Cognitive Elements in Exploratory Process Analysis: From Novice to Expert

Gythe Stans[1,2,*]

[1]UHasselt — Hasselt University, Faculty of Business Economics, Agoralaan, 3590 Diepenbeek, Belgium
[2]UHasselt — Hasselt University, Digital Future Lab, Agoralaan, 3590 Diepenbeek, Belgium

Abstract

While process mining research has primarily focused on algorithms, the cognitive processes of analysts are less explored. This doctoral research aims to identify key cognitive elements involved, compare novices to experts, and explore interventions to help novices develop expert skills. The research is structured around three key questions: (1) how can the cognitive components involved in exploratory process analysis be mapped; (2) how do these cognitive elements differ between novices and experts; and (3) which interventions effectively support novices in developing expert-like cognitive elements. We aim to develop a mapping tool grounded in Design Science Research, and to apply it to capture and compare cognitive patterns across expertise levels. Ultimately, we aim to design targeted interventions to optimize the cognitive development of process analysts.

Keywords

Exploratory Process Mining, Cognitive Elements, Expertise, Personalized Training System

1. Introduction

Since the pioneering work of van der Aalst [1, 2, 3, 4], process mining has seen important growth and received significant recognition. Much of the domain's research has primarily focused on the development of tools and algorithms [5, 6], driving advancements in automated process discovery [7], conformance checking [8], and enhancement techniques [9]. Relative to this extensive research on the algorithmic aspect of the field, little attention has been paid to the individual within the process.

This has given rise to recent studies on the Process of Process Mining (PPM), which emphasize the importance of the analyst's role [10]. Within this emerging field, the focus often lies on the behavior of the analyst and the actions the individual performs, with less emphasis on the cognitive processes driving these actions. Sorokina et al. (2023) contribute to the cognitive focus by analyzing and describing the cognitive steps an analyst takes during process mining [10]. They identify four types of cognitive strategies and test several hypotheses to explore the relationship between this strategy and either a higher performance or the expertise of the analyst.

The individualized perspective becomes even more critical in exploratory process analysis (EPA). Despite limited research in this area and the absence of a clear, widely accepted definition, insights from exploratory data analysis (EDA) research could offer a valuable foundation for adaptation. EDA is described as more of an art than an exact science [11], and research has demonstrated that analysts exhibit diverse behaviors and pursue varying goals during exploratory process mining [12]. These findings highlight the importance of the individual analyst, suggesting that cognitive strategies and behaviors play a crucial role in shaping the quality and direction of EPA.

Despite the growing recognition of the analyst's role in process mining, particularly within the PPM domain [10], there remains a gap in understanding and supporting the cognitive processes behind EPA. While recent studies have identified cognitive strategies in the process [10], these insights have not yet been translated into practical interventions or educational frameworks. The literature is mostly limited to cognitive behaviors, without exploring the cognitive elements behind the behavior or how these

BPM 2025: 23rd International Conference on Business Process Management, 31st Aug - 5th Sep 2025, Seville, Spain
*Corresponding author.
gythe.stans@uhasselt.be (G. Stans)
0009-0009-8275-0879 (G. Stans)

can be effectively taught or optimized. This is specifically critical in EPA, where the lack of structure and predefined goals calls for cognitive flexibility. Although the research on the cognitive strategies of analysts suggests that the findings could inform future training approaches [10], there is an absence of how to effectively and efficiently train non-experts to adopt expert-like cognitive elements.

This research project aims to bridge this gap by identifying the cognitive elements involved in EPA by comparing novices and experts. Additionally, the project explores targeted interventions to support the optimization of the cognitive aspect of novice analysts.

The paper is organized as follows: Section 2 reviews related work, Section 3 outlines the research design and work packages, and Section 4 presents initial steps and preliminary results.

2. Related work

As process mining gains traction in academia and industry, research has increasingly shifted from a purely technical to a human-centered perspective. Sorokina et al. (2023) use cognitive science to identify four strategies that analysts use and link them to performance and expertise [10]. Zerbato et al. (2022) explore real-world usage through interviews, revealing diverse strategies and challenges analysts face [13]. Additionally, Zerbato et al. (2023) introduce the ProMiSE project to better understand analyst practices, and to assist novices during analysis [14].

Zimmerman et al. (2022) identify 23 challenges that analysts encounter across different phases of process mining projects and discuss existing methods that may help in addressing them. These challenges make up a broad range, from technical issues to organizational barriers. Based on this study, Zimmerman et al. (2024) validate these challenges by conducting a survey. They explore how analysts are currently handling them in practice, and emphasize the need for better support systems.

Ammann et al. (2025) focus on the cognitive aspect of analysts as well by researching how individuals use process mining in organizations. By looking at the behavior, thoughts, and emotions of the individual, they have identified four types of analysts. They suggest that training and tools for process mining should be tailored to these types of analysts [15].

While most research takes a broad view of process mining, exploratory process mining remains underexplored, despite its importance in early project stages. Zerbato et al. (2021) address this gap by examining how analysts interact with event logs during this phase, revealing diverse strategies and behaviors that underscore its cognitive flexibility and complexity [12]. Zerbato et al. (2023) address the challenge of tracking the ad hoc nature of exploratory analysis, proposing a support system to improve transparency, reproducibility, and data awareness.

Although research on exploratory process analysis (EPA) is limited, the extensive work on exploratory data analysis (EDA) offers valuable insights. Tukey (1977) compares EDA to detective work [16]. The analyst plays a key role in uncovering patterns, anomalies, and insights. Like a detective, an analyst does not only need the right tools, but also a deep understanding. Good (1983) further describes EDA as more art than science [11], reinforcing the importance of the individual in the process.

While, despite these insights, the cognitive dimension in (exploratory) process mining has been underexplored, broader research has explored the cognitive factors influencing performance or expertise. For example, Endsley (2018) explains the concept of situation awareness (SA) and its three hierarchical levels [17]. He mentions several cognitive features that influence the level an individual can reach, such as an individual's prior knowledge and expectations.

Cognitive modeling also emphasizes the role of knowledge. The ACT-R theory identifies two key types: *declarative knowledge* ("knowing what"), which is conscious and can be verbalized, and *procedural knowledge* ("knowing how"), which is unconscious and demonstrated through actions [18, 19, 20].

The ACT-R theory also considers conditions that determine when procedural knowledge is applied. If a condition applies, the corresponding production is executed. This aligns with *conditional knowledge*, introduced by Paris & Lipson (1983), which captures the strategic aspect of knowledge [21]. Conditional knowledge refers to knowing when and why to apply certain actions. Together with declarative and procedural knowledge, it enables individuals to select appropriate actions to achieve specific goals.

For the second type of knowledge, the procedural knowledge, Fitts and Posner (1967) outline three stages: cognitive, associative, and autonomous [22]. In the cognitive stage, actions are deliberate and effortful. In the associative stage, performance becomes smoother, and in the autonomous stage, tasks are performed efficiently with minimal thought.

In summary, recent studies have started exploring the cognitive side of (exploratory) process mining, but we still know little about how these skills develop, or differ between novices and experts. While some work has identified analyst types and strategies, there has been little focus on how to help analysts grow. This project takes a step in that direction by mapping the cognitive elements during analysis, comparing novices and experts, and exploring ways to better support learning. Our goal is to deepen theoretical insights and lay the groundwork for more effective training for future analysts.

3. Research design

3.1. Research questions

The research project focuses on understanding and analyzing the individual's cognitive aspect in an exploratory process analysis (EPA), and facilitating the optimization of the cognitive aspect to gain expertise more effectively and efficiently. The research is divided into three research questions.

- RQ1: How can the cognitive elements of exploratory process analysis be extracted?
- RQ2: How do the cognitive elements from novices differ from those of experts in EPA?
- RQ3: Which interventions effectively and efficiently help to reshape and optimize the cognitive elements of a novice towards those of an expert?

3.2. Course of research

3.2.1. Work package 1: Development of a tool to map the cognitive elements on exploratory process analysis

To find an answer to the first research question 'How can the cognitive elements of exploratory process mining be extracted?', two sub-questions are defined. First, "*What is exploratory process analysis?*", and second, "*Which cognitive elements influence the quality of an exploratory process analysis?*".

Although EDA has received significant attention since Tukey (1977) [16], its application to process analysis remains relatively underexplored. Only a few studies have addressed this intersection [12, 23], and a clear definition of what EPA has not been established. Therefore, the first part of this package aims to define EPA within our research context. We will adapt existing definitions and insights from EDA to build a suitable foundation for our domain.

This definition serves as a guideline, helping to identify the relevant activities and actions to focus on through the further course of the research. It will clarify the key characteristics that are relevant in our specific context, forming the foundation for the following work packages.

Once the definition has been established, we aim to identify the cognitive elements that influence the quality of an EPA. The goal is to pinpoint the elements that make a significant difference in either the process or the outcome of the analysis. For example, if two individuals perform the same task, and one demonstrates a more developed particular cognitive element, we expect to observe noticeable differences in how the analysis is conducted or in the results produced.

These cognitive elements will be identified through reviewing the literature and experiments. These experiments entail an EPA, where the participant is expected to perform the analysis using a think-aloud approach. A description of the task and a dataset are given, which includes a certain pattern that the participant should find. After the participants have performed the task, the experiments are transcribed. Based on these transcriptions, we look for certain patterns in the cognitive process of the participants to find the cognitive elements that seem to be used while doing the exploratory analysis. After coding the transcriptions, we categorize these codes through application of grounded theory methodology, with the aim to find the relevant cognitive elements used in the process.

Once the relevant cognitive elements have been identified, we proceed to the next phase of the work package, which focuses on developing a method for mapping these elements. This phase will be conducted using the Design Science Research methodology. We will begin by developing a manual mapping technique through an iterative process of experimentation and refinement. This method will be applied across a series of experiments as described in the first phase. After each experiment, the technique will be reviewed and adjusted to enhance its effectiveness and efficiency. This trial-and-error method allows us to optimize the mapping technique based on empirical insights. The resulting manual method will then serve as the foundation for the subsequent step of automating the mapping process. In this step, we explore the potential for an agent to mimic the behavior and actions of the human during the manual mapping method, possibly with the help of a Large Language Model.

3.2.2. Work package 2: Analysis of the cognitive elements of experts versus novices

The second work package aims to answer the second research question, 'How do the cognitive elements from novices differ from those of experts in exploratory process analysis?' Based on the automated mapping method developed in the first work package, this phase focuses on applying the method to systematically capture the cognitive elements of novice and expert participants, and consequently comparing the results. The goal is to uncover significant differences and similarities in how these two groups approach EPA. The scope of this work package will be refined based on insights gained from the first work package, as we build on the results from that phase.

Participants will be recruited and assigned an EPA task, during which their cognitive elements will be mapped using the developed method. After data collection, the participants will be categorized along a continuum of expertise. This will be based on how the process is executed, the efficiency, the effectiveness, and the results of the analysis. Given that expertise is not binary but exists on a spectrum and is most likely multi-faceted, we will focus on the participants at the extreme ends of the continuum. We focus on the participants who most clearly represent novices and experts.

For each of the two groups, a prototypical model of the cognitive elements will be constructed by identifying common patterns and structures in the elements of the participants within each group. These prototypical models form a representative framework for understanding how novices and experts typically think and their cognitive elements during an EPA.

The final step in this work package involves a comparative analysis of the two prototypical models. This comparison should highlight key differences and potential similarities in the cognitive models. We focus specifically on identifying the gaps and deficiencies in the cognitive elements of novices. These insights can serve as a guide for future training or support mechanisms that help novices or non-experts develop the cognitive elements to reach a higher level of expertise.

3.2.3. Work package 3: Design and evaluation of interventions to improve the cognitive elements and increase performance

This work package addresses the third research question, regarding the interventions that help to reshape and optimize the elements of a novice. Building on earlier insights, this phase explores targeted interventions to address cognitive gaps in novices. The goal is to support the development of expert-level thinking and improve performance in EPA through actionable interventions.

We aim to lay the groundwork for an intelligent tutoring system that uses these targeted interventions to train users and reshape their cognitive models toward expert-level performance. This involves exploring the possibilities of an adaptive agent that can identify deficiencies in a user's cognitive model and generate targeted exercises to address them. To achieve this, the agent developed in the first work package could be extended to detect deficiencies by comparing the user's cognitive model with the prototypical expert model from the second work package.

We aim to identify the interventions that will eliminate or minimize the weaknesses in the cognitive model of an individual. The design of these interventions will be based on the (automated) cognitive mapping method from the first work package, the comparative analysis from the second work package, and relevant literature on cognitive training and learning. To design effective interventions, we aim to link specific cognitive deficiencies to targeted training strategies and exercises. For each commonly observed weakness in the second work package, we will identify the most suitable techniques or interventions that can address and help overcome the deficiency. These techniques and interventions will be drawn from existing literature and established training programs.

Once this list of interventions is developed, it can form the foundation for future personalized training programs or tutoring systems. These systems could support individuals in strengthening and improving their cognitive models, and thus enhancing their performance in exploratory process analyses.

4. Preliminary results

In the first work package, three think-aloud experiments were conducted during an individual process analysis. The results were open-coded, revealing four categories. These were refined through a literature review on cognitive modeling and mental models. These categories led to the four key cognitive elements that we believe influence the analyst's performance in EPA: beliefs, and three types of knowledge: declarative, procedural, and conditional knowledge.

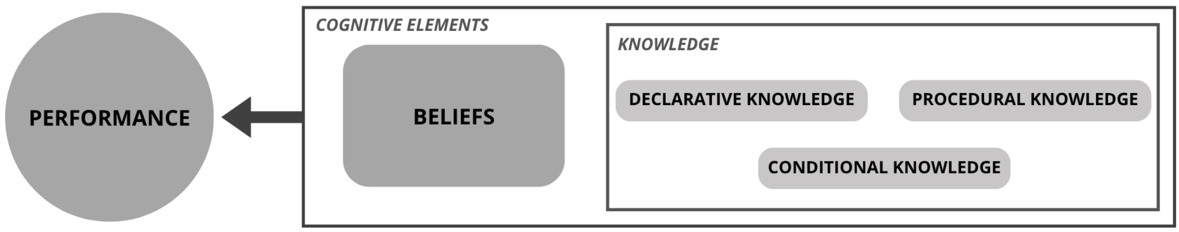

Figure 1: Key cognitive elements during exploratory process analysis

Beliefs are the analyst's internal ideas about the world, the process, the data, or process mining techniques. They guide attention and interpretation during analysis, influencing which elements are noticed or prioritized. Similar to expectations in situation awareness theory [17], these beliefs shape where analysts look for information and how they make sense of what they see.

The **declarative knowledge** entails the knowledge of facts, concepts and information. This represents the 'knowing what' [20], such as the ability to explain what an event log is.

Procedural knowledge, or 'knowing how' [19], refers to understanding and executing specific tasks, such as knowing how an algorithm works or how to calculate the mean duration of an activity.

Finally, being able to identify when and why it is relevant to perform certain actions is categorized as **conditional knowledge**. The conditional knowledge is a strategic application of the declarative and procedural knowledge, thus 'knowing when and why' [21]. This knowledge involves a trigger, an action, and a goal. For example, if the analyst sees several outliers (*trigger*), he will create and add a filter (*action*) to extract a subset of the data for further analysis (*goal*).

While declarative and procedural knowledge provide the necessary foundation, beliefs and conditional knowledge have the most impact on performance. The ability to apply knowledge strategically and be guided by well-formed beliefs likely distinguishes competent analysts from those who truly excel.

Acknowledgments

This study was supported by the Special Research Fund (BOF) of Hasselt University under Grant No. BOF24OWB25.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] T. Weijters, W. van der Aalst, Process Mining: Discovering Workflow Models from Event-based Data, in: Belgium-Netherlands Conference on Artificial Intelligence, Citeseer, 2001.

[2] W. M. P. v. d. Aalst, A. J. M. M. Weijters, Process Mining: A Research Agenda, Computers in Industry 53 (2004) 231–244. doi:https://doi.org/10.1016/j.compind.2003.10.001.

[3] W. Aalst, van der, Process Mining : Discovery, Conformance and Enhancement of Business Processes, Springer, Germany, 2011. doi:10.1007/978-3-642-19345-3.

[4] W. Van Der Aalst, Process Mining: Overview and Opportunities, ACM Transactions on Management Information Systems 3 (2012) 1–17. doi:10.1145/2229156.2229157.

[5] W. Van Der Aalst, W. van der Aalst, Data Science in Action, Springer, 2016.

[6] W. M. van der Aalst, Process Mining: A 360 Degree Overview, in: Process Mining Handbook, Springer, 2022, pp. 3–34.

[7] A. Augusto, J. Carmona, E. Verbeek, Advanced Process Discovery Techniques, in: Process mining handbook, Springer, 2022, pp. 76–107.

[8] S. Dunzer, M. Stierle, M. Matzner, S. Baier, Conformance Checking: A State-of-the-Art Literature Review, 2019, pp. 1–10. doi:10.1145/3329007.3329014.

[9] F. A. Yasmin, F. A. Bukhsh, P. D. A. Silva, Process Enhancement in Process Mining: A Literature Review, in: CEUR workshop proceedings, volume 2270, CEUR, 2018, pp. 65–72.

[10] E. Sorokina, P. Soffer, I. Hadar, U. Leron, F. Zerbato, B. Weber, PEM4PPM: A Cognitive Perspective on the Process of Process Mining, in: C. Di Francescomarino, A. Burattin, C. Janiesch, S. Sadiq (Eds.), Business Process Management, Springer Nature Switzerland, Cham, 2023, pp. 465–481.

[11] I. J. Good, The Philosophy of Exploratory Data Analysis, Philosophy of science 50 (1983) 283–295. Publisher: Cambridge University Press.

[12] F. Zerbato, P. Soffer, B. Weber, Initial Insights into Exploratory Process Mining Practices, in: A. Polyvyanyy, M. T. Wynn, A. Van Looy, M. Reichert (Eds.), Business Process Management Forum, volume 427, Springer International Publishing, Cham, 2021, pp. 145–161. doi:10.1007/978-3-030-85440-9_9, series Title: Lecture Notes in Business Information Processing.

[13] F. Zerbato, P. Soffer, B. Weber, Process Mining Practices: Evidence from Interviews, in: International Conference on Business Process Management, Springer, 2022, pp. 268–285.

[14] F. Zerbato, L. Zimmermann, H. Völzer, B. Weber, ProMiSE: Process Mining Support for End-Users., in: CAiSE Research Projects Exhibition, 2023, pp. 60–67.

[15] J. Ammann, L. Lohoff, B. Wurm, T. Hess, How do Process Mining Users Act, Think, and Feel? An Explorative Study of Process Mining Use Patterns, Business & Information Systems Engineering (2025) 1–23. Publisher: Springer.

[16] J. W. Tukey, Exploratory Data Analysis, Addison-Wesley series in behavioral science, Addison-Wesley publ, Reading (Mass.) Menlo Park (Calif.) London [etc.], 1977.

[17] M. R. Endsley, Expertise and Situation Awareness, in: K. A. Ericsson, R. R. Hoffman, A. Kozbelt, A. M. Williams (Eds.), The Cambridge Handbook of Expertise and Expert Performance, 2 ed., Cambridge University Press, 2018, pp. 714–742. doi:10.1017/9781316480748.037.

[18] J. R. Anderson, C. Lebiere, The Atomic Components of Thought, 1998.

[19] J. R. Anderson, Cognitive Psychology and its Implications, 3rd ed., Cognitive psychology and its implications, 3rd ed. (1990) xvi, 519–xvi, 519. Place: New York, NY, US Publisher: W H Freeman/Times Books/ Henry Holt & Co.

[20] J. Hong, Z. Pi, J. Yang, Learning Declarative and Procedural Knowledge Via Video Lectures: Cognitive Load and Learning Effectiveness, Innovations in Education and Teaching International 55 (2016) 74–81. doi:10.1080/14703297.2016.1237371.

[21] S. G. Paris, M. Y. Lipson, K. K. Wixson, Becoming A Strategic Reader, Contemporary Educational Psychology 8 (1983) 293–316. doi:https://doi.org/10.1016/0361-476X(83)90018-8.

[22] P. M. Fitts, M. Posner, Human Performance, Human performance, Brooks/Cole, Oxford, England, 1967.

[23] F. Zerbato, A. Burattin, H. Völzer, P. N. Becker, E. Boscaini, B. Weber, Supporting Provenance

and Data Awareness in Exploratory Process Mining, in: International Conference on Advanced Information Systems Engineering, Springer, 2023, pp. 454–470.