Attentional Characteristics of Anomaly Detection in Conceptual Modeling

Karl-David Boutin, Pierre-Majorique Léger, Christopher J. Davis, Alan R. Hevner and Élise Labonté-LeMoyne

Abstract We use eye tracking to better understand the attentional characteristics specific to successful error detection in conceptual models. This phase of our multi-step research project describes the visual comportments associated with successful semantic and syntactic error identification and diagnosis. We test our predictions, based on prior studies on visual attention in an error detection task, or studies comparing experts and non-experts in diverse tasks, in a controlled experiment where participants are tasked with detecting and diagnosing errors in 75 BPMN® models. The results suggest that successful error diagnostics are linked with shorter total view time and shorter fixation duration, with a significant difference between semantic and syntactic errors. By identifying the visual attention differences and tendencies associated with successful detection tasks and the investigation of semantic and syntactic errors, we highlight the non-polarity of the 'scale' of expertise and allow clear recommendations for curriculum development and training methods.

Keywords Eye tracking · Conceptual modeling · Attentional characteristics

K.-D. Boutin (✉) · P.-M. Léger · É. Labonté-LeMoyne
HEC Montréal, Montréal, Québec, Canada
e-mail: karl-david.boutin@hec.ca

P.-M. Léger
e-mail: pierre-majorique.leger@hec.ca

É. Labonté-LeMoyne
e-mail: elise.labonte-lemoyne@hec.ca

C. J. Davis
University of South Florida, Saint Petersburg, FL, USA
e-mail: davisc@mail.usf.edu

A. R. Hevner
University of South Florida, Tampa, FL, USA
e-mail: ahevner@usf.edu

F. D. Davis et al. (eds.), *Information Systems and Neuroscience*,
Lecture Notes in Information Systems and Organisation 29,
https://doi.org/10.1007/978-3-030-01087-4_7

1 Introduction

Business process modeling has become a central activity in IS practice [1]. Conceptual models facilitate communication about business domains and their processes [2–4]. Such models have become a primary medium used in design activities. This phase of our research strives to deepen understanding of visual attention during error detection tasks [5]. While researchers have explored the variations between novice and experienced modelers [6], the differences in the visual attention between successful and unsuccessful error detection tasks in conceptual modeling are yet to be deeply investigated.

For this research, we employ the Business Process Modeling Notation (BPMN^{R}), an international standard for business processes. Its popularity in commercial settings prompted its selection for this phase of our work. Visual notations such as BPMN are composed of visual syntax—symbolic vocabulary and grammar—and visual semantics—elements that give meaning to each symbol and symbol relationship [7, 8]. However, while evaluating notations or their usage, the syntactic component is rarely discussed [8]. This presents an opportunity to contribute to the literature by comparing both the semantic and syntactic error identification process. The main objective of this study is to explore the differences in the attentional characteristics between successful and unsuccessful diagnostics in a detection task. We use eye tracking to monitor the visual attention of subjects as they search for and diagnose semantic and syntactic errors in a controlled experiment.

2 Prior Research and Hypotheses Development

Studies regarding the difference in the visual attention in an error detection task conclude that, on average, errors are fixated more often and longer than irrelevant information [9–11], and that a high number of fixations on the stimulus is correlated with an ineffective search [11, 12]. Furthermore, the longer the participant spends looking for an error, the lower the chances of success [9], possibly due to too much cognitive resource being drawn away for the encoding of the stimulus. Studies that compare novices and experienced modelers point to attentional characteristics that might be associated with expertise, and thus, generally, with better performance [1, 13].

Several meta-analyses that use eye tracking to compare experts and novices in a range of domains conclude that those classified as experts spend less time looking at stimuli before fixating relevant areas or anomalies [5, 14–16]. More efficient scan patterns [13–15] or more detailed and completed schemata [17, 18] are offered as explanations. Experts also tend to have fewer fixations, suggesting less cognitive effort to decipher and understand the stimuli [13, 15], and shorter fixation durations [5], which are also associated with lower cognitive processing effort [11]. Therefore, we propose three study hypotheses:

H1 *Successful error detections in conceptual modeling will require less time spent looking at the stimulus than unsuccessful error detections.*

H2 *Successful error detections in conceptual modeling will require, in total, fewer fixations than unsuccessful error detections, but with a higher proportion of fixations on the error.*

H3 *Successful error detections in conceptual modeling will require, on average, shorter fixation duration than unsuccessful error detections, but with longer fixation duration on the error.*

3 Research Method

Our experiment was conducted on a sample of 18 participants (7 males, 11 females) with different ages and experience. All our participants were offered a $20 gift card as a compensation for their participation. The research was approved by our institution's Research Ethics Board (REB), and each participant signed a consent form.

3.1 Research Design and Protocol

In order to confirm our hypotheses, we tasked our participants with identifying and diagnosing errors in conceptual models written in BPMN. Each participant had to inspect five (5) distinct sets of 15 models (for a total of 75 models), where each set represented a business process scenario (e.g. airport check-in process). An example can be seen in Fig. 1. Five (5) versions of each scenario were presented as BPMN 'sentences'. These were further manipulated to present three (3) versions: one with no known errors; one with a known semantic error, and one with a known syntactic error [7].

To mitigate the effect of prior knowledge of the business domains of the models [19, 20], the stimuli were created using simple scenarios, well-known to all potential participants. Furthermore, we limited the range of symbols used and the scope of the 'sentences' to 10–12 elements, favoring numerous accessible models rather than more complex and domain knowledge-dependent stimuli. To train the participants and mitigate the effect of learning through the trial, the experiment started with a short presentation on BPMN [21]. The symbols used in the study, as well as the two different types of errors, were shown and explained. The training was concluded with a practice task where participants were shown three (3) models, each one with a different type of error. Just like in the real task, the participant had to pinpoint the error and to diagnose the error type, both by clicking in the corresponding area on the stimuli. The correct location of the error, as well as the right error type, was revealed after each practice model. To avoid bias, the models used in the practice task were not related to the sets of models used later in the experiment, and the conditions were the same as in the experiment [21].

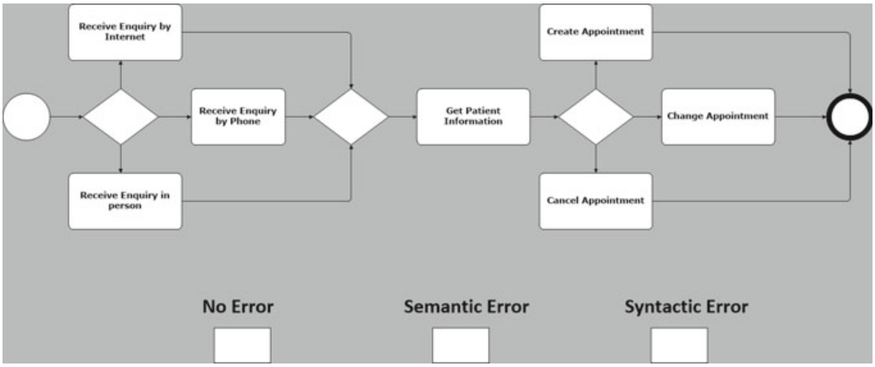

Fig. 1 Example of a model with boxes to indicate type of error detected

The participants then started the first task with their first set of models. The fifteen models included in the set were shown in random order and without any time limit. After identifying and diagnosing the error in a model, participants had to manually advance to the next model, using the space bar on their keyboard. They then proceed to the next set of models until they completed the five (5) sets.

3.2 Apparatus and Measures

Eye tracking (Red 250, SensoMotoric Instruments GmbH, Teltow, Germany) was used to gather the behavioral measures throughout the experiment, at a sampling frequency of 60 Hz. The number of fixations, which is the stabilization of the eye on an object [13], and their duration were gathered for each area of interest (AOI), as the literature tends to agree that fixation is related to the cognitive processing of visual information [13, 22]. The fixation duration threshold was set at 200 ms [11, 23]. The time before the first fixation in an AOI and the total view time of a stimulus were also collected. One (1) to three (3) AOIs were placed on the correct choice of error type and on the actual location of the error(s). For each participant, the eye tracker was calibrated to a maximum average deviation of 0.5°, using a 9-points predefined calibration grid.

4 Preliminary Results

We briefly present several preliminary results from our study. Hypothesis 1 states that successful identification and diagnosis of errors in conceptual models will take less time than unsuccessful answers. A linear regression with mixed model and a two-tailed level of significance is performed to compare the Total View Time for each value of the variable Answer (i.e. if the participant successfully diagnosed the error, Answer = 1, if not, Answer = 0). Results suggest that successful detection of error, including models without an error, is linked with lower total time spent on each stimulus ($B = -0.3934$, $p < .0001$). Furthermore, successful detection of semantic errors shows a faster time to first fixation on the area of interest (i.e. the zone containing the error) ($B = -0.3333$, $p = .0027$). However, there are no significant results linking the detection of syntactic or no errors with the time to first fixation.

Hypothesis 2, which stipulates that successful error detection will be linked with fewer fixations, but with a higher proportion of fixations on the error, is tested using a Poisson regression with mixed model and a two-tailed level of significance of Fixation Count on Answer. A significant relation is found between successfully detecting an error in a model and lower fixation count ($B = -0.4402$, $p < .0001$). Moreover, greater proportions of fixation are allocated to the zone containing semantic errors ($B = 0.5448$, $p < .0001$) and syntactic errors ($B = 0.9379$, $p < .0001$). However, while correct diagnosis of semantic errors are linked with a decrease in the number of fixation in the areas of interest ($B = -0.04956$, $p = .0897$), the opposite is found for the successful detection of syntactic errors, where more fixations on the AOIs are required ($B = 0.3654$, $p < .0001$).

In order to test Hypothesis 3, we apply a linear regression with mixed model and a two-tailed level of significance of Fixation Durations on Answer, allowing us to identify a significant correlation between successful diagnostics and shorter fixation duration ($B = -0.373$, $p < .0001$). However, fixations in the AOIs are longer for successful diagnosed semantic errors ($B = 0.1654$, $p = .0061$) and syntactic errors ($B = 0.4436$, $p < .0001$), which indicate that the participants, when successfully identifying the errors, spend more time on the errors, but less time on the rest of the stimuli.

5 Discussion and Conclusion

Our preliminary results suggest that H1, H2 and H3 are partially supported. Significant links between successfully detecting an error and a lesser amount of time spent on a stimulus (H1), and between an accurate diagnostic and shorter fixation duration are found (H3). Furthermore, H2, which proposed fewer fixations and a greater proportion of fixations on the errors when successfully detecting an error, is supported. However, successful detection of syntactic errors is significantly associated with a greater number of fixations in AOIs, which suggests that the error was detected, but the correction response is inhibited [9], possibly due to a higher level of complexity in syntactic errors. No significant link between syntactic errors and the number of fixations in the entire stimulus is found. Thus, this study presents evidence that the characteristics of visual attention of experienced modelers, such as lower number and duration of fixations, are generally related with successful error detection. On the other hand, attributes normally associated with novices, such as higher time spent on a stimulus or higher fixation duration, lead to unsuccessful error detection.

Our research so far offers both theoretical and practical contributions. The differences between the process and repertoire of attentional characteristics in the detection of semantic errors and syntactic errors reinforce Moody's [8] propositions about their complex interdependence. Syntactic errors require more attentional fixation than semantic errors. This finding runs contrary to our expectations and highlights the need for further studies to more fully articulate the differences and the metrics that might be used to measure them. The next phase of our research will address this challenge. From a practical standpoint, deeper insights into differences between the attentional characteristics will offer guidance to the evolution of BPMN and other notations and recommendations for curriculum development and training methods.

Limitations of this exploratory phase of our study center on the models used as stimuli for our experiment. While we tried to minimize the impact of domain-specific knowledge by using processes well-known to all potential participants, it is virtually impossible to negate the influence of the variation of domain familiarity between participants. Another limitation can be found in our sample. A bigger and more equally distributed sample will allow more complex statistical analyses and provide more significant findings. Finally, no measure was taken to evaluate the 'stopping rule', which is when a subject decides to terminate his information search because he judges that he has enough information to complete his task [24, 25]. The next step of our research should evaluate this concept in order to better understand our eye tracking data, especially the measures linked to the view time.

References

1. Recker, J.C., Dreiling, A.: Does it matter which process modelling language we teach or use? An experimental study on understanding process modelling languages without formal education. In: Toleman, M., Cater-Steel, A., Roberts, D. (eds.) 18th Australasian Conference on Information Systems, Toowoomba, Australia, pp. 356–366 (2007)
2. Parsons, J., Cole, L.: What do the pictures mean? Guidelines for experimental evaluation of representation fidelity in diagrammatical conceptual modeling techniques. Data Knowl. Eng. **55**(3), 327–342 (2005)
3. Gemino, A., Wand, Y.: Comparing mandatory and optional properties in conceptual data modeling. In: Proceedings of the Tenth Annual Workshop on Information Technologies and Systems, Brisbane, pp. 97–102 (2000)
4. Gemino, A., Wand, Y.: Evaluating modeling techniques based on models of learning. Commun. ACM **46**(10), 79–84 (2003)
5. Gegenfurtner, A., Lehtinen, E., Säljö, R.: Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. Educ. Psychol. Rev. **23**(4), 523–552 (2011)
6. Shanks, G.: Conceptual data modelling: an empirical study of expert and novice data modellers. Australas. J. Inf. Syst. **4**(2) (1997)
7. Davis, C.J., Hevner, A.R., Labonte-LeMoyne, É., Léger, P.-M.: Expertise as a mediating factor in conceptual modeling. In: Information Systems and Neuroscience, pp. 85–92. Springer, Berlin (2018)
8. Moody, D.: The "physics" of notations: toward a scientific basis for constructing visual notations in software engineering. IEEE Trans. Softw. Eng. **35**(6), 756–779 (2009)

9. Van Waes, L., Leijten, M., Quinlan, T.: Reading during sentence composing and error correction: a multilevel analysis of the influences of task complexity. Read. Writ. **23**(7), 803–834 (2009)
10. Henderson, J.M., Hollingworth, A.: The role of fixation position in detecting scene changes across saccades. Psychol. Sci. **10**(5), 438–443 (1999)
11. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J.: Eye Tracking: A Comprehensive Guide to Methods and Measures. OUP, Oxford (2011)
12. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. **24**(6), 631–645 (1999)
13. Yusuf, S., Kagdi, H., Maletic, J.I.: Assessing the comprehension of UML class diagrams via eye tracking. In: Program Comprehension, 2007. ICPC'07, 15th IEEE International Conference, pp. 113–122 (2007)
14. Krupinski, E.A.: The importance of perception research in medical imaging. Radiat. Med. **18**(6), 329–334 (2000)
15. Reingold, E.M., Sheridan, H.: Eye movements and visual expertise in chess and medicine. Oxf. Handb. Eye Movements, 528–550 (2011)
16. Sheridan, H., Reingold, E.M.: Expert vs. novice differences in the detection of relevant information during a chess game: evidence from eye movements. Front Psychol. **5**, 941 (2014)
17. Glaser, R.: Education and thinking: the role of knowledge. Am. Psychol. **39**(2), 93 (1984)
18. Lurigio, A.J., Carroll, J.S.: Probation officers' schemata of offenders: content, development, and impact on treatment decisions. J. Pers. Soc. Psychol. **48**(5), 1112 (1985)
19. Gemino, A., Wand, Y.: A framework for empirical evaluation of conceptual modeling techniques. Requirements Eng. **9**(4), 248–260 (2004)
20. Birkmeier, D., Kloeckner, S., Overhage, S.: An empirical comparison of the usability of BPMN and UML activity diagrams for business users. In: ECIS, p. 2 (2010)
21. Bavota, G., Gravino, C., Oliveto, R., De Lucia, A., Tortora, G., Genero, M., Cruz-Lemus, J.A.: Identifying the weaknesses of UML class diagrams during data model comprehension. In: International Conference on Model Driven Engineering Languages and Systems, pp. 168–182. Springer, Berlin (2011)
22. Just, M.A., Carpenter, P.A.: Eye fixations and cognitive processes. Cogn. Psychol. **8**(4), 441–480 (1976)
23. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychol. Bull. **124**(3), 372 (1998)
24. Browne, G.J., Pitts, M.G.: Stopping rule use during information search in design problems. Organ. Behav. Hum. Decis. Process. **95**(2), 208–224 (2004)
25. Nickles, K.R., Curley, S.P., Benson, P.G.: Judgment-Based and Reasoning-Based Stopping Rules in Decision Making Under Uncertainty, vol. 7285. University of Minnesota (1995)