Experimental Study toward Estimation of a Learner Mental State from Processes of Solving Multiple Choice Problems Based on Eye Movements

Kazuaki Kojima*, Keiich Muramatsu & Tatsunori Matsui
Faculty of Human Sciences, Waseda University, Japan
*koj@aoni.waseda.jp

Abstract: Recently, the research area of intelligent educational systems has increasingly approached Educational Data Mining, which addressed a variety of learner aspects such as mental states. Although data of eye movements is promising in realizing mental states of learners, we have to empirically understand features of eye movements in advance due to difficulty in handling them. Toward the final goal of automatic estimation of learner mental states based on eyes, this study experimentally described changes of eye movements by confidence of correct answers. In our experiment, participants were asked to answer multiple-choice problems and respond to questionnaires about the problems, with the results indicating that transitions of eyes in initial processes where all choices were scanned differed depending on the levels of confidence.

Keywords: Educational data mining, mental state, confidence, eye movements, multiple-choice problems

Introduction

Recently, the research area of intelligent educational systems has increasingly approached Educational Data Mining (EDM), attempts to develop methods for exploring data from educational settings and adapt those methods to understand learners and their settings [1]. Such EDM studies have addressed a variety of learner aspects. Besides knowledge structures, those aspects include affective states such as confidence or confusion. Development of such EDM methods that automatically and directly estimate learner mental states would allow teachers to know situations of the learners in distant education, or systems to sophisticate their supportive interventions.

One important data generally used in EDM is information of eyes. Data from eyes is considered to be promising in realizing mental processes that can instantly change in a short activity (e.g., solving of a single problem). Accordingly, it is adopted in exploring human cognitive processes or mental states. In fact, some ITS studies have adopted techniques for measurement of eye movements (e.g., [2]). Furthermore, methods to automatically estimate cognitive processes or mental states from eyes have been implemented [3, 5]. Due to unstability in obtaining data from eyes and absence of a general method to analyze it, however, eye data is considerably difficult to handle in any ways. Hence, the automatic estimation basically needs specific models of eye transitions according to structures of tasks. To automatically estimate mental states in problem solving through observation of eyes, we have to understand features of eyes in advance based on data empirically collected.

Toward the final goal of automatic estimation of learner mental states, this study experimentally described changes of problem solving processes by confidence of correct answers as a mental state based on eye movements. To provide data description not depending on task domains, we adopted multiple-choice problems that only needed recall of knowledge. This study, therefore, aimed to produce qualitative descriptions useful in direct and bottom-up analysis of learner behaviors.

1. Method

In this experiment, participants were asked to answer multiple-choice problems and respond to questionnaires about the problems. Eye movements of the participants were recorded while they answered the problems. Confidence in each problem was evaluated based on the questionnaires. We then explored relationships between confidence and features of the eye movements.

1.1 Tasks and Procedures

This experiment used 30 four-choice problems that questioned encyclopedic knowledge such as history or geography. To observe behaviors in reading of texts and searching of answers, the problems that require no mental processing of integration and reasoning were created by the authors.

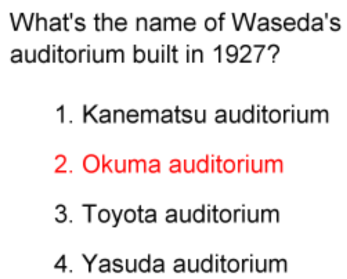


Figure 1. Example of problems presented by the program

Each of the problems was presented on a full screen of a PC monitor by a program implemented by the first author. Figure 1 shows examples of the problems presented. Participants responded each of the problems in the following procedures.

1. Reading a text: A button labeled "*proceed next*" appeared in a position where a text of each problem would be presented. When a participant clicked the button with a mouse, a text of a problem was shown. S/he was instructed to click a button labeled "*I have read the text*" beneath the text as soon as s/he had read it.
2. Responding a pre-questionnaire: The text was once hidden and a dialog window showing the questionnaire that asked to what extent the answer was familiar (*pre-familiarities*) appeared. The participants responded to it by selecting one of "I know its answer, I can answer without choices (*recall*)", "I can remember its answer from choices (*recognition*)", "I don't know, but I may be able to guess from choices (*guesstimate*)", and "I have no idea (*no-idea*)".
3. Selecting an answer: The text and four choices were presented after response of the questionnaire. Each of the choices turned red when the mouse cursor entered. The red choice was selected as the answer when it was clicked.
4. Responding post-questionnaires: The problem was then removed and questionnaires about the problem were presented. The first questionnaire asked whether or not the

participants selected the choice that they had judged to be the answer. The second again asked the familiarity for the answer (*post-familiarities*). In the post- questionnaires, choices in the questionnaire here were altered to “I know the answer, I could answer without the choices (*recall*)”, “I remembered the answer from the choice (*recognition*)”, “I don’t know, but guessed from the choices (*guesstimate*)”, or “I had no idea (*no-idea*)”. The third asked *evaluations* for each choice. Every choice was evaluated with “it’s definitely the answer/ not the answer”, “it’s probably the answer/ not the answer”, or “I cannot judge”.

Each participant in the experiment was seated in front of a desk where a PC monitor (the resolution was 1280 x 1024 pixels) was set up. S/he was asked to answer problems with a mouse as quickly and correctly as possible, and to respond to the questionnaires after answering of each problem. To train for the experimental tasks, s/he answered one problem and its questionnaires prior to answering the 30 problems.

While each participant engaged in selecting answers, eye movements were recorded with EMR-AT VOXER produced by nac Image Technology Inc. The sampling rate of the record was 60 frames/sec. Data in each frame included values of x and y coordinates on the screen.

1.2 Analysis

Data of the 30 problems of each participant was categorized into groups according to the pre- and post-questionnaires. The number of alternatives in finally selecting an answer of each problem was estimated with responses of the evaluations of the third one of the post-questionnaires: answering from *one* (three choices were evaluated not to be the answer), *two*, *three* or *all* (nothing was evaluated not to be the answer). The groups were formed by combining the alternatives, pre-familiarities and post-familiarities.

Data of eye movements were used to analyze transitions of positions where participants watched. Fixations of eye movements were computed from the data. In this study, a fixation was defined as a sequence of six or more serial frames (0.1 sec or more) whose range was smaller than a circle of a 22-pixel radius (the angle was 0.8 degree or smaller) according to the experimental setting. Targets that the participants watched and their transitions were estimated with vertical positions of fixations. Prior to the analysis, we removed frames of blinks from the data, and then corrected and smoothed it.

In answering a single problem, first of all, each participant must have read its text and all four choices in turns. We defined this initial process where the all choices were viewed as *initial scanning*, and extracted transition patterns of fixations in it.

2. Results

Ten undergraduate students participated in the experiment. Four of them were excluded from the analysis because they inappropriately responded to the questionnaires due to misunderstanding of instructions.

2.1 Groups of Problems

Data of 180 problems was obtained from the six participants. In the aspect of the alternatives, answers were selected from one alternative choice in 58.1% of the problems, from two in 16.8%, from three in 3.9% and from all in 21.2%. Hereafter, data[1] of

[1] The label “one/recall-recall” indicates a group of problems whose alternative was one, whose pre-familiarity was recall, and whose post-familiarity was recall.

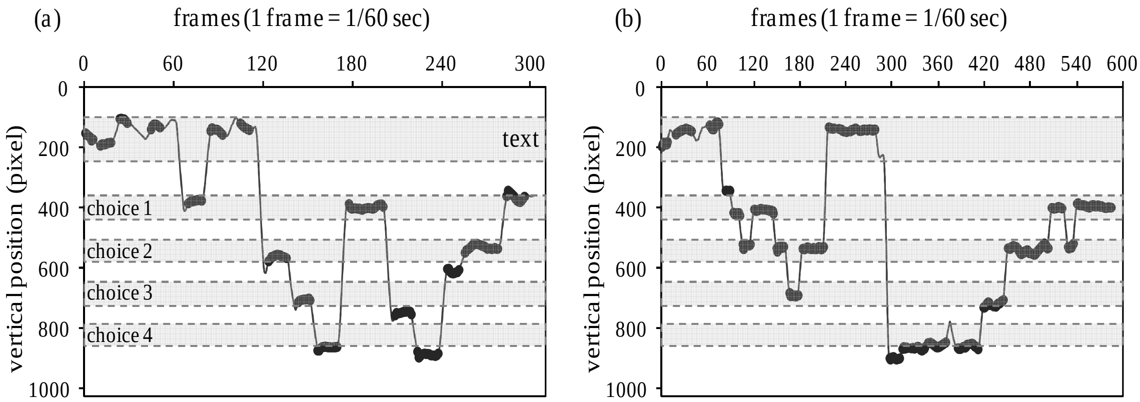
one/recall-recall, one/recognition-recognition, one/guesstimate-guesstimate, all/guesstimate-noidea and all/noidea-noidea was analyzed because data frequencies of the other groups were few.

It was assumed that the degrees of confidence of correct answers got higher in the order of the groups described above. Actually, the proportions of correct answers basically got higher in the order (97.3% in one/recall-recall, 100% in one/recognition-recognition, 46.7% in one/guesstimate-guesstimate, 27.3% in all/guesstimate-noidea and 20% in all/noidea-noidea).

2.2 Transactions of fixations

Figure 3 presents examples of eye movements obtained from the participants. Transitions of targets that the participants watched were reproduced with vertical positions of fixations. In the case (a) of the figure, the participant first read a problem text for about 1 second. He again saw the text after viewing choice 1, and then viewed choices 2, 3 and 4 in turns. He finally selected choice 1 as the answer after examining the four choices. Transitions in initial scanning in this case were [t (text) > 1 > t > 2 > 3 > 4]. Transition patterns in initial scanning were categorized into the following three.

1. *Simple sequential scanning (simple)*: Four choices were simply scanned in turns in initial scanning.
2. *Sequential scanning including returning (returning)*: Like as the case mentioned above, returning to a text was included during scanning of four choices in turns.
3. *The others*: Transitions of choices were not orderly but varied. An example of this is (b) in Figure 3 (The transitions were [t > 1 > 2 > 1 > 2 > 3 > 2 > t > 4]).



(Lines indicate a vertical position of eyes in each frame, and bolded lines that of fixation)
Figure 3. Examples of eye movements

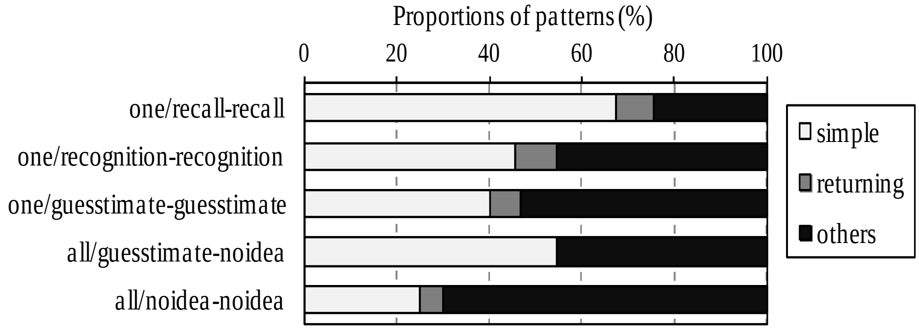


Figure 4. Proportions of transition patterns

Figure 4 indicates the proportions of each transition pattern in each group. We examined differences of the numbers of the patterns among the groups by the chi-square test, however, the differences were not significant ($\chi^2(8)=12.93$, *n.s.*). We examined differences between one/recall-recall and all/noidea-noidea, with the results indicating that the differences were significant ($\chi^2(2)=11.36$, $p<.01$). The residual analysis indicated that the number of simple was high and that of others were low ($p<.01$) in recall/one, and the number of simple was low and that of others was high ($p<.01$) in no-idea/three.

3. Discussion and Future Work

It was assumed that confidence of correct answers was the highest in working on the problems in one/recall-recall. In this case, most transition patterns of eyes in initial scanning were simple. This indicates that the participants scanned choices in turns when confidence was high. On the other hand, transitions of eyes varied or were disordered in all/noidea-noidea, where confidence was the lowest. Thus, analysis of initial scanning must be useful in directly estimating confidence in an earlier stage of problem solving.

Our experiment revealed that the transitions were differed depending on confidence. This fact must indicate that purposes of viewing choices were changed along with confidence. When confidence was high in answering to a problem, viewing choices was considered to be an action to search an answer a participant had already had, whereas it was an action to understand a problem or consider its answer when confidence was low. In many cases of low confidence, initial scanning included transitions to return to texts from choices. Eye transactions were disordered because of comparison between a text and a choice, or between choices. In text reading, it has been documented that backtracking of eye movements occurs when the reader does not understand the text [4].The backtracking in text reading is similar to the disordered eye transactions observed in the current study. Therefore, it is considered that initial scanning served as problem understanding when confidence was low. One important future work is further collection of empirical data to study more diverse cases, such as problems answered in middle confidence. Of course, it is also important task to implement computational models to estimate confidence based on transaction patterns of things that learners watch. Based on our finding, such a model may be implemented with simple techniques because it has only to distinguish a normative and orderly process in initial scanning.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 22300294.

References

[1] Baker, R. S. J. D. and Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1, 3-17.

[2] Conati, C. and Merten, C. (2007). Eye-Tracking for User Modeling in Exploratory Learning Environments: An Empirical Evaluation. *Knowledge-Based Systems*, 20, 557-574.

[3] Nakayama, M. & Takahasi, Y. (2008). Estimation of Certainty for Multiple Choice Tasks Using Features of Eye-movements, *Proceedings of Symposium on Eye Tracking Research & Applications* (pp. 73-76).

[4] Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124, 372-422.

[5] Salvucci, D. D. (2001). An Integrated Model of Eye Movements and Visual Encoding, *Journal of Cognitive Systems Research*, 1(4), 201-220.