Real-Time Business Process Model Tailoring: The Effect of Domain Knowledge on Reading Strategy

Sven Vermeulen(✉)

Department of Business Informatics and Operations Management, Ghent University, Ghent, Belgium
Sven.Vermeulen@UGent.be

Abstract. Due to the use of thousands of often very complex process models, having them immediately usable towards their purpose is of great economic benefit. In order to maximize usability, process models must be intuitive and easily understandable. In other words, processing the information contained within the process models must enable a successful completion of the task for which the model is being used. Recently, research efforts into the effects of user characteristics on understandability have increased. However, current limitations create promising research possibilities, particularly with regard to the use of realistic process models and direct data collection techniques. This thesis will contribute to the existing body of knowledge by investigating domain knowledge as a fundamental user characteristic and utilizing eye-tracking as a direct data collection method while using realistic, complex process models. The end goal of this research is to propose an automatic process model tailoring technique, with the aim of enhancing a user's understanding and thus their performance. As of now a pilot study has indicated the existence of distinct reading strategies, which establishes the viability of the proposed future work.

Keywords: Business process modeling · Process model understandability
Process model visualization · Domain knowledge · Process model tailoring
Reading strategy

1 Introduction

The performance of a business is in large part achieved by modeling the business processes on which the enterprise relies [1]. Due to their communicative and supportive function for knowledge transfer, quality control and regulation [2], they are required to be intuitive and easily understandable [3]. Process model understandability is a field investigating the factors that influence this comprehension. Uncovering these effects facilitates a common understanding of processes and helps improve the quality of models [1]. Recently, studies on the cognitive aspects of process models have increased, establishing the importance of the effects of user characteristics (e.g. [4, 5]). We follow the reasoning that the way an individual receives, selects, organizes and uses the information visualized in a process model is an example of intrapersonal information behavior [6].

C. Debruyne et al. (Eds.): OTM 2017 Workshops, LNCS 10697, pp. 280–286, 2018.
https://doi.org/10.1007/978-3-319-73805-5_30

Process models are used and must therefore be understood by a variety of people whose user characters' can have an even bigger influence compared to model factors [7].

In our work we will contribute to current research in three studies. First, we will study the effects of domain knowledge on reading strategy and process model understandability. Efforts into uncovering how users read models are very limited and recent although it has been established that looking at model elements is highly correlated with the individual's thinking process and performance [8]. Our work is innovative as it improves upon existing research by investigating the reading process, using a realistically complex model and applying eye-tracking as a direct data collection technique. Given the stadium at which our research currently stands, the metrics' thresholds (e.g. CFC) are yet to be determined, based on which a model will be deemed "complex".

In our second study, we intend to investigate and apply these factors to the idea of tailoring process models [9]. The static way in which processes are currently visualized is a major shortcoming as it prevents individual and customized perspectives on business processes [1, 10], hampering performance. Based on the results of this second study, we will work towards a fully automatic approach that will tailor process models to specific user characteristics as a third contribution.

2 Related Work

2.1 Process Model Reading

In the domain of process model reading, Figl and Strembeck [11] focus on modeling direction and the effect on process model comprehension. The authors remarked a lack of complexity in the models used in the experiment. We intend to improve upon this by using a realistic, complex model. We further argue that since the experiment did not contain a direct data gathering method with regard to reading a model, such as eye-tracking, the deeper cognitive principles remained unknown. Specifically aimed at model inspection for quality issues, Haisjackl et al. [12] investigated the strategies undertaken by model readers in search of errors. In this study, data collection was done using a think-aloud protocol for which we argue the same limitations apply.

2.2 Investigating Domain Knowledge

Only four studies could be identified investigating domain knowledge in the context of process model understandability. These findings are supported by a very recent literature review [1]. None of these studies reported an effect of domain knowledge on model comprehension. Either researchers have kept domain knowledge constant (as a control variable) (e.g. [5, 13]), have used homogenous groups of participants [14] or domain knowledge was measured using a self-reported scale [15]. The key problem is that when the extent of knowledge in a domain is itself unknown, no basis exists for a person to estimate what they don't know. In other words, an individual is unable to assess their own level of domain knowledge as their understanding of the domain affects their understanding of their own knowledge level [16].

For all studies mentioned in this chapter, we argue that the data collection method does not allow a direct observation and registration of the reading strategy and underlying cognitive processes. Furthermore, as explained by Figl et al., domain knowledge has as of yet not been a factor under study in the process modeling domain even though relations with performance are well established in other fields [17]. Therefore, utilizing eye-tracking as a data collection method would be our fourth contribution, next to our study of domain knowledge, reading strategy and the use of realistic models.

2.3 Process Model Tailoring

Configuring process models so they only provide relevant information for a particular user is not a new idea (e.g. [18]). Within this field of study, however, the effects of different configurations of the process on a user's understanding have received less attention. More specifically, in line with our work, the idea of tailoring the model to personal factors [9] has not yet been empirically assessed. In [1], potential future work on tailoring process models is limited towards variables for which interaction effects with various forms of representation have been found. Given the limitations of research regarding the use of domain knowledge as a factor, we aim to contribute to both areas (i.e. determining the effect of domain knowledge on comprehension and tailoring process models to personal factors such as learning style or personality).

3 Research Questions and Methodology

Our end goal is to develop a technique for dynamically visualizing process models based on user characteristics. Specifically, we intend to personalize these models based on domain knowledge and reading strategy. To this end, we first need to establish an understanding of the effects of domain knowledge on process model reading and understandability. Experts and novices acquire and integrate information in different ways [19]. The extent of tacit knowledge – general and domain specific skills acquired over time – separates experts from novices [20]. As experience and ability form an individual's internal state of knowledge, they determine performance.

A preliminary experiment will determine the presence of different reading strategies and their potential relation with domain knowledge [21]. A process model reading strategy can be defined as "the set and sequence of reading actions performed to translate the model into meaning". Eye-tracking will be used as a direct data collection method. It uncovers the mental processes underlying task behavior as eye movement paths, eye fixations and pupil dilations reflect the internal processing of information [22]. The actions that will be recorded are: *fixations,* where the eye stops for a brief moment on a model element; *saccades*, where the eye is quickly positioned on a new model element, *regression,* where the eye returns to an already visited element and *perceptual span*, which refers to the size of the visual window processed at each fixation. These actions will be analyzed individually (e.g. total number and average duration of fixations). Furthermore, recording the sequence of the elements themselves as read by each participant, a reading strategy will be interpretable as a sentence, on which text reading analysis techniques can be applied. Research that employs eye-tracking has high external validity as it is a non-invasive process tracing technique. This is also supported by [1] who consider the use of eye-tracking as the one neurophysiological tool that has been successfully applied in the area of process model comprehension.

Two process models will be used in the form of BPMN-diagrams, each a realistic representation of a process in a distinct domain: one on complaint handling and one on the treatment of stroke in healthcare. Both models were created in collaboration with employees active in the process. A set of domain related questions will be presented to each participant determining their level of domain knowledge. This set will be the result of discussions with longtime process participants, who, given their experience with the domain, will serve as proxies for domain experts. This approximation is necessary because a user can only be deemed an expert by other experts, for which we would also need the set of questions, causing a vicious cycle to start. The set will be deemed complete if no further questions are provided by new proxies. By creating a set that is as representative and long as possible, the risk of choosing a sample which includes "a preponderance of questions to which the subject happens to know the answers" [16] is avoided.

Participants will be domain experts in exactly one of the domains depicted by one of the models. By that logic, participants can be treated as novices for the model depicting the domain they are not an expert in. In order to determine reading strategies and differences between domain knowledge, each participant will receive both models. Whichever model they see first will be randomized to avoid bias. We will collect data on a sample of at least 70 participants (i.e. around 35 per domain) to maximize significance (following [23]). These are still to be determined, but will most likely be employees currently active in the used processes.

We focus on comprehension and problem solving measured by questions on semantics and pragmatics. This study does not aim to investigate differences in modeling knowledge and thus has no need to include syntax related questions. Performance will be determined by the results on these tests. We except domain experts to outperform novices. Therefore we hypothesize:

H1a: Domain experts will have higher performance on comprehension than novices.
H1b: Domain experts will spend less time reading the process model than novices.

Experts will rely on their prior knowledge to build an understanding of the model. Therefore we expect them to require less time, resulting in lower fixation rates and longer saccades. This is supported by the information-reduction hypothesis, which proposes that expertise optimizes the amount of processed information by neglecting task-irrelevant information and actively focusing on task-relevant information [24]. This effect is also expected to result in a different reading strategy and performance.

H2: Domain experts will exhibit a different reading strategy from novices.
H3: A correlation can be found between reading strategy and performance.

Participants will be required to read the entire model without a questionnaire. We argue that the specificity of the reading process is too much dependent on question choice [8] and therefore does not fully represent the cognitive processes that would otherwise unfold. The majority of research on process model understandability provides a questionnaire adapted to the factors under study. Consequently, those model elements not included require separate research and might fragment the body of knowledge. Since there will be no question list provided, participants will likely adopt a certain reading strategy to process the model in a way that best fits their learning style without being directed through the model as they are given the liberty to explore and process the model as they chose. This is in line with research on reading strategies of data models where participants were "simply asked to give a complete specification of each data model" [25].

The second stage starts after participants are convinced of understanding the model. They are only then provided with the aforementioned questionnaire. To avoid memorization in stage one, the model will be kept in view while solving questions. That way, participants will not feel the need to memorize for short-term recollection, but will instead use this first reading as an initial step of a learning process. Again, eye-tracking will be used to collect data in terms of fixation rates, saccades and reading strategies. In line with research on data models [25] and process models [12] we expect to be able to identify distinct reading strategies in both stages and differences in performance based on those strategies.

In the following phase, based on the findings of the experiment, we will develop a solution in the form of a collection of propositions towards personalized visualizations of process models. These will include guidelines and patterns to optimize process models with regards to a user's domain knowledge and reading strategy in order to enhance understandability. We aim to study the results of certain model manipulations on understandability for different users and reading strategies to conclude on a set of guidelines for each. Further in the thesis, we will work towards an automatic approach that will adapt models in real-time based on a user's domain knowledge, reading strategy and the corresponding set of modeling guidelines. To determine the success of a process configuration, a user's performance will be monitored with the use of intermediate questions.

4 Exploratory Study Preliminary Results

A pilot study was conducted with the single goal of determining the existence of distinct reading strategies. In case none of the participants exhibited dissimilar reading behaviour on the same process model understanding task, there would have been need to adjust our hypotheses and/or research track of future experiments as these rely on this initial assumption. A small set of users (10) was invited by the researcher to participate. These were chosen to ensure adequate variance in the user characteristics (i.e. modelling expert and modelling novice, abstract learner and structured learner,..). Using the SMI Red250Mobile eye-tracker (binocular mode at 250 Hz sampling rate with average operating distance of 65 cm) data was collected on reading sequence and characteristics (e.g. saccades, fixations, regressions, dwell time,..). Both types of data were used to define a user's reading strategy. At the point of writing, this data has only been inspected visually by comparing scan paths and fixation sequences side by side on dual monitors. This visual analysis already made evident the existence of differences in reading strategy.

However, we plan on continuing the analysis by converting fixation sequence data into strings on which string similarity measures (e.g. Jaccard index,..) will be applied.

5 Conclusion

Within the large body of knowledge on process model understandability, the growing importance of the effects of user characteristics is evident [1], establishing the need for further study. As an answer to current research opportunities, this thesis will contribute by investigating the effects of domain knowledge on reading strategy and process model understandability. This work will provide an innovative addition to research on four aspects: (1) using eye-tracking as a direct data gathering method in order to uncover participants' cognitive processes; (2) using realistically, complex process models so as to recreate as much of a realistic setting as possible; (3) investigating domain knowledge as the independent variable and (4) studying differences in reading strategy. These findings will ultimately lead to an automatic tailoring solution, which would personalize process models depending on the user, thus enhancing their performance.

Acknowledgements. This Ph.D. project will be funded by Ghent University (Belgium) under the supervision of Prof. Dr. Manu De Backer (administrative promoter) and Prof. Dr. Amy Van Looy (daily supervisor).

References

1. Figl, K.: Comprehension of procedural visual business process models. Bus. Inf. Syst. Eng. **59**, 41–67 (2017). https://doi.org/10.1007/s12599-016-0460-2
2. de Oca, I.M.M., Snoeck, M., Reijers, H.A., Rodríguez-Morffi, A.: A systematic literature review of studies on business process modeling quality. Inf. Softw. Technol. **58**, 187–205 (2015). https://doi.org/10.1016/j.infsof.2014.07.011
3. Dehnert, J., Van Der Aalst, W.M.P.: Bridging the gap between business models and workflow specifications. Int. J. Coop. Inf. Syst. **13**, 289–332 (2004)
4. Mendling, J., Strembeck, M., Recker, J.: Factors of process model comprehension-Findings from a series of experiments. Decis. Support Syst. **53**, 195–206 (2012). https://doi.org/10.1016/j.dss.2011.12.013
5. Turetken, O., Rompen, T., Vanderfeesten, I., Dikici, A., van Moll, J.: The effect of modularity representation and presentation medium on the understandability of business process models in BPMN. In: La Rosa, M., Loos, P., Pastor, O. (eds.) BPM 2016. LNCS, vol. 9850, pp. 289–307. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45348-4_17
6. Heinrich, L.J., Riedl, R., Stelzer, D., Sikora, H.: Informationsmanagement: Grundlagen, Aufgaben, Methoden. De Gruyter, München (2014)
7. Reijers, H.A., Mendling, J.: A study into the factors that influence the understandability of business process models. IEEE Trans. Syst. Man Cybern. **41**, 449–462 (2011). https://doi.org/10.1109/TSMCA.2010.2087017
8. Petrusel, R., Mendling, J.: Eye-tracking the factors of process model comprehension tasks. In: Salinesi, C., Norrie, M.C., Pastor, Ó. (eds.) CAiSE 2013. LNCS, vol. 7908, pp. 224–239. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38709-8_15

9. Aysolmaz, B., Reijers, H.A.: Towards an integrated framework for invigorating process models: a research agenda. In: Reichert, M., Reijers, H.A. (eds.) BPM 2015. LNBIP, vol. 256, pp. 552–558. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42887-1_44
10. Kolb, J.: Abstraction, Visualization, and Evolution of Process Models. Ulm University (2015)
11. Figl, K., Strembeck, M.: On the importance of flow direction in business process models. In: Proceedings of the 9th International Conference on Software Engineering and Applications (ICSOFT-EA 2014), pp. 132–136 (2014)
12. Haisjackl, C., Soffer, P., Lim, S.Y., Weber, B.: How do humans inspect BPMN models: an exploratory study. Softw. Syst. Model 1–19 (2016). https://doi.org/10.1007/s10270-016-0563-8
13. Bera, P.: Does cognitive overload matter in understanding BPMN models? J. Comput. Inf. Syst. **52**, 59–69 (2012)
14. Recker, J., Dreiling, A.: Does it matter which process modelling language we teach or use? An experimental study on understanding process modelling languages without formal education, pp. 356–366 (2007)
15. Recker, J., Reijers, H.A., van de Wouw, S.G.: Process model comprehension: the effects of cognitive abilities, learning style, and strategy. Commun. Assoc. Inf. Syst. **34**, 199–222 (2014)
16. Borgatti, S.P., Carboni, I.: On measuring individual knowledge in organizations. Organ Res. Methods **10**, 449–462 (2007)
17. Hambrick, D.Z., Engle, R.W.: Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. Cogn. Psychol. **44**, 339–387 (2002)
18. Santos, E., Pimentel, J., Castro, J., Finkelstein, A.: On the dynamic configuration of business process models. In: Bider, I., Halpin, T., Krogstie, J., Nurcan, S., Proper, E., Schmidt, R., Soffer, P., Wrycza, S. (eds.) BPMDS/EMMSAD -2012. LNBIP, vol. 113, pp. 331–346. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31072-0_23
19. Libby, R., Luft, J.: Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. Acc. Organ. Soc. **18**, 425–450 (1993)
20. Schmidt, F.L., Hunter, J.E.: Tacit knowledge, practical intelligence, general mental ability, and job knowledge. Curr. Dir. Psychol. Sci. **2**, 8–9 (1993)
21. Patig, S.: A practical guide to testing the understandability of notations. In: Proceedings of the Fifth Asia-Pacific Conference on Conceptual Modelling, vol. 79, pp. 49–58. Australian Computer Society, Inc. (2008)
22. Weber, B., Pinggera, J., Neurauter, M., et al.: Fixation patterns during process model creation: Initial steps toward neuro-adaptive process modeling environments. In: 49th Hawaii International Conference on System Sciences, pp. 600–609 (2016)
23. Petrusel, R., Mendling, J., Reijers, H.A.: How visual cognition influences process model comprehension. Decis. Support Syst. **96**, 1–16 (2017). https://doi.org/10.1016/j.dss.2017.01.005
24. Haider, H., Frensch, P.A.: Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. J. Exp. Psychol. Learn. Mem. Cogn. **25**, 172 (1999)
25. Nordbotten, J.C., Crosby, M.E.: The effect of graphic style on data model interpretation. Inf. Syst. J. **9**, 139–155 (1999). https://doi.org/10.1046/j.1365-2575.1999.00052.x