Differentiating between Encoding and Processing during Diagnostic Reasoning: An Eye tracking study.

Anja Klichowicz (anja.klichowicz@psychologie.tu-chemnitz.de)
Technische Universität Chemnitz, Department of Psychology, Cognitive and Engineering Psychology,
Wilhelm-Raabe-Strasse 43, D-09120 Chemnitz, Germany

Agnes Scholz (agnes.scholz@psychologie.tu-chemnitz.de)
Technische Universität Chemnitz, Department of Psychology, Cognitive and Engineering Psychology,
Wilhelm-Raabe-Strasse 43, D-09120 Chemnitz, Germany

Sascha Strehlau (sascha.strehlau@yahoo.de)
Technische Universität Chemnitz, Department of Psychology, Cognitive and Engineering Psychology,
Wilhelm-Raabe-Strasse 43, D-09120 Chemnitz, Germany

Josef F. Krems (josef.krems@psychologie.tu-chemnitz.de)
Technische Universität Chemnitz, Department of Psychology, Cognitive and Engineering Psychology,
Wilhelm-Raabe-Strasse 43, D-09120 Chemnitz, Germany

Abstract

When finding a best explanation for observed symptoms a multitude of information has to be integrated and matched against explanations stored in memory. Although assumptions about ongoing memory processes can be derived from the process models, little process data exists that would allow to sufficiently test these assumptions.

In order to explore memory processes in diagnostic reasoning, 29 participants were asked to solve a visual reasoning task (the Black Box paradigm) where critical information had to be retrieved from memory.

This study focused on differentiating between processes that take place during the encoding and the evaluation of symptom information by comparing eye movement measures (the number of fixation and fixation duration per dwell).

Results will be discussed in light of existing theories on sequential diagnostic reasoning. Further, it will be discussed to which extent eye movements can be informative about memory processes underlying sequential diagnostic reasoning.

Keywords: diagnostic reasoning; eye tracking; process tracing; encoding-processing differences

Introduction

In sequential diagnostic reasoning multiple pieces of information have to be combined to find a best explanation for observed symptoms (e.g., Johnson & Krems, 2001). It is a complex cognitive process since the reasoner generates an undefined number of explanations for any number of observations (Johnson & Krems, 2001). Nevertheless, understanding this process is a major goal of research concerning reasoning and problem solving because of its high practical relevance. For instance, in the medical context, a complete understanding of diagnostic reasoning can help to save lives by improving the process of forming the right diagnosis (Mehlhorn, Taatgen, Lebiere, & Krems, 2011). But there are more applications such as finding the error in a technical system like a car or a computer (Johnson & Krems, 2001; Krems & Zierer, 1994; Mehlhorn et al., 2011). For instance, imagine you experience a loss in power of your car. Later, you witness some blue smoke coming from your exhaust pipe. Furthermore, you feel that recently, your car needs more oil as usual and your “check engine” light turns on. By combining these observations, you come up with the explanation that your car has an engine damage.

Diagnostic reasoning involves the processing of a number of observations and explanations. Often, the reasoner does not have all the necessary information available at once, but receives them in a sequential order. The reasoner then needs to integrate the symptom information into a situation model containing symptoms and explanations (Johnson & Krems, 2001; Johnson-Laird, Byrne, & Schaeken, 1992).

Besides this complexity, people are generally able to solve problems (Johnson & Krems, 2001), but how do they successfully engage in this demanding task? A number of process models (e.g., TAR: Johnson & Krems, 2001; TEC: Thagard, 1989; HyGene: Thomas, Dougherty, Sprenger, & Harbison, 2008) provide assumptions about ongoing memory processes. For instance, TAR assumes that *encoded* symptoms have to be *evaluated* concerning their fit with the current model of explanations (Johnson & Krems, 2001) and HyGene states that newly *encoded* information has to be judged concerning its implication for existing explanations (e.g., Thomas et al., 2008). Thus, in order to test these process assumptions, it is necessary to disentangle encoding and processing. Therefore, the process of diagnostic reasoning needs to be made visible. So far, different methodological approaches exist to trace memory processes in higher order cognitive tasks like judgement and decision making (e.g., Glaholt & Reingold, 2011; Schulte-Mecklenbeck, Kühberger, & Ranyard, 2011).

Recently, eye tracking is employed as a process tracing method to assess memory processes (see Glaholt & Reingold, 2011; Jahn & Braatz, 2014; Schulte-Mecklenbeck et al., 2011). Advanced hardware and improved understanding of its measures make it possible to get better insights (Jahn & Braatz, 2014; Renkewitz & Jahn, 2012; Scholz, von Helversen, & Rieskamp, 2015).

Monitoring eye movements allows the study of cognitive processes while participants interact with present objects without causing any restrictions and with minimal intrusions on the participant's behavior (Glaholt & Reingold, 2011). But interpreting eye tracking measures is challenging. The growing number of measures and their variety of applications complicate the assignment of specific measures to specific cognitive processes (Holmqvist et al., 2011).

Glaholt and Reingold, (2011) present a way to differentiate between encoding and processing in decision making by analyzing eye movement data. They asked participants to choose one picture out of a set of nine. They categorized dwells (defined as the sum of fixation durations of all fixations from the moment the gaze enters an area until it exits this area) into first-visit dwells and re-visit dwells, whereas re-visit dwells are the repeated viewing of an alternative. They found that *mean fixation durations per dwell* increases over time. The *number of fixation per dwell* is increasing between first-visits and re-visits as well. This indicates processing, such as evaluation of the current stimuli. The researchers interpreted these results as evidence that early in the process, participants merely screened the stimuli, moving to processing of the stimuli as they try to reach a decision. Even very early research finds evidence that the deliberate processing of information results in more and longer fixations (Loftus & Mackworth, 1978). As screening/*encoding* is assumed to happen implicitly without conscious control (Betsch, Hoffmann, Hoffrage, & Plessner, 2003) few fixations are plausible and in line with assumptions by Horstmann, Ahlgrimm, and Glöckner, (2009). *Increased* number of fixations, on the other hand, reflect processes such as the *evaluation* of an alternative which emerge later in the process of decision making (Horstmann et al., 2009). In contrast to Glaholt and Reingold (2011), Horstmann et al., (2009) did not find evidence that processing is associated with longer fixation durations.

Previous research on differences in eye movement patterns between encoding and processing of information in memory focuses on decision making. We assume that it is possible to use the knowledge about processes of decision making in the context of diagnostic reasoning as well. Encoding and processing differences are part of many models and assumptions. Whenever there is something to decide, we encode much information in a short period of time and subsequently evaluate and process this information (Glöckner & Betsch, 2008). As stated by TAR or HyGene (Johnson & Krems, 2001; Thomas et al., 2008), we assume that the same holds true when finding a best explanation for a set of symptoms.

Congruently, we assume that encoding and processing can be differentiated by the means of eye tracking measures, i.e. analyzing fixation duration and number of fixation per dwell. Following Glaholt and Reingold, (2011) and Horstmann et al., (2009) we assume that the number of fixation per dwell is increased during processes which appear later in the reasoning process such as symptom evaluation compared to the encoding of information.

In order to test our hypothesis, we used a task where information was presented in a sequential order. Remember the example with the engine trouble, people most likely witness one symptom such as the power loss at first and subsequently watch for more symptoms which are usually discovered one after another. In addition we need a task with the complexity comparable to that of everyday life problems. There were many symptoms pointing in the direction of engine damage, some of which might be related. The blue smoke is easily explained as soon as you know about the increased oil consumption. The car might burn some oil. On the other hand, we needed to control for prior knowledge of participants to get comparable results. Additionally, the task has to be learnable in an experimental setting. These requirements are met in Black Box task (BBX, Johnson & Krems, 2001). The Black Box task is a reasoning paradigm in which all information is, just as in our car example, *visuospatial* in nature. The symptoms are related to different areas of the car such as the exhaust pipe, the cockpit, or the dipstick in the motor compartment. Therefore, the Black Box task is especially suited to study memory processes during diagnostic reasoning with eye movements.

Method

Participants solved the Black Box task while their eye movements were recorded. In the Black Box paradigm, the participants' task is to determine a hidden state of a device using indirect evidence that needs to be combined following specific rules. Participants first learn and then apply the rules to get insight into the device by explaining observations with a combination of causes.

Participants

Twenty-nine students enrolled at Technische Universität Chemnitz took part in the experiment. One participant had to be excluded due to technical problems. Of the remaining participants 20 were female and 8 were male with a mean age of M = 22.3 (SD = 3). All participants had normal or corrected to normal vision.

Task and Apparatus

The participants' task is to locate hidden atoms in the Black Box which consists of a 10 x 10 grid by watching where light rays enter and exit the box. Participants do not see the path of the light rays. They only see the entrance and exit position of the light rays. As shown in Figure 1 each atom has a field of influence (a circle around the atom). Hitting this field, the light rays get reflected and exit the box, depending on where in the box an atom is located.

Participants have to place the atoms based on the information drawn from the rays that were shot in the box. The participant has to place as many atom markers at the grid as needed to explain the ray pattern of one trial without placing more atoms as absolutely necessary. During one trial, the participant watches a fixed number of rays in sequential order one at a time displayed by numbers marking the entrance and exit of the current ray (Figure 2). Participants have to remember where the rays entered and exited the Black Box as well as the atom locations of already set atoms, because, just as in the car example, symptoms might be related. For instance, it is possible that a previous pattern needs to be remembered to explain the current one (as illustrated in Figure 2: Ray five can be explained by ray three.) Therefore, the paradigm allows to measure memory processes that take place in sequential diagnostic reasoning. The participant decides when to move on to the next ray by pressing the space bar. How many rays are left during one trial is shown at the upper left corner of each Black Box (cue). An exemplary trial is pictured in Figure 2.

The study recorded gaze data using a binocular IViewX RED eye-tracking system from SensoMotric Instruments with a sampling rate of 120 Hz. Data was analyzed with BeGaze 3.0, Microsoft Excel 2007 and IBM Statistics 23 (SPSS).

Stimuli were presented on a 22-inch computer screen using EPrime 2.0 software with a resolution of 1680 × 1050 pixels. All subjects were seated at a distance of 600 to 800 mm in front of the screen.

Procedure

Each participant was tested individually. After an initial instruction phase, in which participants were familiarized with the rules of the Black Box, two training phases followed. During the first training phase, rays and atoms were visible throughout the trial and therefore did not have to be remembered. In the second training phase participants solved the trial under test conditions (memory-based). That is, they only saw where the current ray of light entered and exited the Black Box and the cue in the upper left corner (showing how many rays are left in the current trial). They saw the atom and its field of influence as soon as they placed it by using the mouse until they demanded a new ray shot into the Black Box. No prior rays or atoms were present. Each training phase consisted of seven trials. A calibration and 48 test trials followed. The experiment ended after a short survey containing demographic questions. Participants needed 53 to 127 minutes to complete the entire experiment (*M* = 85.2 min; *SD* = 21.1).

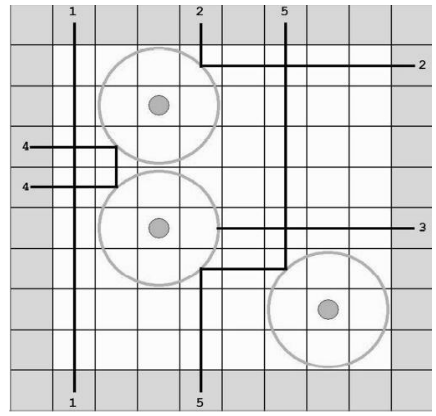

Figure 1: Rules in the Black Box.
If the ray does not hit any atom, it goes straight through the Black Box (1). If a ray hits the field of influence at an angle, it is reflected 90 degree. This results in a L-pattern (2). If a ray hits a field of influence of an atom straight forward, it is absorbed (3). The ray does not exit. Hitting a second field of influence after being already reflected the light ray path results in an U-pattern (4) or a Zick-Zack-pattern (5). Note, light ray paths and atom locations are not visible for the participant and have to be inferred. In this example all five observations can be explained by three atoms.

Analysis

To test our hypotheses, quadratic Areas of Interests (AOIs) were drawn around each square of the grid resulting in 100 separate AOIs, also called gridded AOIs (see Holmqvist et al., 2011). Furthermore, we coded the AOIs where the rays entered and exited the Black Box, where they hit the field of influence of an atom and where atoms had to be set.

Following Glaholt and Reingold, (2011) we identified the order of dwells in the AOIs and compared first-visit and re-visit dwells. First-visits are termed the first dwell in each AOI. Re-visits are all *repeated dwells in that AOI*. Re-visit dwells integrate all dwells from the second till the last dwell per AOI. In this point we slightly differ from the methods applied by Glaholt and Reingold, (2011) who defined *every* dwell as a re-visit that follows after any AOI is viewed a second time, even if the gaze hits an AOI not viewed yet. In decision making the transition between encoding the alternatives and their evaluation is believed to be the point in time where the first alternative is viewed a second time (Russo & Leclerc, 1994). Since our material is relatively complex we expect participants only to be able to encode information if they actually look at it. If the gaze shifts to a different AOI, information need to be encoded and subsequently evaluated as well. Therefore, for diagnostic reasoning, we expect that the first dwell into an AOI reflects encoding and that all subsequent dwells in that same area represent evaluation processes since participants test the implication of the information for the current explanation (e.g., Johnson & Krems, 2001).

We expect every first-visit of an area to map onto the encoding of that area. Even if a dwell consist of only one fixation; we are confident that participants encoded this location sufficiently during their first visit since we defined a fixation with a minimum duration of 80 ms which is enough time to gather information since 50 to 60 ms are assumed enough time to encode even words (Rayner, 2009). During every re-visit of that same area we expect participants to process this location since it is already encoded.

For every dwell we calculated the mean number of fixations and the mean fixation duration. We aggregated first-visit dwells and re-visit dwells over every symptom presentation, trial and participant.

Results

Participants tended to look mostly at task relevant areas. Aggregated over symptom presentations, trials and participants 12.08 % of the fixation time was on average directed to the AOI containing the atom (SD = 9.05). The participants looked mostly to the AOI where the rays hit the field of influence of an atom (M = 21.16 %, SD = 14.76) and 14.63 % of the fixation time fell on average to the two AOIs covering the rays (SD = 7.00). 5.24 % of the time participants fixated on the AOI with the cue in the upper left corner (SD = 4.15). Taken together, participants looked more than 50% of the time to direct task-relevant AOIs which covered only about 5 % of the grid. Fixation duration weighted by the number of AOIs, participants looked significantly longer to the atoms ($t(27) = 19.82$, $p < .001$, $d = 3.81$), to the field of influence of the atom ($t(27) = 20.38$, $p < .001$, $d = 3.92$), to the rays ($t(27) = 20.45$; p = <-001, $d = 3.80$) and to the cue in the upper left corner ($t(27) = 12.86$, $p < .001$, $d = 2.47$) than to all the other AOIs.

They found correct solutions for most of the trials (M = 71.3%, SD = 0.13). Therefore, we conclude that participants have understood task instructions and engaged in diagnostic reasoning as we intended.

The calibration procedure reached a satisfying accuracy with a mean deviation of $M = 0.54°$ (SD = 0.53).

We assumed that encoding and processing can be differentiated by the means of the eye tracking measures analyzing dwells concerning fixation duration and fixation frequency. More precisely, our hypothesis predicted that fixation duration and number of fixations per dwell is increased during re-visits compared to first-visits, assuming that first-visits represent encoding whereas re-visits map onto processing.

Over all, we find longer fixations during re-visits than first-visits ($M_{first\text{-}visits}$ = 267 ms, SD = 45, $M_{re\text{-}visits}$ = 279 ms, SD = 51, $t(27)$ =- 1.94, p = .06, d = 0.37) and we find significant more fixations during re-visit dwells compared to first-visit dwells ($M_{first\text{-}visits}$ = 1.25, SD = 0.09, $M_{re\text{-}visits}$ = 1.31, SD = 0.10, $t(27)$ =- 4.11, $p < .001$, d = 0.79).

An analysis of the different pattern separately shows that there are differences. To illustrate this, we exemplary report the L-pattern and the absorption in Table 1.

The increase in fixation duration and the significant increase in number of fixations per dwell during re-visits compared to first-visits support our hypothesis that encoding and processing can be differentiated by the means of the eye tracking measures as Glaholt and Reingold, (2011) assumed.

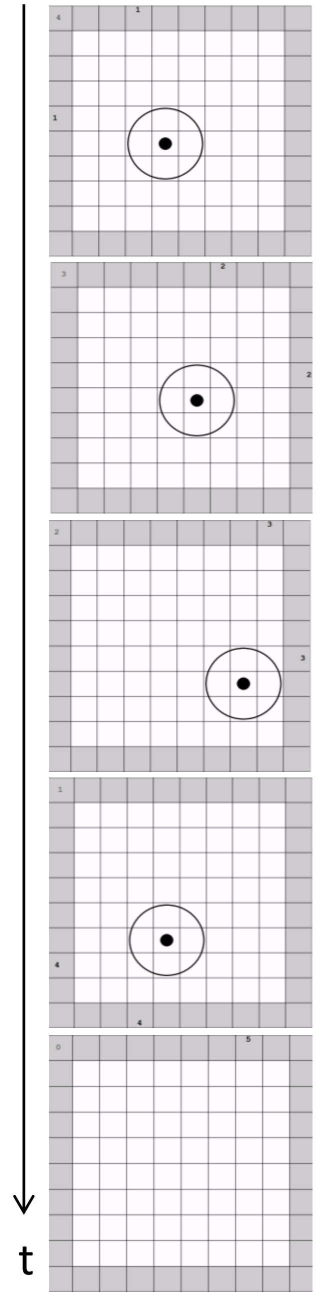

Figure 2: Exemplary trial
The trial consists of five rays shot into the Black Box sequentially. For each ray the participant can place atoms to explain them. When finished with one ray the participant decides when to move on to the next one. The number (cue) in the upper left corner shows how many more rays are to come in the current trial. In this example the participant should notice that no new atom need to be placed for ray five since this ray can be explained by the atom already set for ray three.

Table 1: Analysis of L-pattern and Absorption

		L-pattern	Absorption
Mean fixation duration per dwell	First-visit *M (SD)*	277ms (51)	255 ms (44)
	Re-visit *M (SD)*	279 ms (51)	280 ms (64)
	t-test	$t(27) =- 0.27$, $p =$ .79 , $d = 0.05$	$t(27) =- 2.31$, $p = .03$, $d = 0.44$
Number of fixations per dwell	First-visit *M (SD)*	1.26 (0.09)	1.19 (0.11)
	Re-visit *M (SD)*	1.34 (0.13)	1.27 (0.15)
	t-test	$t(27) =- 4.46$, $p <$ .001, $d = 0.83$	$t(27) =- 2.59$, $p = .02$, $d = 0.50$

Discussion

This study aimed to find a measure which is able to describe reasoning processes differentiating between encoding and processing. Therefore, we employed eye tracking as a process tracing method because it is an objective and fine grained measure to human behavior. Following research by Glaholt and Reingold, (2011) and Horstmann et al., (2009) we assumed that encoding and processing can be differentiated by analyzing dwells concerning fixation duration and fixation frequency.

Horstmann et al., (2009) and Glaholt and Reingold, (2011) showed that a higher number of fixations might be an indicator for more deliberate thinking and processing. Since we assume encoding to be more intuitive and number of fixations increases during re-visits when processing takes place, our results support the research of Horstmann and colleagues (2009) as well as Glaholt und Reingold, (2011). Over all patterns as well as for L-pattern and absorption individually increases the number of fixation with medium to high effects. In line with the research by Horstmann et al., (2009) the analysis of fixation duration draws a less clear picture. It might be that there is not such a clear difference in fixation duration between (more intuitive) encoding and (deliberate) processing as Horstmann et al., pointed out. The researchers assume that intuitive and deliberate processes may have similar underlying mechanisms as indicated by no clearly distinct differences in fixation durations. The differences between encoding and processing might not be qualitative in nature but mainly quantitative regarding information sampling. However clear differences between L-pattern and absorption might indicate that both patterns are not strictly similar processed. Horstmann et al., (200) state that there may still be crucial differences regarding intuitive and deliberate processes, which cannot be shown by this measure. Since absorptions only identify the row or column of the atom location, a second ray of light is needed to determine the exact location in the Black Box, making the absorption a more complex pattern. It remains to future research to investigate the differences in processing of different rules. Therefore the analysis of U-pattern and Zick-Zack- patterns may be useful, as they are more complex as well. They are not included in the results of this paper since those patterns are not used frequently enough throughout the experiment to produce a reliable data basis.

We rule out the possibility that participants did not process during re-visit dwells but encoded new features of already encoded information. Since a single AOI contained simple material with not much features like colors or different shapes we consider this alternative very unlikely in this specific task.

Further, we are confident that participants actually encoded and processed what they were looking at. The fact that participants reached fairly good response accuracies in spite of a difficult and strongly memory based task speaks in favor of this assumption. There was no evident reason to separate eye movements and attention while solving the Black Box task.

Since looking back to associated, but emptied spatial locations can facilitate memory retrieval (e.g., Scholz, Mehlhorn, & Krems, 2016), it is an interesting questions whether there is a connection between gaze behavior and accuracy during diagnostic reasoning, too. Do people who solve the task successfully show a different gaze pattern? Since our participants reached a fairly high response accuracy in solving the trials, it is difficult to draw causal conclusions to this question from our data. Future research should address this question by careful manipulations.

We were able to replicate Glaholt and Reingolds (2011) results that an increased number of fixations to an already viewed area is an indicator for evaluation and processing of information. However, it is difficult to make a clear statement regarding the mean fixation duration per dwell.

In conclusion, eye movements show a high potential to assess memory processes during higher-order reasoning and thinking (see also Jahn & Braatz, 2014; Renkewitz & Jahn, 2012; Scholz et al., 2015). In particular, the number of fixations per dwell may be used as process tracing measures to asses processing during diagnostic reasoning. We found increased number of fixations during a dwell to indicate processing of information. The fact that we found a measure that indicates evaluation processes allows for a more detailed testing of process assumptions derived by process models on diagnostic reasoning such as TAR (Johnson & Krems, 2001), HyGene (Thomas et al., 2008) or TEC (Thagard, 1989). All these models make assumption about the manner in which the reasoner links observation or symptoms to a set of explanations. Thereby the models differ in their predictions to how this link is processed. TEC for instance, describes this process as a judgment to which extend symptoms and explanations are coherent (Thagard, 1989). Thereby, the encoding and evaluation of all information to reach a judgement is supposed to happen parallel. HyGene on the other hand, gives insight about the way explanations are generated and used to asses new information (Thomas et al., 2008). HyGene states that new information are encoded sequentially and have to be judged concerning their implication for existing explanations. TAR describes diagnostic reasoning as a deliberate comprehension process in which observations are sequentially interpreted and integrated into a mental model (Johnson & Krems, 2001). TAR assumes that one information is encoded and evaluated before new information is gathered, making encoding and processes closely intertwined. All models have in common that they assume perception-action cycles with different encoding and processing phases. To describe reasoning, we need to know *how* people *use* information to reach causal conclusions. Therefore the first step is to know when people simply gather information and when they actually use it for evaluation, comparison or processing.

The process tracing measure presented in this paper may help to disentangle encoding and processing assumptions stated by different models giving more insight into diagnostic reasoning by providing a possibility to identify processing phases and test for precise research questions.

Acknowledgments

We thank Martin Baumann for helpful comments.

References

Betsch, T., Hoffmann, K., Hoffrage, U., & Plessner, H. (2003). Intuition Beyond Recognition: When Less Familiar Events Are Liked More. *Experimental Psychology*, *50*(1), 49–54. doi:10.1026//1618-3169.50.1.49

Glaholt, M. G., & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 125–146. doi:10.1037/a0020692

Glöckner, A., & Betsch, T. (2008). Modeling option and strategy choices with connectionist networks : Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Makingg*, *3*(3), 215–228.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye Tracking: A comprehensive guide to methods and measures* (1rst Editi.). New York: Oxford University Press.

Horstmann, N., Ahlgrimm, A., & Glöckner, A. (2009). How distinct are intuition and deliberation ? An eye-tracking analysis of instruction-induced decision modes. *Judgment and Decision Making*, *4*(5), 335–354. doi:10.2139/ssrn.1393729

Jahn, G., & Braatz, J. (2014). Memory indexing of sequential symptom processing in diagnostic reasoning. *Cognitive Psychology*, *68*, 59–97. doi:10.1016/j.cogpsych.2013.11.002

Johnson, T. R., & Krems, J. F. (2001). Use of current explanations of multicausal abductive reasoning. *Cognitive Science*, *25*, 903–939. doi:10.1016/S0364-0213(01)00059-3

Johnson-Laird, P. N., Byrne, R. M., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, *99*(3), 418–439. doi:10.1037/0033-295X.99.3.418

Krems, J. F., & Zierer, C. (1994). Sind Experten gegen kognitive Täuschungen gefeit? Zur Abhängigkeit des confimation bias von Fachwissen. *Zeitschrift Für Experimentelle Und Angewandte Psychologie*, *XLI*(1), 98–115.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology. Human Perception and Performance*, *4*(4), 565–572. doi:10.1037/0096-1523.4.4.565

Mehlhorn, K., Taatgen, N. a., Lebiere, C., & Krems, J. F. (2011). Memory activation and the availability of explanations in sequential diagnostic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1391–1411. doi:10.1037/a0023920

Rayner, K. (2009). *Eye movements and attention in reading, scene perception, and visual search. Quarterly journal of experimental psychology (2006)* (Vol. 62). doi:10.1080/17470210902816461

Renkewitz, F., & Jahn, G. (2012). Memory indexing: A novel method for tracing memory processes in complex cognitive tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1622–1639. doi:10.1037/a0028073

Russo, J. E., & Leclerc, F. (1994). An Eye-Fixation Analysis of Choice Processes for Consumer Nondurables. *Journal of Consumer Research*, *21*(2), 274. doi:10.1086/209397

Scholz, A., Mehlhorn, K., & Krems, J. F. (2016). Listen up, eye movements play a role in verbal memory retrieval. *Psychological Research*, 149–158. doi:10.1007/s00426-014-0639-4

Scholz, A., von Helversen, B., & Rieskamp, J. (2015). Eye movements reveal memory processes during similarity- and rule-based decision making. *Cognition*, *136*, 228–246. doi:10.1016/j.cognition.2014.11.019

Schulte-Mecklenbeck, M., Kühberger, A., & Ranyard, R. (2011). The role of process data in the development and testing of process models of judgment and decision making. *Judgement and Decision Making*, *6*(8), 733–739.

Thagard, P. (1989). Explanatory Coherence. *Behavioral and Brain Sciences*, *12*, 435–502.

Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review*, *115*(1), 155–185. doi:10.1037/0033-295X.115.1.155