The Acquisition of Mental Representations under Uncertainty: An Eye Movement Study

Johanna Renker* Gerhard Rinkenauer*

*Leibniz Research Centre for Working Environment and Human Factors, TU Dortmund
Germany (Tel: 0049 231 1084 460; e-mail: renker@ifado.de)

Abstract: Users interact with technical systems based on their mental model of the system. Both, the mental model itself and the situation inhere uncertainty, because mental models are only a reduced representation of the reality and complex systems do not provide a complete set of information. However, users are able to develop strategies to cope with uncertainty and finally, to reach satisfying results. In this study we investigated the development of mental models during the performance of an uncertain task via eye movements. Main results showed decreasing visual search activity parallel to an increasing learning level and hence decreasing subjective uncertainty. In addition, eye movement parameters seem to be able to differentiate between high and low performers in the way that low performers showed more visual search activity. These findings may be of relevance for human-machine interaction by using eye movement patterns to detect individual learning problems and the degree of the user’s subjective uncertainty. Based on the eye movement analysis, the information content of the technical system could be adapted to the user’s learning curve to improve the interaction even in the context of uncertain systems and tasks. However, further studies are necessary to proof the stability of the concept.

Keywords: Eye tracking, mental model, probability concept, uncertain environment

1. INTRODUCTION

Although guidelines for the design of interfaces exist human-computer interaction is still characterized by erroneous operations, especially when interacting with a technical system for the first time without reading any instructions. Users who are initially confronted with a technical system develop an internal representation or mental model of the system which serves as the basis for their interactions with the system. Such models are a reduced representation of the reality and can be considered as knowledge structures in the long-term memory that represent the user’s understanding of the system functioning (Durso and Gronlund, 1999). According to Johnson-Laird and Byrne (2012) the mental model theory assumes that “mental models represent explicitly what is true, but not what is false” and therefore might lead to systematic errors. The latter authors also assume that erroneous conclusions increase with the number and the complexity of mental models and that invalid inferences during reasoning based on the mental model are only refuted by counterexamples. Such assumptions suggest that mental models never reflect a complete image of the reality, but rather a reduced representation of what is true. In addition, complex technical systems also cannot provide complete information of the system state. Thus, the mental model as well as the technical system inhere a certain amount of uncertainty users have to deal with. However, it depends on the user’s personality if uncertainty is used as a motivating factor or as an inhibiting factor (Smithson, 2009). There are different strategies to cope with uncertainty. Users might search for additional relevant information to reduce uncertainty or defer the task until additional information is available. Suppressing uncertainty includes tactics of ignoring or distorting ambiguous information (Lipshitz and Strauss, 1997).

Visual search strategies to collect relevant information can be assessed by eye movement analysis. Eye movements as quick, frequent and automatic actions (Spivey and Dale, 2011) reflect to which part of information attention is drawn and thereby might gain insights into patterns or strategies of information collection and additionally into the development of mental models. In contrast to other common methods used in learning research to probe cognitive activities like interviews based on the think-aloud protocol, eye-tracking data inform about online cognitive activities without conscious awareness (Lai et al., 2013). Just and Carpenter (1980) already postulated a causal relationship between attention and working memory by formulating the eye-mind assumption positing “that there is no appreciable lag between what is being fixated and what is being processed.” (p.331). In human-computer interaction more fixations per area of interest indicate that this area is more striking or important than other ones (Poole, Ball and Phillips, 2005) and longer fixation durations indicate difficulties in encoding information or a higher engagement (Just and Carpenter, 1976). Goldberg and Kotval (1999) studied eye movements during the use of good and poor interfaces. They found higher numbers of overall fixations and longer scan path lengths leading to a less efficient search. In conclusion, they emphasize the importance of interface design to enhance user’s visual search strategies. For example, in process control it is assumed that device knowledge as well as display
knowledge is necessary to build up an adequate mental model about a technical system (Kluwe, 1997). However, even if the interface is well designed, some additional requirements have to be fulfilled to develop accurate mental models based on visual stimuli. First, attention has to be directed towards the new visual stimuli to be able to encode the visual information (Mulligan, 1998). Second, there has to be enough capacity in visual working memory to store information (see for a review Brady, Konkle and Alvarez, 2011). If both conditions are met, visual information can be processed this means knowledge will be acquired, processed, stored, retrieved and thus, an internal representation can be developed. Schumacher and Czerwinski (2014) report three stages during the acquisition of mental models. In the first pretheoretic stage, similar instances in memory are used to understand the system performance viz. this stage bases on earlier experiences. In the second stage called “experiential stage”, causal relationships emerge to a certain amount and first abstractions seems to take place. Finally, the expert stage is characterized by an abstraction across different system representations. Users in this stage are able to recognize patterns, to retrieve former knowledge and to transfer knowledge. Therefore, experts’ mental models are more abstract and consider fewer irrelevant details than novices’ models.

Aim of the present study was to gain deeper insights in the mental model acquisition and representation of uncertain information. Here, uncertainty in the sense of task uncertainty was defined as a lack of knowledge about the cause-and-effect relationship (cf. Thompson, 1967). Stimulus-outcome associations had to be learned in an uncertain setting via outcome feedback in order to develop the most accurate mental model of the task as possible. The pretheoretic stage in this task did not contain the search of similar instances in memory, but rather the accumulation of new information, because prior knowledge about the underlying processes of the task – a probability concept - did not exist. Thus, all participants started on the same uncertain stage and had to develop a completely new mental model. Therefore, feedback mechanisms were essential to understand the probability structure and to improve decision strategies.

In this study we assumed that eye movement patterns inform about the stage of learning and the participants’ degree of subjective uncertainty during the performance of the task in addition to the conscious understanding which was established at the end of the experiment. In more detail we expected extensive information search at the beginning and more focussed visual search at the end of the experiment. Thus, gaze shifts, fixation frequency and fixation duration should decrease in parallel to an increasing task performance. In addition to these within effects we expected differences between high and low performing participants. The two groups should show distinct eye movement patterns whereby it was expected that high performer show less visual search behaviour than low performer because uncertainty is already reduced and knowledge acquired. Furthermore, we hypothesized stronger effects, especially stronger learning effects at the beginning of the experiment than at the end according to Rieskamp and Otto (2006).

2. METHOD

2.1 Participants

A total of 18 participants (9 female, mean age=24.78 years, SD=3.12 years), mainly psychology students from the Technical University Dortmund, participated for course credits or refund (10€ per hour) and provided written informed consent. All participants had normal vision and were right-handed.

2.2 Apparatus

Participant’s eye movements were recorded with a SMI RED 500 remote eye tracking system using the corneal reflection technique to detect the gaze position on the screen. This system records with a sampling rate of 500 Hz and allows an accuracy of 0.4°. Stimuli were presented on a white background of a 27 inch monitor set to a resolution of 1920x1080. Participants were seated 70 cm from the computer screen with head positioned on a chinrest to minimize head movement.

2.3 Materials and design

Stimuli were constructed based on a new developed task: the occluded visual spatial search task. This task was divided into two consecutive parts: a prediction and a reaction part. First, participants observed dark grey geometric objects (circle, triangle, square) with a diameter of 2 cm disappearing into a quadratic dark room (20x20 cm) with three exits (2 x 0.5 cm). Each object was associated with a higher probability to one of the exits. Participants had to predict as accurate as possible with the arrow keys of the keyboard at which of the exits these objects will reappear within max 2.8 sec. If the colour intensity of the reappearing object has been changed, participants had to press the appropriate arrow key again. If the colour was the same as before, participants had just to observe the situation (see Fig. 1 for a schematic description). In order to perform the task as accurate as possible participants had to learn the underlying probability concept of the task. As already mentioned every object was associated to one of the exits with a higher probability (74%) and to the other exits with a corresponding lower probability (11%). In 4% of the cases objects reappeared at the bottom entrance to create a rare occurrence. Auditory feedback indicated erroneous task performance e.g. if the arrow key was pressed too late or too early. Overall 324 trials of the visual spatial search task were tested subdivided into four blocks to investigate learning effects. Between each block there was a fixed pause of 2 minutes. The eye tracker was calibrated with a 9-point calibration before every block and feedback about the percentage of correct predictions was displayed after every block.

2.4 Procedure

Each participant was tested in a single session which lasted about one and a half hour. A computer version of the D2-test
of attention (Brickenkamp, 1994) was performed before and after the visual spatial search task to check for concentration decline as a confounding variable. We controlled also for motivation with the subscale “interest” of the Questionnaire of Current Motivation (QCM) on a seven-point rating scale ranging from “disagree” to “agree” (Rheinberg, Vollmeyer and Burns, 2001). Each participant completed a practice block (18 trials) of the visual spatial search task with geometric objects that were different from the experimental ones before the actual run started. The conscious understanding of the probability concept used in the visual spatial search task was inquired at the end of the experiment via questionnaire.

![Image](image.png)

Fig. 1. Schematic description of the visual spatial search task:
Participants fixate a fixation cross for 750 ms, before the object appears at the bottom entrance. The object stays 1 sec at the bottom entrance and then moves inside the dark room. The object takes 1 sec from the position at the bottom entrance until it completely disappears in the room. Participants predict the exit of reappearance within 2 sec as soon as the object is disappeared into the room. Finally, the object moves out of the room within 2 sec and stays at the exit position for 1 sec. Thus, the task lasts 7 750 ms.

2.5 Data Analysis

All trials with missing predictions and/or reactions and trials with less than 65% of the eye movement data were excluded from data analysis. Furthermore, all data points within a radius of 100 pixel around the middle of the room were not considered in data analysis. The saccade-detection algorithm of SR Research (Tatler, 2007) was used to detect fixations and a minimum fixation duration was set to 50 ms. Gaze shifts between four equal areas of interests (AOI) around the three exits and the entrance were counted.

Statistical analysis was performed using a two-way repeated measures ANOVA with the within-subjects factors prediction (true, false) and block (1-4). Independent variables were three eye movement parameters: number of gaze shifts, fixation frequency, fixation duration and judgement time defined as the time interval from the starting point of the prediction – when the object disappeared into the room - until participants actually made their prediction. Performance defined as the number of correct predictions was used as further dependent variable. The effect of control variables was measured using correlation analysis and repeated measures ANOVA. Whenever necessary, violations of sphericity were corrected by using Greenhouse-Geisser or Huynh-Feldt ε. An adequate method to analyse low and high performing participants would be random coefficient modelling. However, this analysis would require larger sample sizes to provide a sufficient statistical power (cf. Hox, 1998). Therefore, we computed analogous to Kluwe (1997) a hierarchical cluster analysis with regard to task performance, viz. the number of correct predictions in order to split participants in performance groups (high performer vs. low performer). The minimum threshold concerning the sample size for running a cluster analysis is known to be 25 participants (k=number of variable) and thus, complies with the tested sample size of the current experiment (cf. Forman, 1984). Independent t-tests were run to compare eye movement parameters (number of gaze shifts, fixation frequency, fixation duration) between low and high performer.

3. RESULTS

Repeated measures analyses of variance showed that all eye movement parameters decreased significantly over blocks: fixation frequency \( F(3;51)=5.32, p=0.012, \eta_p^2=0.238 \) (Fig. 2A) as well as fixation duration \( F(3;51)=4.00, p=0.045, \eta_p^2=0.190 \) (Fig. 2B) and number of gaze shifts \( F(3;51)=11.16, p=0.001, \eta_p^2=0.396 \) (Fig. 2C). Judgement time also decreased significantly over blocks \( F(3;51)=11.61, p=0.001, \eta_p^2=0.285 \) (correct predicted: Block1 \( M=437, SD=184 \); Block4 \( M=346 \), \( SD=140 \); false predicted: Block1 \( M=446, SD=194 \); Block4 \( M=357 \), \( SD=143 \)). In parallel, the number of correct predictions increased over blocks \( (Block1: M=47.11, SD=9.72; Block4: M=54.78, SD=8.75) \), but only the increase from Block1 to Block2 was significant \( F(1;17)=17.14, p=0.001, \eta_p^2=0.505 \). Furthermore, judgement time was significantly shorter \( F(1;17)=6.77, p<0.019, \eta_p^2=0.406 \) and the number of gaze shifts \( F(1;17)=37.86, p<0.001, \eta_p^2=0.690 \) (Fig. 2C) as well as fixation frequency \( F(1;17)=23.69, p<0.001, \eta_p^2=0.582 \) (Fig. 2A) was significantly lower for correct predicted than for false predicted trials indicating more search behaviour during false predictions.

<table>
<thead>
<tr>
<th>Tab. 1. Cluster statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster</td>
</tr>
<tr>
<td>-number of gaze shifts</td>
</tr>
<tr>
<td>fixation frequency</td>
</tr>
</tbody>
</table>

Note. 1=low performer, 2=high performer.
A hierarchical cluster analysis was used to group participants into low performers (n=6) and high performers (n=12) based on the number of correct predictions. Low performers showed significantly more gaze shifts t(16)=3.057, \( p=0.025 \) and had a higher fixation frequency t(16)=3.392, \( p=0.004 \) than high performers (see Tab. 1). However, fixation duration did not differentiate between the clusters t(16)=0.348, \( p=0.732 \).

Tab. 2. Memory retrieval and behavioural probabilities

<table>
<thead>
<tr>
<th>Object &amp; Position</th>
<th>Subjective Probability Concept</th>
<th>Performed Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>left</td>
<td>16 %</td>
<td>7 %</td>
</tr>
<tr>
<td>top</td>
<td>71 %</td>
<td>89 %</td>
</tr>
<tr>
<td>right</td>
<td>13 %</td>
<td>4 %</td>
</tr>
<tr>
<td>left</td>
<td>16 %</td>
<td>5 %</td>
</tr>
<tr>
<td>top</td>
<td>14 %</td>
<td>8 %</td>
</tr>
<tr>
<td>right</td>
<td>70 %</td>
<td>87 %</td>
</tr>
<tr>
<td>left</td>
<td>66 %</td>
<td>87 %</td>
</tr>
<tr>
<td>top</td>
<td>18 %</td>
<td>9 %</td>
</tr>
<tr>
<td>right</td>
<td>16 %</td>
<td>4 %</td>
</tr>
</tbody>
</table>

Note: More likely object-exit associations are shown in bold.

Overall, the participants’ estimation of the probability concept via questionnaire was close to the given probabilities. Participants estimated the relation of the object to the exits on average with 67% for the more likely exit and 15.5% for the unlikely exits. Thus, lower probabilities were overestimated and higher probabilities were underestimated. Behavioural data showed a reversed pattern. Participants predicted the exit with the higher probability more often than the object actually reappeared there (see Tab. 2). This corresponds with the optimal response strategy - predicting the exit with the higher probability - to maximize the number of correct predictions (see TTB - Take The Best heuristic: Dougherty, Franco-Watkins and Thomas, 2008). Response strategies did not change after the rare occurrence in 82% of the cases (in 193 of 235 trials) viz. the response in the next trial after the rare occurrence with the same object as in the rare occurrence trial still matched with the optimal response strategy participants mainly used.

The analysis of control variables showed that concentration of participants was not affected by the visual spatial search task. Participants performed significantly more cases of the D2-test after the experiment than before \( F(1;17)=47.32, \ p<.001 \). During the pre-test they performed 71% of the cases (\( M=207.39; \ SD=30.32 \)) and during the post-test 79% of the cases (\( M=230.83; \ SD=33.55 \)) presumably due to a learning effect. The error rate did not change from the first to the second run of the D2-test \( F(1;17)=0.05, \ p=.828 \).

The Pearson product-moment correlation coefficient showed a non-significant (\( \alpha=0.05 \)) unilateral correlation (\( r=-.368, \ p=0.066 \)) between the subscale interest of the QCM (\( M=3.82; \ SD=0.95 \)) and performance – the number of correct predictions. Therefore, interest in the task did not seem to affect the performance results to a great extent.

Fig. 2. Fixation frequency (A), fixation duration (B) and number of gaze shifts (C) decreasing over blocks for correct predicted (true) and false predicted (false) trails. Error bars show the standard error.
4. DISCUSSION

In this study we investigated eye movement patterns during the acquisition of mental models under uncertain task conditions. Participants performed a newly developed occluded visual spatial search task improving their performance by learning an underlying probability concept. Despite of the task uncertainty, participants were able to develop a mental representation of the probability concept that was nearly accurate. It seems that participants showed both strategies described in the literature to cope with uncertain situations: information search and suppression of ambiguous information (Lipshitz and Strauss, 1997). The reported results suggest that participants search for relevant information at the beginning of the task extensively due to the high uncertainty. During the subsequent course of the task, increasing expertise led to more focused visual search, lower error rates and lower uncertainty. Eye movement patterns also enabled to differentiate between high and low performer whereby low performer showed more visual search activity. In addition, behavioural data showed that participants ignored ambiguous information and maintained their developed decision strategy.

As common in literature we also found that participants underestimated high probabilities and overestimated low probabilities while estimating the subjective probability concept due to a tendency to the mean (Beuer-Krüssel and Krumpal, 2009). However, overall probability estimation was close to the real probability concept, even if probabilities as numbers are quite abstract. It seems that participants were able to adapt their conscious probability estimation according to their decision strategy. They presumably took into account that they made some false predictions, especially while predicting the likely exit more often. Furthermore, the high probability of 74% was facilitating the probability estimation by pointing clearly into one direction. Jungermann, Pfister and Fischer (2010) illustrate in their book one possible relation between verbal and numeric terms within 25%-steps ranging from 0% – never to 100% - always. They suggest that 75% is associated with “often” and that might be the anchor for the estimation in the current experiment. Further studies might investigate probability distribution which are harder to differentiate and thus more uncertain.

As hypothesized eye movement patterns informed about the stage of learning and thereby about the degree of subjective uncertainty, viz. the number of gaze shift, the fixation frequency and the fixation duration decreased over blocks in parallel with the increase of knowledge about the cause-and-effect relationship. Learning effects seem to be stronger from Block1 to Block2 corresponding to the findings of Rieskamp and Otto (2006). Additionally, as shown in Figure 2A fixation frequency was higher at the beginning of the experiment where new information was presented which might be due to the assumption that only during fixations new information is processed (Rayner, 2009). These findings might indicate that the development of the mental model mainly took place rather early (in the first two blocks) presumably due to the low complexity and high consistency of the task (cf. Ackermann, 1988). After the second block good performer might already be in the expert stage where participants are able to recognize patterns and to retrieve former knowledge for the development of an optimal decision strategy under uncertainty according to Schumacher and Czerwinski (2014). Eye movement patterns seem to give also insights into the individual performance. Participants with low performance showed according to Goldberg and Kotval (1999) a higher fixation frequency indicating inefficient visual search behaviour and more gaze shifts presumably due to a lower learning rate. These participants might have more difficulties to cope with task uncertainty and to use coping strategies appropriately. Another reason for these results might base on interindividual differences with regard to working-memory capacity and information-processing abilities (Jipp, 2016; Sweller, 2003). Participants could be supported by guiding their attention via a model’s eye movement that improves visual search and enhances interpreting relevant information (Jarodzka et al., 2013).

Not only eye movement patterns changed over time, also judgement time decreased over blocks. Participants made their predictions considerably faster at the end of the experiment than at the beginning. That might indicate a kind of proceduralization in decision making (Pomerol and Brézillon, 2003) viz. rules of the developed decision strategy based on the mental model were automatically retrieved from memory. Such a rule could be: whenever the circle appears predict the top exit. Thus, every object might be associated with the appropriate likely exit neglecting small probabilities. This tactic would lead to an optimal performance according to the TTB heuristic in the current experiment (Dougherty, et al., 2008). An alternative explanation may be a three phase account as suggested by Ackermann (1988). The first phase demands learners’ general ability due to cognitive information processing. In the second phase knowledge is compiled and thus, perceptual speed is needed. In the third phase noncognitive psychomotor skills are essential as processes become more autonomous which might be comparable with the proceduralization of decision making also found in the current study. Performing the occluded visual spatial search task might also be divided in these phase. First, information of the presented stimuli have to be processed demanding general abilities as in the first phase described by Ackermann (1988). Then, associations between these stimuli (objects and exits) have to be learned and thus, perception becomes crucial for the task performance comparable with the second phase. Finally, predictions were made automatically requiring no further cognition as described earlier matching with Ackermann’s third phase.

In research, uncertainty is usually equated with probability (Smithson, 2009). Thus, the new occluded visual spatial search task was developed and designed around a probability structure. On the one hand, the task has some advantages. First, it seems that the task fulfills the earlier mentioned requirements. The appearance of the object at the bottom entrance as well as the movement seems to gain attention via bottom up activation according to the pop-out effect (cf. Wolfe, 1994). Furthermore, only one object per trial – one chunk – has to be processed in visual working memory, so that there should be sufficient capacities. Second, the task
enables to compare eye movements that are mainly automatic with behavioral data and the conscious understanding of the task – the subjective probability concept. Third, the visual spatial search task allows to vary the degree of uncertainty by choosing a different probability distribution of the probability concept. On the other hand, there might be a bias concerning the shape of the objects. The association between the circle and the top exits seems to be easier to learn than the association between the square and the left exit. One reason for this might be a top-to-bottom bias due to the gravity leading objects to fall down (Spalek and Hammad, 2005). A circle could be a flying object and might fit into this bias facilitating learning whereas the square does not allow such a connection. Further studies should use different kind of objects e.g. Gabor-figures without information about the direction to avoid such biases. Practitioners may consider the reported results to detect individual learning problems and the degree of the user’s subjective uncertainty. Subsequently, they could adapt the information content to the user’s individual learning curve to improve human-computer interaction even in the context of uncertain systems and tasks. One main advantage of the eye-tracking technique in comparison to oral reporting is that users do not have to be aware of problems and do not have to comment them which is often time-consuming and might lead to misinterpretations.

5. CONCLUSION

In the current study, we investigated the development of mental models under uncertainty via eye movements with the newly developed occluded visual spatial search task which seems to be a valid method enabling to study different levels of consciousness. The findings of this study suggest that eye movement patterns inform about the stage of learning and about the degree of participants subjective uncertainty during the acquisition of mental models. However, uncertainty concerning the technical system is not reflected by the behavioural data as participants use heuristic to make their decisions and rather ignore low probabilities. Future research is necessary for a deeper understanding of eye movement patterns in relation to the development of mental models in an uncertain environment. Future studies should also take in account the effect of interindividual differences in cognitive abilities (cf. Jipp, 2016) on the mental model acquisition.

ACKNOWLEDGMENTS

This research was supported by the Research Training Group 1855 “Discrete Optimization of Technical Systems under Uncertainty” of TU Dortmund University funded by the German Research Foundation (DFG). We would like to thank H. Mussmann for technical assistance and M. Behrendt for the test management.

REFERENCES


