty, 1966; Klingner et al., 2010; Van Gerven et al., 2004). For example, Kahneman and Beatty (1966) showed that pupil dilation is associated with working memory load. Participants had to memorize a string of digits or a list of words and report those back (immediate recall) or had to transform a string of

digits (add one to each digit). Their data on the digit strings showed that with the presentation of each additional digit, pupil dilation increased, while with reporting back each digit, it decreased. Moreover, pupil dilation increased more steeply with the more demanding tasks of learning word lists or transforming digits than with learning digit strings. Hyönä and colleagues (1995) used pupil dilation to investigate variations in cognitive load during translation tasks. They showed that variations in cognitive load during a translation task were reflected in pupil size: More difficult words to translate resulted in higher levels of pupil dilation than words that were easy to translate. Klingner and colleagues (2010) investigated the effect of auditory versus visual task presentation on pupil dilation with three different tasks and found that while patterns of dilation were similar for auditory and visual presentation for all three tasks, the magnitudes of pupil response were greater for auditory presentation than for visual presentation, suggesting the latter is less cognitively demanding. Van Gerven and colleagues (2004) investigated the usefulness of the pupil response as an indicator of cognitive load in young and aging adults. They used a memory-search task, consisting of two phases. In the encoding phase, participants had to memorize strings of one to six digits (none occurred more than once). In the search phase, participants had to judge whether single-digit probes belonged to the memory set. For both young adults and elderly participants, pupil dilation systematically increased with the length of the string of digits in the encoding phase (i.e., with task difficulty), but in the search phase, pupil dilation was only sensitive to task load variations for the young adults, which suggests this measure may not always be suitable in studies with elderly participants.

Metacognitive Processes

Monitoring Learning and Comprehension

Metacognitive judgments play an important role in self-regulated learning, because such judgments, for example, of whether information has been sufficiently learned or not, affect the allocation of study time and choices about items to select for further studying (Metcalfe, 2009).

Kinnunen and Vauras (1995) assessed children's monitoring of *comprehension* during reading by means of eye tracking. The need for comprehension monitoring was enhanced by causing difficulties in text processing in certain sentences, for example, by adding a nonsense word or a word that made the sentence inconsistent with general knowledge or with a prior sentence. They assumed that comprehension monitoring would be associated with higher reading time and a higher number of regressions (i.e., looking back) to difficult passages in the text. Comprehension was assessed by a text summary provided by the students after reading. Results indeed showed that reading complex sentences lead to higher reading times and more regressions compared to regular sentences. Moreover, this effect was stronger for high-achieving students. Graesser, Lu, Olde, Cooper-Pye, and Whitten (2005) also created a cognitive disequilibrium in participants who read illustrated texts about devices by presenting a breakdown scenario that was assumed to result in question asking, and investigated the relationship between question asking and eye movements. They showed that deep comprehenders tended to formulate better questions and fixate on fault-related components just before or during question formulation. In sum, eye-tracking data can provide detailed insight into the metacognitive process of comprehension monitoring when studying texts.

Roderer and Roebers (2010) conducted an eye-tracking study of confidence judgments. Children were shown easy and difficult Kanji symbols of which they had previously learned the meaning or new ones that they could not recognize. The children were asked to select the correct meaning from four alternatives. Subsequently, a *confidence* rating followed and they were asked to indicate how confident they were of their answer by pointing at one of five smileys (ranging from a very sad looking one to a very happy looking one). In addition to this explicit confidence judgment provided by pointing, the authors measured implicit judgments based on the eye movement data from the phase

before the explicit judgment was provided (i.e., looking at the confidence judgment "category" that attracted a maximum of fixation time during confidence scale presentation). They found a high correlation between explicit and implicit confidence judgments, suggesting that eye tracking can be used as a measure of confidence judgments.

Monitoring Information About Other Students' Knowledge in Collaborative Learning

Sangin, Molinari, Nüssli, and Dillenbourg (2008) used eye tracking to investigate how students' monitored and used information about *other students'* knowledge in collaborative learning in a computer-based environment. Participants created concept maps in dyads. One group of participants had an awareness tool available that provided information on the other person's knowledge. Results showed that looking at this knowledge awareness tool (KAT) was positively related to learning. When combined with verbal data from the episodes in which participants looked at the KAT, it was found they looked at the KAT for three reasons: when they were seeking for specific knowledge, when their peers provided information, or when their peers provided cues regarding their existing or nonexisting knowledge.

Self-Explaining

Conati, Merten, Muldner, and Ternes (2005) used eye-tracking data to estimate metacognitive behavior (more specifically, self-explanation), while students performed a task in a computer-based mathematics learning environment. They also asked participants to think aloud. Afterwards, the verbal data were coded in terms of whether or not they contained self-explanations. Then, time on task data (obtained from log files) and eye-tracking data (gaze shifts) were related to each of these episodes that did and did not contain self-explanations. The assumption was that self-

explanations would take more time and would be accompanied by gaze shifts between graphs and formulas. Results show that time on task had the highest sensitivity, while eye-tracking data had the highest specificity for predicting self-explanations. In this study, an algorithm was used for analyzing eye-tracking data, which has the benefit over verbal data that it can be analyzed and used *online* (i.e., during learning). Provided eye-tracking data can be coupled to cognitive or metacognitive processes with great sensitivity and specificity, such algorithms could be used to adapt a computer-based learning environment in real time to the learner's cognitive or metacognitive state (e.g., by providing self-explanation prompts when learners do not spontaneously self-explain).

Registering Off-Task Behavior

Mind wandering, that is, a focus of attention on internal processes rather than on processing the external environment, seems to be associated with an increase in eye blinks (Smilek et al., 2010). Smilek and colleagues (2010) had participants read a text during which they were randomly probed ten times by an auditory stimulus to report whether they were on task (i.e., reading) or mind wandering, which could be task related (e.g., thoughts relevant to the text) or unrelated (e.g., thoughts about room temperature or meals). In the 5 s before the probes, participants blinked more when they were mind wandering than when they were on task, and participants made less fixations on the text (even when corrected for blink time). Using a comparable self-report and prompting procedure, Reichle, Reineberg, and Schooler (2010) investigated mindless reading, in which the eyes keep moving across the page but the individual is mind wandering. They found that, compared to normal reading, fixations were longer during mindless reading and were also less affected by characteristics of the text, presumably due to the absence of cognitive processes that normally direct eye movements during reading.

These findings suggest that eye-tracking data may provide interesting information on whether

or not participants are on task in computer-based learning environments. A problem of course is that mind wandering may concern task-related thoughts, which are probably highly relevant for learning (e.g., for making inferences beyond the literal text) and that there is (as yet) no way to distinguish such task-relevant episodes of mind wandering from task-unrelated episodes solely based on the eye movement data.

Limitations of Eye Tracking in Studying Cognitive and Metacognitive Processes: Adding Verbal Reports

The studies discussed above show that eye fixation data can provide interesting information about participants' (visual) attention allocation: They tell us where a participant was looking, in what order, and for how long, and how much they were blinking. However, these data require a substantial amount of inferences about underlying cognitive processes, as they do not explain *why* a participant was looking somewhere for a certain amount of time or in a certain order. To reduce the amount of inferences required by the researcher, eye movement data can be complemented with concurrent verbal reports (i.e., thinking aloud; Ericsson & Simon, 1993; for a combination with eye tracking, see, e.g., Van Gog, Paas, & Van Merriënboer, 2005a). The central assumption behind the use of thinking aloud data is "that it is possible to instruct subjects to verbalize their thoughts in a manner that does not alter the sequence and content of thoughts mediating the completion of a task and therefore should reflect immediately available information during thinking" (Ericsson, 2006, p. 227).

However, even if verbalizing thoughts does not alter those thoughts, a potential drawback of asking participants to think aloud during task performance in combination with eye tracking is that this has been suggested to affect their eye movements. For instance, in complex tasks the speech planning process has been shown to alter the allocations of eye movements (e.g., Holsanova, 2008), and, on average, oral reading increases

fixation duration and reduces saccade length compared to silent reading for skilled English readers (Rayner, 2009), and concurrent reporting is suspected to slow down task performance (Karpf, 1973) and might therefore lead to more eye movements.

As an alternative to concurrent reports, retrospective verbal reports could be used. However, compared to concurrent reports, retrospective reports tend to suffer from omission of information due to forgetting and from fabulations (e.g., Kuusela & Paul, 2000). Cueing a retrospective report with information from the task performance process might prevent forgetting and fabulation (Van Someren, Barnard, & Sandberg, 1994). Most eye-tracking software allows not only for recording but also for replaying the records of eye movements as an overlay on the stimulus or computer screen recording, and such replays of eye movement records may provide an excellent cue for retrospective reports (Van Gog, Paas, Van Merriënboer, & Witte, 2005b; see also Hansen, 1991; Russo, Johnson, & Stephens, 1989). Van Gog and colleagues found that both concurrent and cued retrospective reporting resulted in quantitatively more information than retrospective reporting without a cue. Interestingly, cued retrospective reporting also resulted in a higher number of metacognitive statements in the protocols than concurrent and retrospective reporting.

Cued retrospective reporting might provide a valuable alternative to concurrent reporting, not just because it cannot affect eye movements as concurrent reporting has been suggested to do but especially for research with novice participants or with instructional materials that make concurrent reporting impossible. For novices, because they have little prior knowledge, tasks often impose a high cognitive load, and as a result, they may stop verbalizing their thoughts during concurrent reporting (Ericsson & Simon, 1993). Indeed, in the study by Van Gog and colleagues (2005b), participants who had lower performance and experienced higher cognitive load on the tasks (i.e., who had lower expertise) also indicated that they preferred cued retrospective reporting over concurrent reporting (reported in

Van Gog, 2006). Not only learners' expertise level but also the type of learning material provided can have consequences for which verbal reporting technique to choose. For instance, instructional materials that are widely used in computer-based learning environments with which concurrent reporting is not possible are animations or videos that contain spoken text.

Cued retrospective reporting has been used, for instance, in problem-solving or information-search tasks in which mouse and keyboard operations were also recorded (Brand-Gruwel, Van Meeuwen, & Van Gog, 2008; Schwonke, Berthold, & Renkl, 2009; Van Gog et al., 2005b), and the replays could therefore cue memory of both overt actions (i.e., via mouse clicks that occurred on the screen) and covert processes (i.e., via the display of eye movements) that occurred during task performance. However, it has also been used with animations or videos in which no overt actions such as mouse clicks were required and the eye movements constituted the sole cue (De Koning et al., 2010; Jarodzka, Scheiter et al., 2010).

Enhancing Cognitive and Metacognitive Processes in Computer-Based Learning Environments

Eye tracking can also be applied to improve the design of components of computer-based learning environments. For example, Buscher, Cutrell, and Morris (2009) recorded the eye movements of participants surfing on several hundreds of Web pages. Based on these data, they developed a model that successfully predicts the saliency of single Web page elements, which can inform designers of (instructional) Web pages. Kammerer and Gerjets (2010) found that the design of a Web search engine influenced the thoroughness of information search. The authors recorded participants' eye movements while they searched information using either a traditional list search engine or a novel search engine, in which search results were presented in a tabular format. Participants searching the tabular format were found to look

at more search results, that is, they evaluated the information resulting from the search more thoroughly.

In addition, eye tracking may be used to reveal what the differences are in successful and unsuccessful problem solvers' attention allocation, and this information may then be used to develop cues or instructions to support learners in computer-based environments. For example, Grant and Spivey (2003) showed that participants who were successful at solving Duncker's radiation problem (an insight problem) attended relatively more to a certain area in the picture than unsuccessful problem solvers. In a second experiment, they showed that incorporating a perceptual cue to draw attention to this area led to an increase in successful problem solving. A similar approach was taken by Schwonke and colleagues (2009), using worked examples on probability calculation that consisted of multiple representations (text, tree diagram, and mathematical equation). They showed that conceptual understanding after example study was positively associated with more extensive processing of the tree diagrams and negatively with transitions from text to equations (skipping the diagrams). This suggested that the diagrams played an important role in learning from the worked examples. In a second study, Schwonke and collaborators (2009) provided half of the participants with instruction on how the representations were functionally related, which had a strong effect on learning that was partially mediated by allocation of visual attention to the diagrams.

Next to this indirect route of informing the design of components of computer-based learning environments, eye tracking may also be applied in more direct ways, for instance, in the design of examples. Modelling examples in computer-based learning environments often consist of screen captures of a human model performing a task, and depending on the type of task, the model may also provide a verbal explanation of why she/he is doing what she/he is doing (e.g., McLaren, Lim, & Koedinger, 2008; for a review of research on modelling examples, see Van Gog & Rummel, 2010). Often, the model is an expert on the particular task she/he is demonstrating.

In this case, a problem might arise, especially in examples in which information is transient: Eye-tracking research has shown that with increasing knowledge or expertise on a task, individuals fixate faster and relatively more on relevant information (e.g., Charness, Reingold, Pomplun, & Stampe, 2001; Haider & Frensch, 1999; Jarodzka, Scheiter et al., 2010; Van Gog et al., 2005a). In other words, there might be a discrepancy in attention allocation between the learner and the model, and if the learner does not attend to the right information at the right time, understanding might be compromised, for example, because the information is no longer available for further processing (in case of transience) or because the explanation by the model is more difficult to follow when the learner is not attending to the same information as the model.

Therefore, Van Gog, Jarodzka, Scheiter, Gerjets, and Paas (2009) investigated whether incorporating a display of the eye movements made by the expert model in screen-capture modelling examples with or without spoken explanations could guide students' attention and enhance their learning of a procedural problem-solving task. In contrast to their expectation, they did not find a positive effect of displaying eye movements, although results suggested there might be benefits on transfer. They even found a negative effect when the modelling examples contained both eye movements and spoken explanations, presumably because the verbal explanations were sufficient to guide learners' attention in this task. Using examples of a more perceptual task (learning to classify fish locomotion patterns) with a spoken verbal explanation, in which the verbal explanation was less likely to be sufficient to guide learners' attention, Jarodzka, Van Gog, Dorr, Scheiter, and Gerjets (2013) did find positive effects of displaying the expert model's eye movements in modelling examples on learning.

Not looking at *learning*, but at a direct influence on *performance*, Litchfield, Ball, Donovan, Manning, and Crawford (2010) investigated the effects of seeing another person's eye movements on a visual diagnosis task in medicine: identifying pulmonary nodules (i.e., a lesion in the lung smaller than 3 cm in diameter) in chest X-rays. The "models" in their study did not behave didactically (i.e., their viewing behavior was natural) and did not provide any additional verbal explanation. They found that novices performed better after seeing the "models" searching for nodules.

Such eye movement modelling examples can be constructed and implemented in computer-based learning environments relatively easily, because eye-tracking software nowadays usually allows exporting a screen capture with a display of eye movements as a digital video file. If eye trackers would become cheaper and would become available in classrooms, other direct uses of eye tracking could be conceived of, for instance, in collaborative learning or problem solving. For example, Velichkovsky (1995) conducted a study on real-time cooperative puzzle problem solving by expert-novice pairs, in which the novice controlled the mouse and could observe the expert's eye movements, so the expert could indicate with his gaze what the novice should do.

Another possible application when eye trackers would be more ubiquitous would be to use eye movement records to stimulate reflection. As mentioned above, the findings by Van Gog and colleagues (2005b) showed that reviewing a record of one's own actions and eye movements (during cued retrospective reporting) resulted in a higher number of metacognitive comments (e.g., statements about the adequacy of the learner's own knowledge, actions, or strategies) than concurrent and retrospective reporting. This occurred rather spontaneously, because the instructions for reporting in each condition (concurrent, retrospective, or cued retrospective) were neutral. It also did not occur frequently; even though the difference was significant, the actual number of metacognitive statements in cued retrospective reporting was not very high. However, these findings do suggest that reviewing a record of one's own actions and eye movements may trigger reflective processes, and therefore it has been suggested that such records might be used as explicit tools for reflection (Van Gog, Jarodzka et al., 2009, Van Gog, Kester et al., 2009) or could be implemented to aid self-assessment (Kostons,

Van Gog, & Paas, 2009). Especially combined with additional metacognitive prompts or scaffolds, this might be an effective tool for fostering reflection.

Finally, as mentioned previously, *real-time* analysis of eye movement data could be applied in intelligent tutoring systems or other adaptive learning environments to monitor students engagement in metacognitive behaviors such as self-explaining and to use that information to dynamically adapt the content offered to students (Conati et al., 2005; see also Merten & Conati, 2006; for a discussion of other potential uses of real-time eye movement analysis in tutoring systems, such as error prediction, detection of undesirable solution processes, and identifying when messages are ignored, see Gluck, Anderson, & Douglass, 2000).

Discussion

In sum, eye tracking is not only a useful tool to study (meta)cognitive processes and cognitive load in computer-based learning environments but can also be used indirectly or directly in the design of components of such environments to enhance (meta)cognitive processes and foster learning. Even though eye movement data are still challenging to collect and analyze and often need to be triangulated with another data source such as verbal data to make inferences about associated cognitive processes, they do provide a unique opportunity to study certain kinds of processes in a level of detail that no other data source provides. For example, screen recordings without eye movement data would only provide information on how long the page in its entirety was attended to, not which specific parts of the page received attention. Or in hypermedia environments, screen recordings would only show what hyperlinks are being clicked on, but not which other links have been previously considered but were not opened.

The use of eye tracking to study cognitive processes in computer-based learning environments is increasing rapidly, but there has been much less eye-tracking research on metacogni-

tive processes. The studies discussed in this chapter do highlight some promising areas in which eye tracking may provide useful information on metacognitive processes, such as monitoring one's own comprehension, monitoring information about other people's knowledge in collaborative learning environments, and predicting when students are or are not making self-explanations (thereby providing options for, for instance, adaptive prompting).

The fact that eye-tracking technology is still advancing rapidly will probably stimulate further research on (meta)cognitive processes in computer-based learning environments. In the last decade or so, eye-tracking equipment has become more affordable and much easier to operate. With further technological advances, analysis of eye movement data may become less cumbersome. For example, a major problem when analyzing data on areas of interest (AOI) in videos is that these AOIs often move about, requiring segmentation of the video into very small segments and then computing AOI data and aggregating them over the whole video (see, e.g., Jarodzka, Van Gog et al., 2013), but software solutions are being developed to enable dynamic AOIs (see, e.g., Papenmeier & Huff, 2010). Software features for displaying eye movement data have already come a long way, such as the option to make integrated digital videos of screen recordings and eye movements, and further developments may open up new avenues for the design of learning tasks in computer-based environments.

References

Amadieu, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*, 376–386.

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 135–146). New York: Cambridge University Press.

Barbato, G., De Padova, V., Paolillo, A. R., Arpaia, L., Russo, E., & Ficca, G. (2007). Increased spontaneous eye blink rate following prolonged wakefulness. *Physiology & Behavior, 90*, 151–154.

Boucheix, J.-M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction, 20*, 123–135.

Brand-Gruwel, S., Van Meeuwen, L., & Van Gog, T. (2008). The use of evaluation criteria when searching the WWW: An eye-tracking study. In A. Maes & S. Ainsworth (Eds.), *Proceedings EARLI Special Interest Group Text and Graphics 'Exploiting the opportunities: Learning with textual, graphical, and multimodal representations'* (pp. 34–37). Tilburg, The Netherlands: Tilburg University.

Buscher, G., Cutrell, E., & Morris, M. R. (2009). What do you see when you're surfing? Using eye tracking to predict salient regions of the web pages. In *Proceedings of the 27th CHI conference* (pp. 21–30). New York: ACM.

Castet, E., & Masson, G. S. (2000). Motion perception during saccadic eye movements. *Nature Neuroscience, 3*, 177–183.

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

Conati, C., Merten, C., Muldner, K., & Ternes, D. (2005). Exploring eye tracking to increase bandwidth in user modeling. In L. Ardissono, P. Brna, & A. Mitrovic (Eds.), *User modeling 2005: 10th international conference* (pp. 357–367). Berlin: Springer.

Conklin, J. (1987). Hypertext: An introduction and survey. *IEEE Computer, 20*(9), 17–41.

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 Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*, 111–122.

Dodge, R. (1903). Five types of eye movements in the horizontal meridian plane of the field of regard. *The American Journal of Psychology, 8*, 307–329.

Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice*. London: Springer.

Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 223–241). Cambridge: Cambridge University Press.

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

Folker, S., Sichelschmidt, L., & Ritter, H. (2005). Processing and integrating multimodal material: The influence of color-coding. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 690–695). Mahwah, NJ: Erlbaum.

Gluck, K. A., Anderson, J. R., & Douglass, S. (2000). Broader bandwidth in student modeling: What if ITS were "Eye"TS? *Intelligent Tutoring Systems—Lecture Notes in Computer Science, 1839*, 504–513.

Graesser, A. C., Lu, S., Olde, B. A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition, 33*, 1235–1247.

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

Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 172–190.

Hannus, M., & Hÿönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology, 24*, 95–123.

Hansen, J. P. (1991). The use of eye mark recordings to support verbal retrospection in software testing. *Acta Psychologica, 76*, 31–49.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*, 717–742.

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

Holsanova, J. (2008). *Discourse, vision, and cognition*. Philadelphia: John Benjamins.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2009). Reading information graphics: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology, 23*, 1215–1226.

Hÿönä, J., Tommola, J., & Alaja, A.-M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 48*, 598–612.

Jarodzka, H., Janssen, N., Kirschner, P. A., & Erkens, G. (submitted). Avoiding split attention in computer-based testing: Is neglecting additional information facilitative? Manuscript submitted for publication.

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.

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K. & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction, 25*, 62–70.

Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–355.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583–1585.

Kammerer, Y., & Gerjets, P. (2010). How the interface design influences users' spontaneous trustworthiness evaluations of web search results: Comparing a list and a grid interface. In C. Morimoto & H. Instance (Eds.), *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications ETRA'10* (pp. 299–306). New York: ACM.

Karpf, D. (1973). *Thinking aloud in human discrimination learning*. Unpublished doctoral dissertation, State University of New York, New York, USA.

Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high- and low-achieving primary school children's reading. *Learning and Instruction, 5*, 143–165.

Klingner, J., Tversky, B., & Hanrahan, P. (2010). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology, 48*, 323–332.

Kostons, D., Van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology, 23*, 1256–1265.

Kuusela, H., & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *The American Journal of Psychology, 113*, 387–404.

Land, M., & Tatler, B. (2009). *Looking and acting: Vision and eye movements in natural behaviour*. Oxford, UK: Oxford University Press.

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.

Louwerse, M. M., Graesser, A. C., McNamara, D. S., & Lu, S. (2009). Embodied conversational agents as conversational partners. *Applied Cognitive Psychology, 23*, 1244–1255.

Mayer, R. E. (2005). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 183–200). New York: Cambridge University Press.

McLaren, B. M., Lim, S., & Koedinger, K. R. (2008). When and how often should worked examples be given to students? New results and a summary of the current state of research. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2176–2181). Austin: Cognitive Science Society.

Merten, C., & Conati, C. (2006). Eye-tracking to model and adapt to user meta-cognition in intelligent learning environments. *Proceedings of the 11th international conference on Intelligent user interfaces* (IUI '06), (pp. 39–46). New York: ACM.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*, 159–163.

Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 507–523). New York: Cambridge University Press.

Ozcelik, E., Arslan-Arib, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*, 110–117.

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers in Education, 53*, 445–453.

Papenmeier, F., & Huff, M. (2010). DynAOI: A tool for matching eye-movement data with dynamic areas of interest in animations and movies. *Behavior Research Methods, 42*, 179–187.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21*(9), 1300–1310.

Richardson, D. C., & Spivey, M. (2004). Eye tracking: Characteristics and methods. In G. Wnek & G. Bowlin (Eds.), *Encyclopedia of biomaterials and biomedical engineering* (pp. 568–572). New York: Marcel Dekker, Inc.

Roderer, T., & Roebers, C. M. (2010). Explicit and implicit confidence judgments and developmental differences in metamemory: An eye-tracking approach. *Metacognition and Learning, 5*, 229–250.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory & Cognition, 17*, 759–769.

Sangin, M., Molinari, G., Nüssli, M. -A., & Dillenbourg, P. (2008). How learners use awareness cues about their peer's knowledge? Insights from synchronized eye-tracking data. *Proceedings of the 8th International Conference on International Conference for the Learning Sciences* (Vol. 2, pp. 287–294). International Society of the Learning Sciences.

Schmidt-Weigand, F., Kohert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction, 20*, 100–110.

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology, 23*, 1227–1243.

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as an indicator and embodiment of mind wandering. *Psychological Science, 21*, 786–789.

Stelmach, L. B., Campsall, J. M., & Herdman, C. M. (1997). Attentional and ocular movements. *Journal of Experimental Psychology. Human Perception and Performance, 23*, 823–844.

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 Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology, 41*, 167–174.

Van Gog, T. (2006). *Uncovering the problem-solving process to design effective worked examples*. Doctoral Dissertation, Open University of The Netherlands, Heerlen, The Netherlands.

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., Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior, 25*, 325–331.

Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2005a). Uncovering expertise-related differences in trouble-shooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology, 19*, 205–221.

Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005b). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology. Applied, 11*, 237–244.

Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*, 155–174.

Van Someren, M., Barnard, Y., & Sandberg, J. (1994). *The think aloud method: A practical guide to modeling cognitive processes*. Amsterdam: Academic Press.

Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition, 3*, 199–224.

Wade, N. J., & Tatler, B. (2005). *The moving tablet of the eye: The origins of modern eye movement research*. Oxford: Oxford University Press.

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