Information Systems 125 (2024) 102424

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/is

On the cognitive and behavioral effects of abstraction and fragmentation in modularized process models

Clemens Schreiber [a], Amine Abbad-Andaloussi [b,*], Barbara Weber [b]

[a] Karlsruhe Institute of Technology, Karlsruhe, Germany
[b] University of St. Gallen, St. Gallen, Switzerland

ARTICLE INFO

Keywords:
Process model understandability
Modularization
Cognitive load theory
Eye-tracking

ABSTRACT

Process model comprehension is essential for a variety of technical and managerial tasks. To facilitate comprehension, process models are often divided into subprocesses when they reach a certain size. However, depending on the task type this can either support or impede comprehension. To investigate this hypothesis, we conduct a comprehensive eye-tracking study, where we test two different types of comprehension tasks. These are local tasks focusing on a single subprocess, thereby benefiting from abstraction (i.e., irrelevant information is hidden), and global tasks comprising multiple subprocesses, thereby also benefiting from abstraction but impeded by fragmentation (i.e., relevant information is distributed across multiple fragments). Our subsequent analysis at task (coarse-grained) and phase (fine-grained) levels confirms the opposing effects of abstraction and fragmentation. For global tasks, we observe lower task comprehension, higher cognitive load, as well as more complex search and inference behaviors, when compared to local ones. An additional qualitative analysis of search and inference phases, based on process maps and time series, provides additional insights into the evolution of information processing and confirms the differences between the two task types. The fine-grained analysis at the phase level is based on a novel research method, allowing to clearly separate information search from information inference. We provide an extensive validation of this research method. The outcome of this work provides a more thorough understanding of the effects of fragmentation, in the context of modularized process models, at a coarse-grained level as well as at a fine-grained level, allowing for the development of task- and user-centric support, and opening up future research opportunities to further investigate information processing during process comprehension.

1. Introduction

Understanding process models is fundamental for carrying out many technical and managerial tasks [1,2]. On the technical side, activities focused on the maintenance, refinement, and re-engineering of process models all depend on one's ability to comprehend the model at hand and use adequate methods and tools to adapt or extend its functionalities [2]. As for the managerial side, process comprehension plays a crucial role in supporting requirement elicitation as well as in fostering effective communication between domain experts and IT specialists [1,3]. To be able to react to continuously evolving requirements, process modelers need initially to understand the current state of a process model before they can adapt it. Here, comprehension can be crucial to increase performance and productivity. Furthermore, even though process models can be generated automatically using process mining [4] or generative machine learning [5] techniques, understanding the generated models remains an essential task.

However, predominant research efforts and tool development within the Business Process Management (BPM) domain concentrate on the inherent characteristics of model artifacts, specifically the graphical representation of processes. This focus has resulted in a limited exploration of the task-oriented perspective [6], which encompasses the types of tasks a process model aims to facilitate. Understanding this perspective is pivotal for discerning the required information and support in a given context. This study addresses this gap by examining two recurrent tasks frequently encountered in practical applications: *local* and *global* tasks. Local tasks focus only on a single part of the model, i.e., a single sub-process, while global tasks encompass global aspects and require an overview of multiple parts of the model, i.e., several sub-processes. Such tasks are for example commonly encountered in software engineering, where software developers need to understand how to deploy and maintain single microservices (local task), while they also need to understand the interdependencies between microservices (global task) [7]. Similarly, in logistics, tasks

* Corresponding author.
E-mail addresses: clemens.schreiber@kit.edu (C. Schreiber), amine.abbad-andaloussi@unisg.ch (A. Abbad-Andaloussi), barbara.weber@unisg.ch (B. Weber).

https://doi.org/10.1016/j.is.2024.102424
Received 19 January 2024; Received in revised form 25 April 2024; Accepted 2 July 2024
Available online 6 July 2024

can relate to single sub-processes (e.g., unloading cargo from a ship) or to multiple sub-processes (e.g., storing the cargo, which encompasses multiple sub-processes, such as registration, storage assignment, and storage) [8].

In both cases a large system is decomposed into smaller interlinked modules, which is commonly referred to as modularization. In this study we envision modularization in the context of information systems configuration, which is commonly used to facilitate the comprehension of these systems. This assumption is however challenged by the existing research on process model comprehension, which shows that modularization does not consistently facilitate the performance of different comprehension tasks [2,9,10]. One potential explanation for this is related to the two opposing effects of *abstraction* and *fragmentation* [2]. Modularization facilitates abstraction by supporting information hiding and pattern recognition, yet simultaneously induces fragmentation by necessitating users to distribute their attention across diverse fragments to locate pertinent information. Consequently, local tasks, confined to a single module, could benefit from abstraction, thereby becoming easier to solve. On the other hand, global tasks, involving multiple modules, may also benefit from abstraction but encounter impediments due to fragmentation, rendering them more challenging to perform. The apparent relationship between modularization, task type, and the cognitive aspects of model comprehension is however not well investigated and requires further empirical validation [2,9,11,12].

Moreover, there is a need to look into the process of comprehending process models to better understand how the processing of information unfolds in different task types [6,13,14]. While there exists evidence that the task type has a significant impact on information search and processing [15–17], this has so far not been investigated in the context of process model comprehension [1,6,13]. Investigating the cognitive processing steps in local and global tasks does not only help to better understand the effects of fragmentation and abstraction, but also to potentially provide better context-driven support during task performance. In this regard, we will look at two key processing steps during comprehension tasks: *search* and *inference* [6,13,18]. While search denotes the identification and separation of relevant from non-relevant information [19], inference refers to the creative cognitive process of inferring new insights after the recognition of relevant information [18]. Knowing the task-dependent characteristics of these two processing steps would potentially allow to dynamically highlight relevant information [12,16], to provide context-driven search assistance [15,17], to give better instructions and guidance during process modeling workshops [14], or to better adjust visual modeling notations [13].

For this reason, this paper extends our previous work [20] by providing a more thorough and detailed examination of how the task type (local or global) impacts process model comprehension during information search and inference phases. Accordingly, the presented analysis in this paper is split into a *coarse-grained* and a *fine-grained* analysis (cf. Fig. 1). The coarse-grained analysis is conducted at a task level and compares the task types based on the differences observed during the overall task execution. The fine-grained analysis is conducted at a phase level and compares the task types based on the differences observed during the two distinct comprehension phases: information search and information inference. We consider these particular comprehension phases since they appear to exhibit significant differences in information processing [1,6,13]. Since the coarse-grained analysis is extensively covered by our previous work [20], we will only provide an overview of these results. The main focus of this paper is the fine-grained analysis, which extends our previous research by providing more detailed insights into the differences between task types when comprehending a modularized process model. The overall goal of our study is thereby to better understand how local and global tasks are affected by abstraction and fragmentation in the context of modularized process models.

The ensuing research questions can be summarized as follows:

RQ1 How do users' comprehension and cognitive load differ between local and global tasks?
RQ2 How do users' search and inference behavior differ between local and global tasks?

While RQ1 was extensively addressed in our previous study [20], RQ2 was so far only considered at a coarse-grained task level. Our previous analysis [20] shows that in global tasks, information search is more complex, and it requires more cognitive effort to integrate the information when compared to local tasks. However, this does not consider how search and inference behavior unfolds over time, depending on the task type. Building on these gained insights, we further deepen our analysis by showing that we can indeed observe a difference between the search and inference phases and that these phases significantly differ based on the task type. Furthermore, we provide a qualitative analysis, showing the differences in information processing based on process maps (depicting structural differences in visual behavior) and a time-series analysis (depicting how the visual behavior evolves). Overall, our study shows how local and global tasks impact comprehension and overall cognitive load, but also provides new insights into the implications of task locality on process model comprehension at a level of granularity that has been, so far, unexplored in the literature. Besides a better understanding of the effects of local and global tasks, the new insights allow for more targeted user support based on comprehension phase and task type, e.g., by providing context specific information or guiding the attention of a user to the relevant information.

Our empirical and qualitative investigations are based on eye-tracking. We design an eye-tracking experiment, consisting of a modularized process model, based on the fragment-based modeling approach [21], and a set of local and global comprehension tasks. Based on the experiment we assess the impact of the task locality on users' task comprehension, overall cognitive load, search behavior, and inference behavior. We cover conventional model comprehension and subjective cognitive load metrics as outlined in existing literature [1,2]. Additionally, we integrate advanced objective cognitive load metrics derived from eye-tracking. This comprehensive set of measures enables a multi-perspective empirical analysis, enhancing the robustness of our conclusions. Finally, the empirical observations are supported by our qualitative analysis of the visual behavior.

Our results demonstrate the importance of the task type, underscoring its significant influence at the coarse-grained, as well as fine-grained level. The coarse-grained analysis reveals that global tasks, in comparison to local tasks, yield lower model comprehension and higher cognitive load, due to the fragmentation effect. This observation is supported by the fine-grained analysis, which shows that in global tasks, both search and inference phase are longer and characterized by more complex visual behavior. Additionally, the fragmentation effect is more pronounced during inference, than during search. These observations are also supported by our qualitative analysis based on process maps and time series.

Overall, this study provides empirical validation for the cognitive and behavioral effects of global and local tasks at both coarse-grained and fine-grained levels. Furthermore, our applied research model serves as a framework for examining users' comprehension, cognitive load, search complexity, and inference complexity for process model comprehension tasks, using a wide array of measures. We additionally introduce and validate a novel research method, which allows for the segmentation of search and inference phases during information processing, thereby opening up opportunities for future research. The findings of our study emphasize the relevance of the task perspective for the analysis of process model comprehension. Based on the fine-grained analysis we also provide new insights regarding the task-dependent information processing during process model comprehension. In this regard, we propose several strategies to mitigate negative effects of global tasks, to especially facilitate process model comprehension concerning information search and inference.

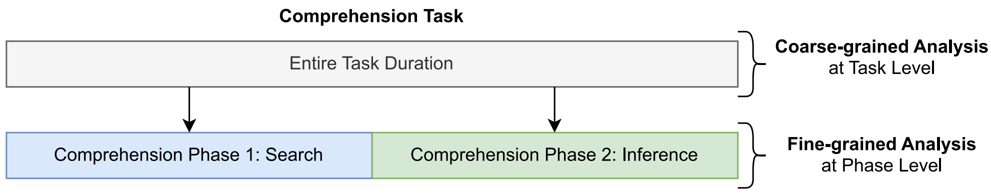


Fig. 1. Overview of the coarse- and fine-grained analysis.

In the remainder, Sections 2 and 3 provide the background and related work respectively. Section 4 explains our research method. Sections 5 and 6 present the findings and discuss them. Finally, Section 7 concludes the paper and delineates future work.

2. Related work

This section presents the related work on studies investigating process model comprehension (cf. Section 2.1) and information processing (cf. Section 2.2).

2.1. Process model comprehension from a task perspective

Over the past decades, a number of factors challenging the comprehension of process models have been identified in the literature [1,22]. Notably, the task type has received increased attention.

The cognitive fit theory of Vessey [23] posits that a good fit between the task at hand and the problem representation (i.e., the artifact presented to the user e.g., process model) facilitates the problem-solving procedure [23,24]. Accordingly, several authors investigated the cognitive fit between different types of tasks and representations. Vessey and Galletta [25], for instance, differentiated symbolic tasks targeting specific data objects and spatial tasks focusing on relationships between data objects. Then, the authors investigated the cognitive fit between these tasks and different tabular and graphical representations. Similarly, Ritchi et al. [26] defined schema tasks that are based uniquely on the process model and non-schema tasks requiring knowledge beyond the process model. Then, the authors examined the extent to which textual and graphical representations support better cognitive fit with these tasks.

Closely related to our work, Dunn and Grabski [24] differentiated *local tasks*, where attention is assumed to be directed to a small area, and *global tasks* where attention is expected to split over several areas. Then, the authors investigated the cognitive fit between these tasks and accounting models (i.e., Debit–Credit-Account, Resources-Events-Agents) represented as Entity Relationship (ER) diagrams, tables, and text. The authors found that cognitive fit is influenced by the locality of information relevant for solving the task. In our study, users' cognitive fit will likely be influenced by the locality of the given tasks when engaging with fragmented process models. Indeed, when solving a local task requiring information within a single fragment users could experience better cognitive fit as they would benefit from the abstraction support of modularization. Conversely, when solving a global task in which the relevant information is distributed over several fragments, the cognitive fit will be limited due to the effect of fragmentation. Such a limitation is likely to impede the problem-solving procedure, hence challenging the users, raising their cognitive load, and eventually affecting their comprehension of the model at hand.

The impact of task locality on modularization was suggested by Reijers and Mendling [9] as a result of some inconclusive findings about modularization. Following that, it was investigated in the context of vertical modularization with a limited set of cognitive load measures by Zugal [2]. Beyond the use of a different modularization approach, this study did not use eye-tracking, and also only conducted a coarse-grained analysis. More recently, a research model based on a multi-modal measurement of cognitive load was suggested for modularized declarative process models [27]. Based on the theoretical background presented in this section together with the insights of the existing literature, in our work, we further substantiate and validate the impact of task locality on horizontal modularization using the cognitive load measures introduced in Section 3.4.

2.2. Information processing in process model comprehension

In the literature users' behavior has been investigated in the context of studies on *model creation* [28–32][1] and studies on *model comprehension* [13,36–41]. Our work relates to the latter set of studies. Therein, the conducted behavioral analyses yielded several insights. For instance, using eye-tracking, Petrusal and Mendling [36] found that the task-relevant parts of imperative process models (in BPMN[2]) receive more attention during comprehension tasks than the non-relevant ones. Haisjackl et al. [40] used think-aloud data and observed that the reading of declarative process models (in Declare [42]) involves typically two strategies. The first is based on the execution order of the model activities while the second relies on the orientation of the model layout. In the same context of declarative process models (in DCR [43]), following an eye-tracking analysis of users' behavior, Abbad-Andaloussi et al. [41] suggested that users follow either an exploratory or goal-oriented strategy when reading these process models. While the aforementioned articles investigated users' behavior during comprehension tasks, they did not specifically focus on the identification of specific behavioral phases in users' data. This contribution was, in turn, made by other studies. Kim et al. [37], for instance, showed that in comprehension tasks involving multiple interlinked diagrams, users first engage in a search process, then develop hypotheses about the target system. These two phases were named perceptual and conceptual integration. The study did, however, not include an eye-tracking analysis and was exclusively based on verbal data analysis. A similar approach was chosen by Shanks et al. [39]. They also used verbal protocols to identify specific phases during task comprehension. As a result, they identified five different behavioral categories, which they termed: preparing, identifying, understanding, articulating, and evaluating. The most common observed sequence of behavior was task understanding followed by identifying the relevant area of the model or preparing the solution before conducting a final evaluation. We could also identify two studies [13,38] using eye-tracking to detect specific phases during comprehension tasks on process models. In [13], Bera et al. identified two relevant cognitive processes, which have a significant impact on process model comprehension performance: Attention (paid to specific model parts) and cognitive integration. To detect attention, fixation count and fixation duration were used. For the detection of cognitive integration, a combination of eye-tracking and verbal protocol analyses were used. In a similar way, Wang et al. [22] show that the searching and integration of process model information can be facilitated by providing additional external information. The conducted eye-tracking study considers three types of information embedding: rule linking (interlinking elements of a process model),

[1] While the listed studies focus on the creation of process models, there also exist some studies focusing on the creation of other conceptual models [33–35].

[2] https://www.omg.org/spec/BPMN/2.0/ (Accessed: 26 March 2024).

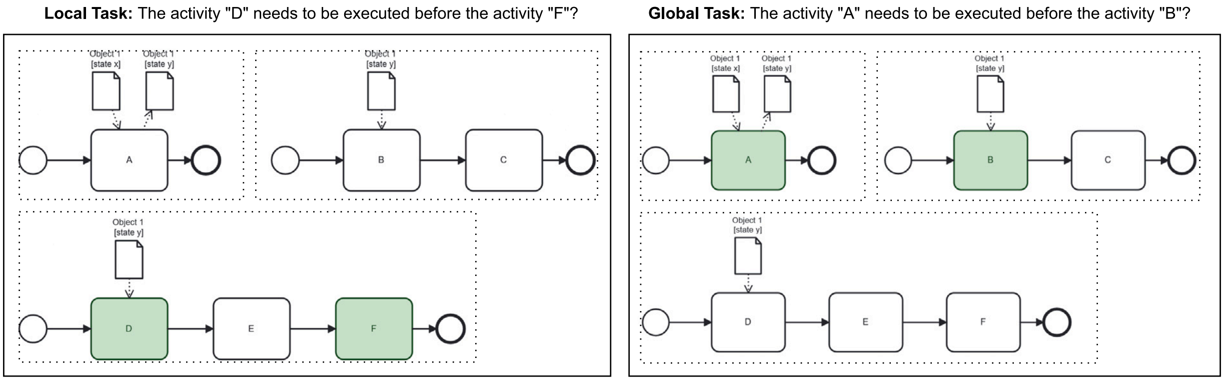


Fig. 2. Comparison between a local and global task based on a modularized process model consisting of three fragments. While both task types benefit from abstraction, the global task is additionally impeded by fragmentation, since the two relevant activities "A" and "B" are distributed among two fragments.

text, and diagrammatic embedding within the model. A recent study by Winter et al. [38] investigated how different gaze patterns occur during comprehension tasks. The study used process models of different complexity and mappings, i.e., representations of the process model elements. Based on heat maps, focus maps, and scan paths analyses, three different reoccurring gaze patterns were identified: orientation pattern, comprehension pattern, and congruence pattern. Our study aligns with studies aiming at investigating behavioral phases in users' eye-tracking data [13,38], with a particular focus on search and inference phases.

3. Background

This section elaborates on the theoretical background relevant to our study. Section 3.1 presents the theories underlying abstraction and fragmentation. Sections 3.2 and 3.3 introduce foundations on process model modularization and fragment-based process modeling. Sections 3.4 and 3.5 provide additional theoretical background on comprehension and cognitive load, as well as search and inference behavior.

3.1. The effects of abstraction and fragmentation in modularized process models

The effect of modularization on the comprehension of business process models is a well-researched subject (for an overview see the work by Zugal [2]). Yet, existing empirical work does not provide conclusive results on whether modularization has a positive or negative impact on model comprehension [1,2]. A potential explanation for the inconsistent results obtained in existing studies is the opposing effects of *abstraction* and *fragmentation* [2,9,44], caused by the spatial separation of modules.

Abstraction arises in modularization from the division of a process model into sub-components (or fragments), allowing users to focus their attention on task-relevant fragments, while irrelevant ones are concealed. This can lead to improved performance in comprehension tasks [45] and support the recognition of patterns, which might be less apparent when looking at the entirety of a larger model [46]. For this reason process modeling guidelines commonly recommend to decompose large process models into smaller sub-components, when they reach a certain size [46]. The guidelines thereby rely on the assumption that abstraction fosters model comprehension and alleviates users' cognitive load. Fragmentation, on the other hand, can hinder the positive effect of modularization to some extent, especially when the task-relevant information is distributed among multiple fragments. In this case, a user needs to shift the attention among the spatially separated sub-components, which can lead to the *split attention effect* [2,47]. The users' attention thereby gets diverted by the separated locations and is exposed to distraction. Indeed, when users need to keep a piece of information in their mind while looking for another one to integrate with, it is likely that their memory will decay after reaching its temporal limits, resulting in increased difficulty in recalling the relevant information that was fixed previously. Consequently, fragmentation can make model comprehension more difficult and cause higher cognitive effort. This is also potentially reflected in the users' search and inference behavior, due to the spatial separation of information.

The two opposing effects of abstraction and fragmentation are well reflected by the distinction of *local* and *global* tasks. While local tasks benefit from abstraction and do not suffer from fragmentation, since they require only information within a single fragment, global tasks also benefit from abstraction, but suffer from fragmentation, since they require information, distributed among multiple fragments (cf. Fig. 2). Accordingly, solving global tasks is expected to pose greater challenges to users compared to local tasks. This conjecture aligns with the cognitive fit theory [23], which asserts that the fit between the task and the problem representation (e.g., a modularized process model presented to the user) influences the problem-solving process [23,24]. Since the task-model fit is lower for global tasks than for local ones (due to fragmentation), the problem-solving process is expected to be more intricate, potentially affecting model comprehension, users' cognitive load, and visual behavior.

Our research contributes to the current body of knowledge concerning the effects of abstraction and fragmentation [2,9,11,12] by (1) investigating the effects in the light of horizontal rather than vertical modularization, (2) employing a diverse set of measures to analyze the effects with respect to users' comprehension and cognitive load, and (3) studying users' information processing based on search and inference behavior, when facing local and global tasks. The three subsequent sections introduce the concepts and measures used to analyze the effects of task locality on task and phase levels in the context of modularized process models.

3.2. Modularization in process modeling

Modularization refers to the systematic decomposition of a system into interconnected modules, each possessing self-contained properties [48]. Within the literature on process modeling [10,49], three types of modularization have been identified: *vertical*, *horizontal*, and *orthogonal*. Vertical modularization involves decomposing the process into sub-processes using a hierarchical structure [10,49]. Conversely, horizontal modularization seeks to partition the process into interconnected fragments. The fragments can be connected in multiple ways, such that they can be executed hierarchically, sequentially, or concurrently [10,24]. Finally, orthogonal modularization partitions a process based on cross-cutting concerns, such as privacy and security, affecting multiple fragments of the model [10,24]. For example, a user could be asked to confirm her identity at multiple stages of a process, for security reasons. In orthogonal modularization, such cross-cutting concerns are depicted in separate modules.

Given that horizontal modularization allows for the separation of process modules without imposing hierarchical relationships or concerns about cross-cutting issues, this approach emerges as more universally applicable. Consequently, it is well-suited for an in-depth exploration of the impacts of task locality on users' comprehension and cognitive load. Although a generally positive influence of horizontal modularization on model comprehension is postulated based on prior research [10], our investigation delves into the specific question of how horizontal modularization influences comprehension tasks, distinguishing between those centered on single modules (local) and those focused on multiple modules (global). To facilitate this exploration, we adopt the fragment-based approach outlined by Hewelt and Weske [21] (cf., Section 3.3) as a representative for this form of modularization.

3.3. Fragment-based process modeling approach

The fragment-based modeling approach poses an extension to the commonly used Business Process Modeling and Notation (BPMN) [21]. Based on this extension it is possible to adapt the case-based modeling paradigm, allowing for flexible process execution [8,50]. The different process fragments are thereby connected based on data constraints, defining input and output conditions for the different process activities. In this way, the order of the fragments is not necessarily bound by any hierarchical constraints and can therefore be considered a horizontal modularization type (cf. Section 3.2). At run time, the fragments can be executed independently from each other, as long as the defined data constraints are fulfilled. These constraints refer to data objects, which are defined by a data type and state. A data object can change its state throughout the process execution, which is in addition to the process model fragments, depicted in a labeled transition system called lifecycle.

Fig. 3 provides an example of the fragment-based modeling approach, as it was used in our experiment. The two fragments "F1_Offload_Container" and "F1_Scan_Container" are connected based on the data object "Container". As soon as the activity "Load container onto the forklift truck" in the first fragment is executed the state of the container changes to "lifted' and the second fragment can be executed. It is important to notice that the second fragment depends on two different data objects, i.e., "Container" and "ERP file". The container lifecycle "C1_Container" additionally shows that the containers' final state is reached when it is "loaded".

In addition to the fragment-based approach, there also exist other modeling techniques, following the case-based modeling paradigm. This includes techniques, linking process fragments based on events instead of data objects (cf. overview by Krumeich et al. [51]), as well as extensions of other modeling languages such as (colored) Petri nets [52]. However, our decision to use the fragment-based approach [21] is based on the popularity of the BPMN language and its clear execution semantic, which can be easily assimilated by readers, even without previous modeling experience. The introduction of the data objects to interlink the fragments, further increases the expressiveness of BPMN, allowing to model real-world behavior in a more precise manner.

3.4. Comprehension and cognitive load

For the coarse-grained analyses (cf. Fig. 1), we focus on the difference between local and global tasks in terms of comprehension and cognitive load.

Comprehension. Comprehension refers to the cognitive process of interpreting an artifact, such as a text or a process model, and constructing a mental representation thereof [53]. In research, there exist many different types of comprehension measures, that are used to investigate the impact of different process model representation factors (e.g., presentation medium, model complexity, nomenclature of activities, task type) on the ability of a user to grasp the information carried by a model (see overview by Figl [1]). We will apply two common measures thereof, which are *comprehension accuracy* and *comprehension efficiency*.

Cognitive Load. The cognitive load theory postulates that when individuals approach the maximal capacity of their working memory, they experience a task at hand to be more difficult, leading to a detrimental impact on task performance and an increased susceptibility to errors or erroneous decision-making [54,55]. In terms of process model comprehension, it would make the comprehension of a depicted process more difficult and would lead to a higher risk of misinterpreting the modeled process behavior. Thus, cognitive load can be assessed *subjectively*, through introspection, or *objectively*, through observable changes in behavior and cognitive states [54,56,57]. One common measure to distinguish between varying levels of cognitive load is the subjective evaluation of *perceived difficulty*, e.g., based on a 5-point Likert scale [58]. Additionally to this subjective measure, we use measures derived from eye-tracking in our study to also capture cognitive load objectively.

Eye-tracking thereby allows to detect visual and behavioral patterns, which might otherwise, not be clear based on verbal protocols [56]. A key concept in eye-tracking is fixations. *Fixations* constitute a time interval with eye movements of very low velocity, implying that the pupil is fixated on a specific position within the visual field [56]. For the investigation of cognitive load at a task level, we will use the *average fixations duration*, which computes the average duration of fixations within a time window [56]. This measure is known to be lower when screening and scanning information, than when conducting intensive mental processing [56].

In addition to average fixation duration, we will also use a *physiological measure* to investigate cognitive load at a task level. Physiological measures can capture humans' reactions in the human body, depending on the cognitive load, required by a given task. Such reactions are triggered by the sympathetic and parasympathetic divisions of the autonomic nervous system, which are responsible for bodily functions [57,59]. The sympathetic division triggers bodily activation in response to heightened mental or physical demands, while the parasympathetic division induces relaxation during decreased demand levels. These occurrences precipitate biological responses such as pupil dilation. Research has shown, that the pupil dilation derived from the *low/high index of pupil activity (LHIPA)* has proven to be a reliable indicator of cognitive load [60]. This index segregates pupil oscillations into low and high frequencies, yielding a measure that is negatively correlated with cognitive load (i.e., a low value corresponds to a high cognitive load) [60].

3.5. Search and inference phases

For the fine-grained analyses (cf. Fig. 1), we focus on the difference between local and global tasks in terms of search and inference phases, to capture the information processing during task execution.

Search. Information search involves the identification and separation of relevant from non-relevant information [19]. In the context of process model comprehension tasks, this equates to the distinction between *task-relevant* and *task-irrelevant* process model activities [36]. How the search is conducted depends on the task type and the applied search strategy [61–63]. Since the search always involves costly resources, such as time and cognitive effort, users will stop search once they assume to have gathered sufficient information [61]. Sufficiency thereby refers to the completeness and correctness of information. In a process model comprehension task, this would imply that a person would stop looking for relevant information in a process model after all task-relevant process activities are found. This of course requires that the task-relevant activities can be clearly recognized [61]. According to Browne et al. [62], to evaluate the sufficiency of information during the search, users apply *cognitive stopping rules* [61]. Notably, for well-structured and decomposable tasks, users often rely on a particular stopping rule called *mental lists*. Following this rule, a person has a mental checklist of items that must be found before stopping the search process [62]. In the context of process model comprehension, this stopping rule might be applied, when asked to explain how a set of model activities relate to each other. Therein, during the search phase,

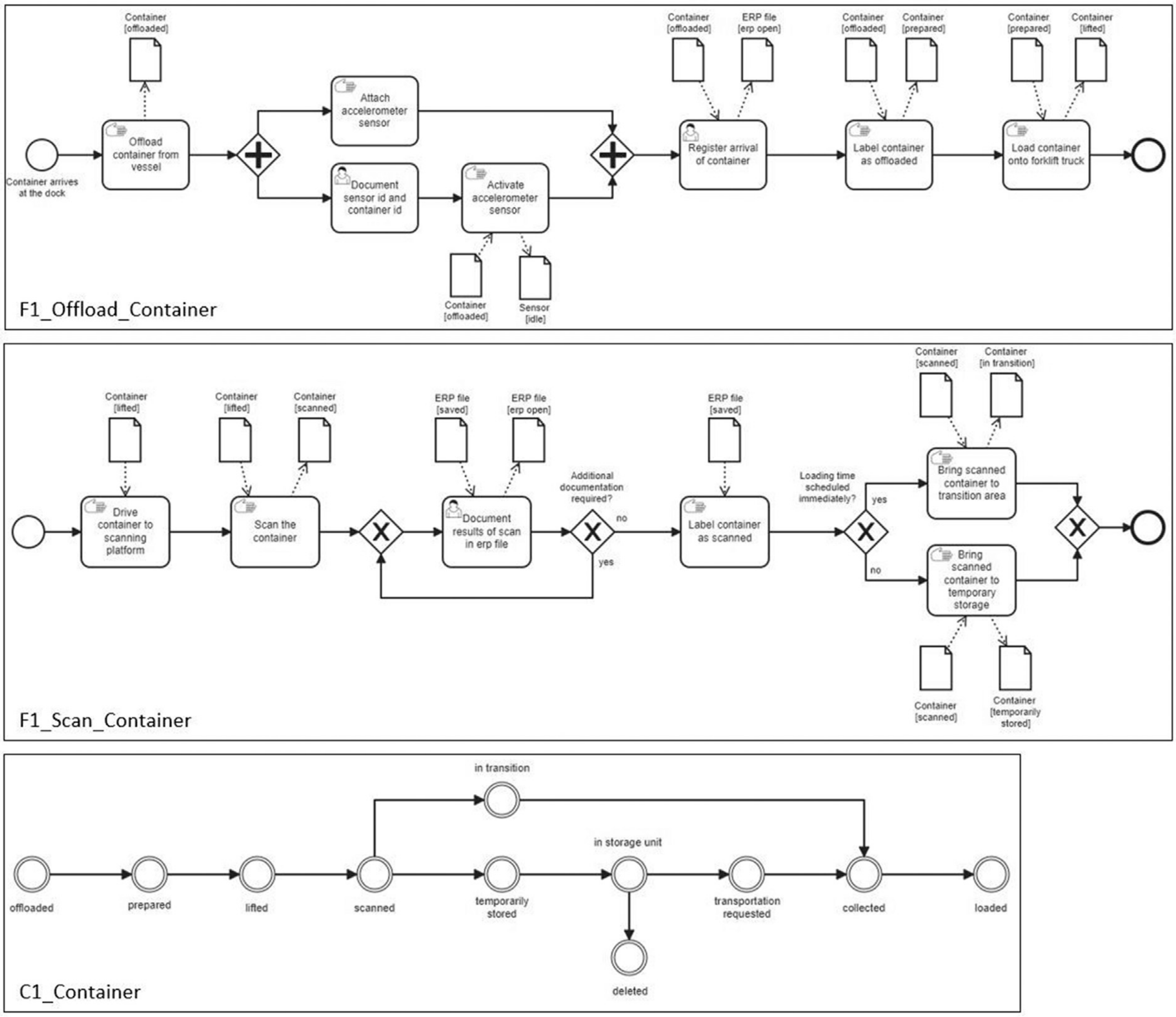


Fig. 3. Example of a horizontally modularized process model following the fragment-based approach. A better resolution is available at https://github.com/promilab/InfoSysTaskType.git.

the user is likely to have a mental checklist of relevant activities that must be found in the model before ending this phase.

Inference. Cognitive inference refers to the creative process of inferring new insights after the recognition of some relevant information [18]. Inference is mostly associated with reasoning [37]. Therein, the integration of the information extracted from the artifact with one's existing knowledge plays a primary role [37]. Inference has been studied as part of many cognitive frameworks. Kim et al. [37] for instance, identified two processes underlying the comprehension of diagrammatic representations (e.g., graphical models): *perceptual* and *conceptual* processes. While the former was associated with search and recognition of relevant information, the latter was linked to reasoning and inferring insights from the given diagram. Similarly, the multimedia learning theory [64] suggests that following the selection of the task-relevant information during a search phase, requires the organization of this information and its integration with existing knowledge, which would take place within the inference phase.

Eye-tracking measures to investigate search and inference. Our study aims at investigating the characteristics of search and inference phases during process model comprehension using visual behavior, captured by eye-tracking measures (cf. Fig. 4). For the analysis of search and inference, fixations can provide important insights. It is common practice to consider *short fixations* (with a duration of < 250 ms), separately from *long fixations* (with a duration of $>= 500$ ms) [65], allowing to differentiate between superficial processing, likely to occur during information search, and deep mental processing, likely to occur during information inference. In addition to fixations, *saccades* can provide insights into the cognitive state. Opposite to fixations, saccades denote eye movements of high velocity implying that the pupil is moving from one location (i.e., fixation) to another [56]. The *average saccade amplitude* computes the average distance traveled by saccades over a time window [56]. This measure is known to shorten when users are involved in an intensive search process or engaged in a careful inspection of an artifact [56].

A further important concept for the analysis of information processing is scan-paths. *Scan-paths* denote a sequence of fixations, reflecting the visual path followed by the user when engaging with an artifact [56]. For the analysis of a scan-path it is often of interest, whether this path includes different areas of interest (AOIs). AOIs thereby define regions in a stimulus that a researcher is interested in to investigate [56], e.g., in the context of process model comprehension, AOIs can correspond to the different elements of a process model. A scan-path includes all *first-time visits* to an AOI, as well as all *returned visits* to an AOI. Building on the scan-path, *scan-path precision* is an important measure to separate search and inference phases. This measure computes the ratio of fixations on the task-relevant areas (e.g., process model activities) to the total number of fixations on the artifact (e.g., process model) [36]. Doing so, the scan-path precision can tell about the extent to which users are particularly focused on the process activities relevant to solving a task. When conducting

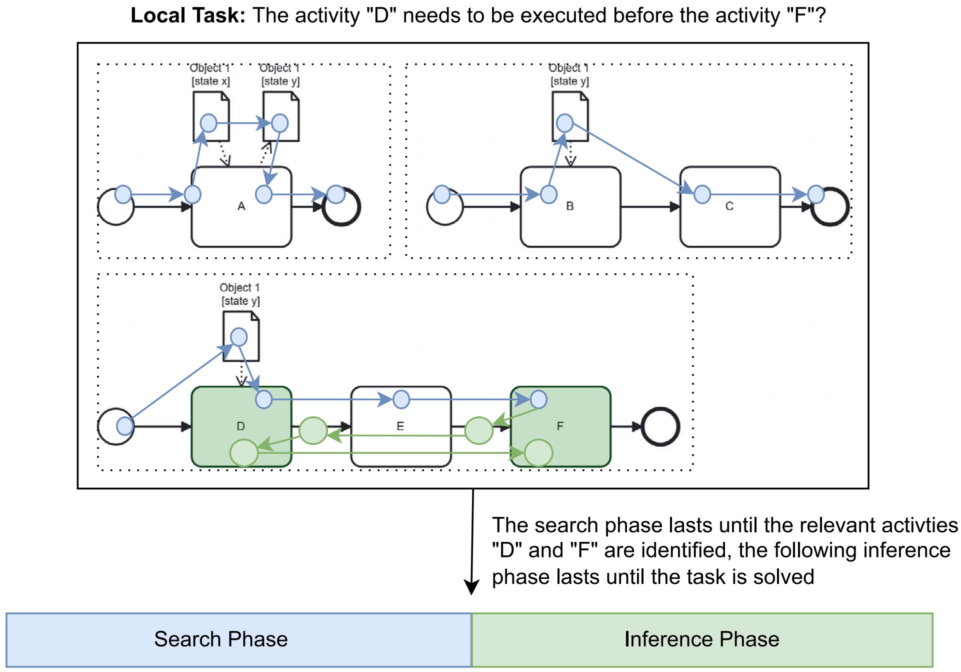


Fig. 4. Distinction between search and inference phase. The blue dots and arrows depict a scan-path during search and the green dots and arrows depict a scan-path during inference. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

search, users are looking for the task-relevant activities among a set of other non-relevant activities. This behavior would, in turn, translate into a low scan-path precision since users' fixations can land on any relevant or irrelevant activity in the model. Conversely, once all the relevant activities are identified, users would limit their attention to these activities to infer a solution to the given task. Herein, the scan-path precision value would get substantially higher. A further relevant metric for the analysis of information processing based on the scan-paths is the *entropy* of observed fixations [56]. Entropy is a measure of the uncertainty regarding the occurrence of a fixation. E.g., when a scan-path only consists of fixations on one single AOI, the entropy value will reach its lowest possible value, which is zero. High entropy values are thereby associated with exploratory behavior (i.e., search), while low values indicate fixations on a restricted area [56]. Finally, scan paths can also be depicted in the form of *process maps* [66,67]. Process maps provide a graphical representation of the scan-paths, allowing to visualize the order of fixations on different AOIs, thereby allowing to identify scan-path patterns. For example, when a user is scanning information, a process map is expected to show linear reading patterns, while during information inference, the process map is likely to show more intricate reading patterns, involving multiple loops between AOIs.

4. Research method

To investigate the effects of abstraction and fragmentation, we have conducted an eye-tracking study following the empirical standard guidelines for experiments.[3] Sections 4.1 –4.3 provide an overview of the study design, the study execution, and the data analysis procedures respectively.

4.1. Study design

4.1.1. Research model

Based on the theoretical background introduced in Section 3, our empirical study aims to investigate the impact of abstraction and fragmentation in modularized process models on users' comprehension and cognitive load at a task level (coarse-grained analysis), as well as search and inference complexity at a phase level (fine-grained analysis). Our research model is depicted in Fig. 5. As *independent variable*, we manipulate the level of fragmentation based on the *locality* of the task, which we separate into two-factor levels with *local tasks* addressing control-flow aspects located within a single process fragment, and *global tasks* addressing control-flow aspects located within two process fragments. As explained in Section 3.1, due to the nature of modularized process models both local and global tasks benefit from abstraction, but only global tasks are impeded by fragmentation. The task locality factor is expected to impact the *comprehension*, *cognitive load*, *search complexity*, and *inference complexity*, which denote the dependent variables. As shown in Fig. 5, we operationalize these theoretical constructs using the measures introduced in Sections 3.4 and 3.5. Following our research model, we formulate the following hypotheses:

Coarse-grained.
· $\mathbf{H_1}$: Global tasks yield lower task comprehension than local tasks.
· $\mathbf{H_2}$: Global tasks yield higher cognitive load than local tasks.

Fine-grained.
· $\mathbf{H_3}$: Global tasks yield more complex search behavior than local tasks.
· $\mathbf{H_4}$: Global tasks yield more complex inference behavior than local tasks.

4.1.2. Material

The experiment comprises a set of *model fragments* that capture a logistics process and a set of comprehension *tasks*, prompting the participants to identify relationships between activities within the process model based on the control-flow. These tasks could either refer to a single process fragment (local task) or to two different process fragments (global task).

Following the fragment-based modeling approach (cf. Section 3.3), the logistics process is divided into 6 process model fragments, interconnected using 3 data objects for which the state changes are represented in their respective life cycles (the complete model is depicted in the Appendix A). The process model fragments are designed following existing process modeling guidelines [46,68] ensuring a carefully aligned layout and a reasonable number of activities and gateways within each model (i.e., from 6 to 8 activities and from 2 to 4 exclusive or parallel gateways). To mitigate the effect of domain knowledge,

[3] https://github.com/acmsigsoft/EmpiricalStandards/blob/master/docs/Experiments.md (Accessed: 26 March 2024).

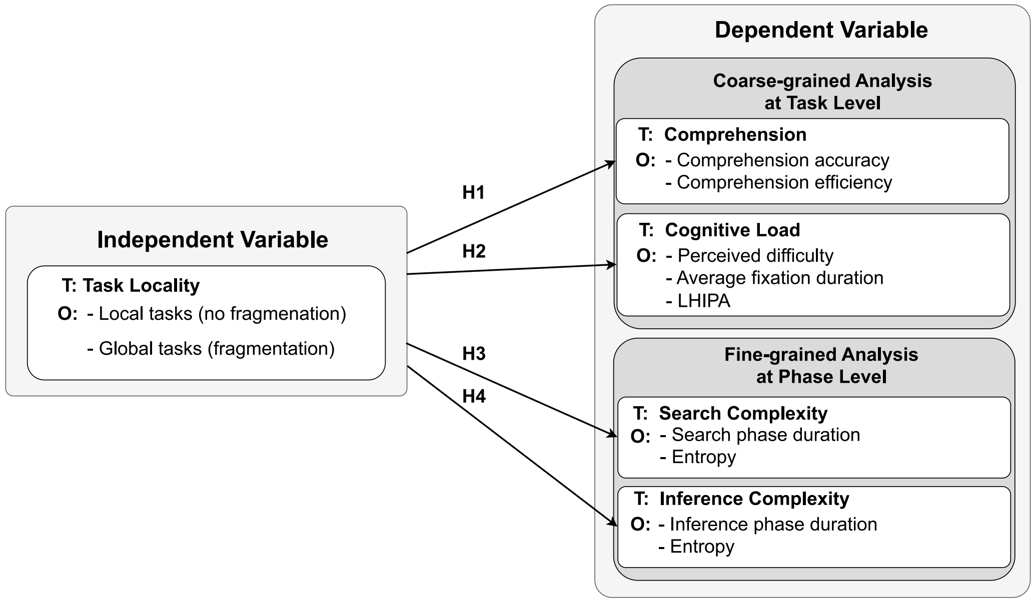


Fig. 5. Research model for hypothesis testing. *T* — Theoretical construct, O — Operationalization of construct.

layman's terms were used in the naming of activities. Overall the process model consists of 41 activities and 46 data-object inputs and outputs. It also covers the most commonly used BPMN constructs, i.e., start and end events, task, sequence flow, parallel gateway, and data-based XOR [69]. The high number of elements combined with the different BPMN constructs results in a relatively high model complexity compared to process models in similar studies [2,70].

The experiment includes 8 tasks, designed to prompt the participants on the relationship (i.e., *sequence flow, exclusiveness, repetition, or concurrency*) between activities located within the same process model fragment (i.e., referred to as *local tasks*) or in different process model fragments (i.e., referred to as *global tasks*). This ensures that the used tasks cover different workflow patterns at both local and global levels, reflecting, in turn, real-world comprehension tasks.

The tasks were formulated as statements following the template shown in Fig. 6. Depending on the semantics of the process model fragments, the participants were asked to evaluate the given statements as correct or incorrect. Throughout the design and execution of our study, we formulate our tasks in such a way that the task-relevant information can be clearly recognized in the process models (cf. Fig. 6). Furthermore, we instruct our participants to continue with search until they have found all task-relevant information, to foster the mental list stopping rule [62]. A search phase thereby lasts from the beginning of a task until the two task-relevant activities are found. The following inference phase lasts until the end of the task (cf. Fig. 4). These design decisions are meant to facilitate the separation between search and inference behavioral phases in the users' data as will be further explained in Sections 4.2 and 4.3.2.

Additionally, to avoid any learning effect, the individual tasks covered distinct aspects of the used fragments. The material deployed in our experiment is available online.[4]

The experiment was conducted within the data collection and empirical investigation framework EyeMind [71]. Fig. 7 depicts a screenshot of the EyeMind user interface used during the experiment to navigate through the tasks. The process fragments and life cycles are provided in different files, which can be accessed through the file explorer shown on the left side of the screen. Note that only a single file can be viewed at a time. The question is shown at the very top of the screen.

[4] https://github.com/promilab/InfoSysTaskType.git

4.1.3. Participants

Table 1 provides an overview of the participants' demographics. 46 participants were recruited for this experiment. 22 participants come from the University of St. Gallen, 17 participants from Karlsruhe Institute of Technology, 4 participants from the research institute Forschungszentrum für Informatik FZI in Karlsruhe, and 3 participants from Promatis an IT company located at Karlsruhe. The participants were aged between 20 and 50 years old with 63% in the range of [20–30]. The participants had different backgrounds. 22 participants were conducting research in academia, 17 participants were students at different levels of bachelor and master educations and 7 were working in the IT industry. On a familiarity scale ([1: unfamiliar, 7: very familiar], 48% of the participants affirmed to be highly familiar with BPMN (in range of [5–7]), while 42% had a low familiarity with BPMN [in range of [1–3]]. To ensure the participants' ability to take the experiment, they were all uniformly familiarized with the used BPMN concepts and how to interpret process models designed following the fragment-based approach. Additionally, a set of test tasks similar to those of the experiment were used to evaluate the participants' skills after the familiarization phase.

4.2. Experiment procedure

The experiment took place in individual eye-tracking sessions with an average duration of 1 h. As depicted in Fig. 8, the sessions started with a familiarization phase where the participants were introduced to the BPMN concepts used in the model fragments and key notions about the fragment-based modeling approach [21]. Afterwards, they were given a quiz to ensure that the participants have a sufficient understanding of the fragment-based modeling approach, and to rectify the BPMN concepts, that challenged the participant's understanding. Based on the quiz, which was conducted orally, none of the participants were deemed unsuitable to continue with the experiment, i.e., every participant could solve all provided tasks correctly, given that they could ask additional questions regarding the fragment-based modeling approach. Then, screening and demographic forms were administrated to the participants to verify their physical ability to join the eye-tracking study as well as to collect basic demographic information (e.g., gender, age range, familiarity with the investigated concepts). Prior to the data collection, the participants were seated in front of an eye-tracking device and instructed about the data collection procedure before starting the calibration of the device to accurately capture their gazes on the screen. As part of the instructions, the participants were asked to avoid head movements. Also, since the same process

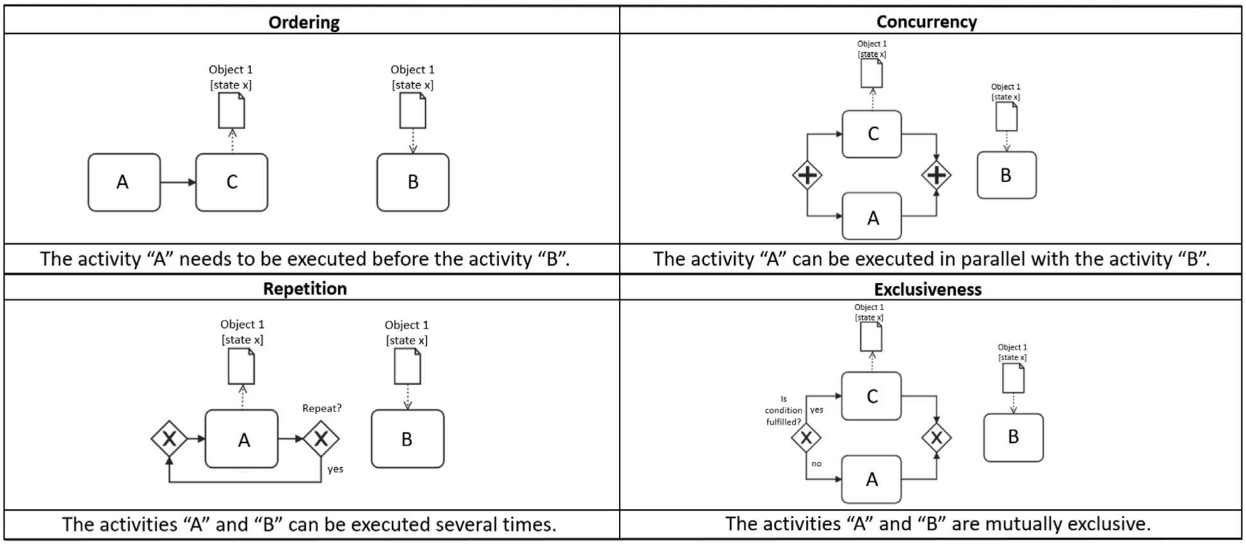


Fig. 6. Four different control-flow patterns and how they are modeled for global tasks based on the fragment-based modeling approach.

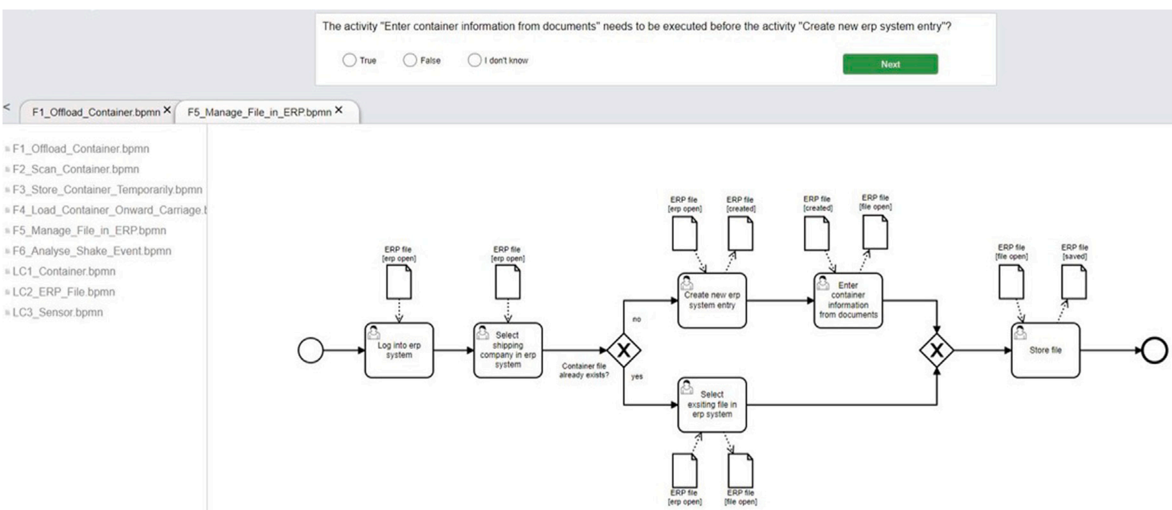


Fig. 7. User interface with an example of an experiment question.

model fragments, were repeated in all the tasks, the participants were instructed to avoid developing an overreaching understanding of the process (including its six fragments and the three life-cycles) from the first task but rather read every task separately, look for the relevant activities and then try to understand how they relate to each other. By doing so, we have also promoted the mental list stopping rule during the search phase (cf. Section 3.5). Consequently, we expect that the participants would first identify all the relevant activities during the search phase, then proceed to infer their relationships during the inference phase. At the data collection, the participants were given a series of tasks displayed in a randomized order. Following each task, the participants were asked to justify their answers and fill out a self-assessment questionnaire of perceived difficulty. For the current study, only the eye-tracking data collected during the task execution, the task solutions, and the self-assessments are used. The verbal explanations will be considered for future research.

4.3. Data analysis

The data analysis encompasses the coarse-grained analysis at a task level, as well as the fine-grained analysis at a phase level. The multi-level analysis is documented through the Python notebooks available online.[5]

4.3.1. Coarse-grained analysis at task level

For the coarse-grained analysis, the comprehension and cognitive load measures introduced in Section 3.4 are calculated at task level. From 46 participants, we obtained 184 data points per factor level (i.e., local tasks or global tasks). To avoid interdependence between the data points coming from each individual, a mean value was calculated for the four tasks capturing each factor level for each participant. This resulted in 46 paired data samples. Due to technical issues with the eye-tracker, the data was further reduced to 44 paired data samples for the fixation-based measures and to 43 paired data samples for the pupil-based ones. In the first two cases the brightness in the room was too high, such that the eye-tracker could not detect the participants' gazes correctly. In the third case, the participant kept moving their head which affected the measurement of their pupil dilation. The remaining data was used to compute the descriptive and inferential

[5] https://github.com/promilab/InfoSysTaskType.git

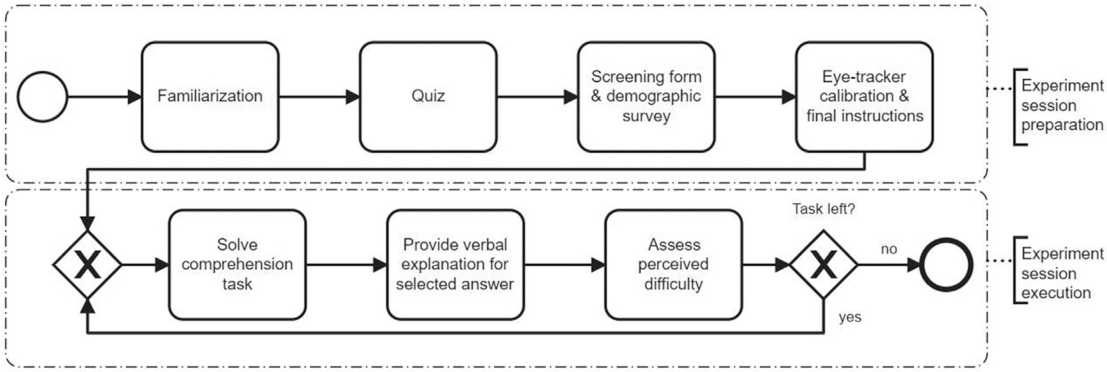


Fig. 8. The different steps during each experiment session.

Table 1
Demographic data of the participants.

Demographics	Groups	Count	Percentage (%)
Education	Business Informatics	15	0.33
	Cognitive Science	1	0.02
	Computer Science	25	0.54
	Industrial Engineering	4	0.09
	Mechatronics	1	0.02
Profession	Research	22	0.48
	Student	17	0.37
	Industry	5	0.11
	IT-Admin	2	0.04
Experience with BPM	0 month	18	0.39
	≤ 1 year	12	0.26
	≤ 2 years	2	0.04
	≤ 5 years	7	0.15
	> 5 years	7	0.15
Familiarity with BPMN	1 (strongly disagree)	9	0.2
	2	7	0.15
	3	3	0.07
	4 (neutral)	5	0.11
	5	8	0.17
	6	9	0.2
	7 (strongly agree)	5	0.11

statistics to investigate our hypotheses (cf. Table 2). We used the non-parametric Wilcoxon Signed-Rank test (single-tailed) for the inferential statistics since it is adequate for comparing paired data samples and does not require the data to be normally distributed. Additionally, we calculated the effect size estimates r^2 for the Wilcoxon Signed-Rank test as proposed by Fritz et al. [72]. The effect size r^2 indicates how much of the variance in the observed dependent variable is explained by the independent variable, i.e., the locality of the task.

4.3.2. *Fine-grained analysis at phase level*

The fine-grained analysis involves three steps. First, we provide a validation of our proposed phase segmentation approach, which allows to separate search from inference phases in our experiment. Second, we conduct a statistical analysis to show that the two phases show different characteristics in terms of cognitive complexity, when comparing local and global tasks. Finally, we conduct a qualitative analysis based on process maps and time series, to support the statistical findings.

Phase separation and validation. The approach followed to infer users' behavior during search and inference phases is summarized in Fig. 9. It shows how the eye-tracking data is segmented into the two phases, followed by a validation of the segmentation. ①At first, the eye-tracking data, collected from the participants, is grouped into trials. Each trial, in turn, refers to the data of an individual participant conducting a single task (e.g., participant P1 conducting task T1, participant P2 conducting task T1). ②In the next step, a cut-mark is assigned to each trial at the point in time where the participant fixated the two relevant activities for solving the task. Hence, the search phase ends and the inference phase starts exactly after the second relevant activity has been fixated (cf. Fig. 4). ③Accordingly, the trial is divided into two phases: a phase prior to the cut-mark (i.e., Phase 1) and a phase after the cut-mark (i.e., Phase 2). Fig. 4 provides an example of the separation of the phases based on the fixations observed during the execution of a local task.

After the segmentation, we conduct an analysis, on whether the phases show indeed differences in terms of cognitive processing. ④For this reason several eye-tracking measures are associated with search and inference behavior, as introduced in Section 3.5, are calculated at the level of each phase. As shown in Fig. 9, when comparing Phase 1 and Phase 2, the participants exhibited shorter average fixation duration, smaller average saccade amplitude, smaller scan-path precision, larger proportion of short fixations ($<$ 250 ms), and smaller proportion of long fixations ($\geq$ 500 ms) in Phase 1 than in Phase 2. ⑤Following the results of the Wilcoxon paired test (i.e., used for pairwise data comparison with no strict assumptions of the normality of the data), all the differences between the measures in Phase 1 and Phase 2 are statistically significant. Considering the trends of these eye-tracking measures, which align with the theoretical underpinnings presented in Section 3.5, and the way the participants were instructed (cf. Section 4.2), Phase 1 is plausibly reflecting search behavior while Phase 2 is plausibly reflecting inference behavior.

Hypotheses Testing. For the fine-grained analysis, the visual behavior measures introduced in Section 3.4 are calculated at phase level. The procedure to select the data for the calculation of the descriptive and inferential statistics at the phase level is identical to the analysis at the task level, resulting in 44 paired data samples to calculate the respective phase duration and entropy for local and global tasks. This time we only needed to remove the two cases from the data, where the gazes were not detected correctly. Again we use the non-parametric Wilcoxon Signed-Rank test (single-tailed) and the effect size estimates r^2 for the inferential statistics.

Process Maps and Time Series. In addition to the hypotheses testing we also apply a qualitative analysis at the phase level to gain more detailed insights on the temporal evolution of the information processing during task execution. The qualitative analysis encompasses process maps and time series graphs, which we derive from a representative participant and task. The process maps are discovered based on PM4Py [73]. To make the structural properties of the process maps more visible the fixations on the different AOIs (represented as rectangles) are encoded numerically. The time series diagrams are derived based on two metrics: first-time visits and entropy. First-time visits track the number of fixations landing on a new AOI, which has not been visited before. Every time a new AOI is fixated, the first-time visits value increases by one. Otherwise, it remains constant. To measure the temporal evolution of entropy we apply a windowing approach [74] with a window size of 30 (fixations). This means the entropy value is calculated after 30 fixations occurred and gets continuously updated each time a new one occurs, while the window moves one step forward, i.e., the window size always remains the same. As a time scale for first-time visits and entropy, we use the number of fixations over time.

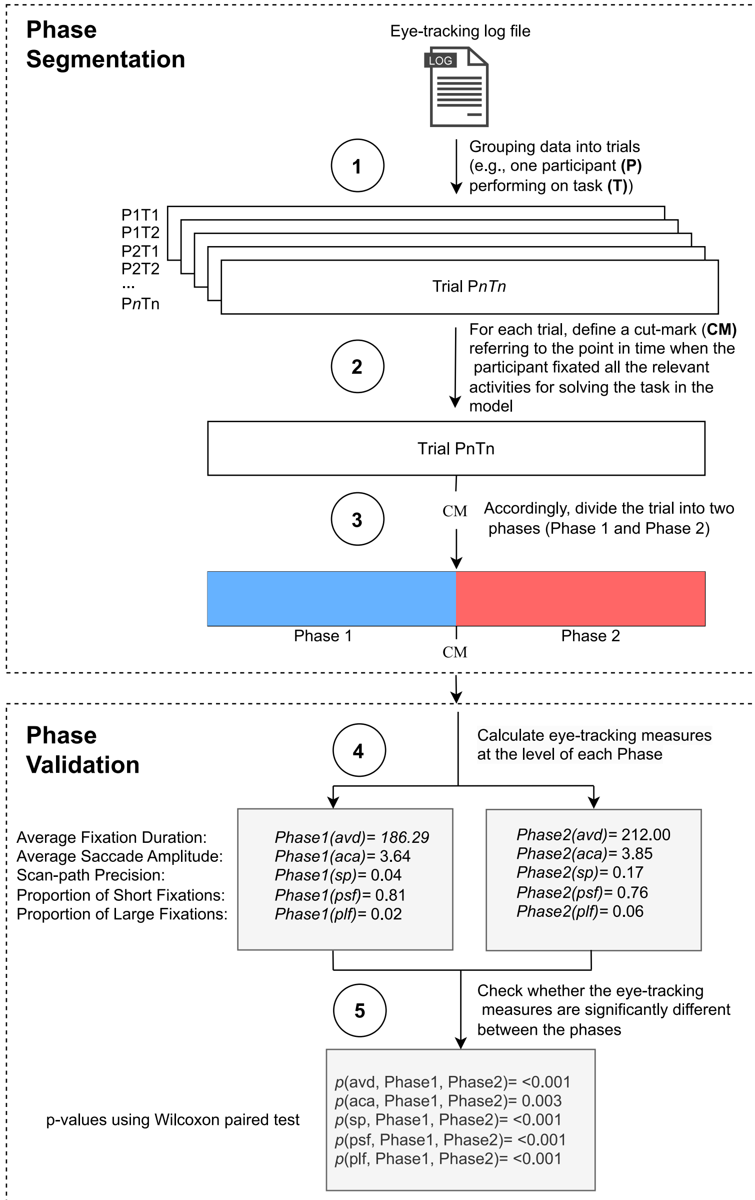


Fig. 9. Approach for the segmentation and validation of the search and inference phases.

Table 2
Descriptive and inferential statistics related to the coarse-grained analysis investigating the impact of local and global tasks on users' performance and cognitive load. N: number of observations, M: calculated mean, SD: calculated standard deviation. Units: Phase duration in seconds, Average fixation duration in milliseconds. A *p-value* $< .05$ means that the pairwise comparison results are significant, r^2 thereby indicates the effect size of the independent variable.

Hypothesis/Construct	Measure	Descriptive			Inferential	
		N	Local M (SD)	Global M (SD)	p-value	r^2
H_1/Comprehension	Accuracy	46	0.978 (0.071)	0.739 (0.247)	< .001	0.575
	Efficiency	46	54.097 (16.829)	117.593 (41.5)	< .001	0.730
H_2/Cognitive load	Perceived difficulty	46	0.62 (0.521)	1.891 (0.772)	< .001	0.754
	Avg. fix. duration	44	194.323 (25.526)	200.966 (25.421)	< .001	0.290
	LHIPA	43	1.185 (0.298)	0.815 (0.216)	< .001	0.668

5. Findings

This section presents the findings of our empirical study. Section 5.1 reports the results of the coarse-grained analysis addressing RQ1. Section 5.2 reports the results of the fine-grained analysis addressing RQ2 (cf. Section 1).

5.1. Coarse-grained analysis at task level

This section presents the results of the coarse-grained analysis investigating the impact of local and global tasks on users' comprehension and cognitive load (addressing RQ1).

Based on the descriptive statistics shown in Table 2, in terms of users' comprehension, *comprehension accuracy* (measured in the range [0:incorrect, 1:correct]) was significantly (cf. p–value in Table 2) lower

Table 3
Descriptive and inferential statistics on the fine-grained analysis investigating how users' search and inference phases differ between local and global tasks in terms of phase duration, cognitive load and scan-path variability. N: number of observations (cf. Section 4.3), M: calculated mean, SD: calculated standard deviation. Units: Phase duration in seconds, Average fixation duration in milliseconds. A *p-value*$< .05$ means that the pairwise comparison results are significant, r^2 thereby indicates the effect size of the independent variable.

Hypothesis/Construct	Measure	Descriptive			Inferential	
		N	Local M (SD)	Global M (SD)	*p*-value	r^2
H_3/Search	Search phase duration	44	28.075 (21.918)	45.084 (24.983)	< .001	0.607
	Entropy	44	3.94 (0.686)	4.65 (0.473)	< .001	0.567
H_4/Inference	Inference phase duration	44	24.856 (19.636)	77.837 (59.236)	< .001	0.755
	Entropy	44	3.154 (0.446)	4.709 (0.626)	< .001	0.731

for global tasks than for local tasks. Similarly, *comprehension efficiency* (measured in seconds) was significantly lower for global than for local tasks. As for the cognitive load, the subjective assessment of *perceived difficulty* (measured in the range [0: "very easy", 4: "very difficult"]) was significantly higher for global tasks compared to local tasks. Similarly, the *average fixation duration* (measured in milliseconds) was significantly higher for global compared to local tasks. Finally, the *LHIPA* score was significantly lower for global than for local tasks. These order relations between the task types are consistent among a vast majority of the participants, independent from their demographic background (an overview on the number of expected and deviating order relations can be found in Appendix B).

Based on the background presented in Section 3.4, our findings support the two formulated hypotheses H_1 and H_2, indicating that global tasks yield lower task comprehension (H_1) and higher cognitive load (H_2) than local tasks, therefore indicating a significant effect of the task locality on comprehension and cognitive load at a task level.

5.2. Fine-grained analysis at phase level

This section reports the results of the fine-grained analysis investigating how users' search and inference behavior differ between local and global tasks in terms of phase duration and entropy (addressing RQ2).

The results in Table 3 summarize the comparison of search and inference behavior between local and global tasks. For the *search* phase, the phase duration (measured in seconds) in global tasks was significantly longer for global tasks than for local tasks. Concerning users' scan-path variability, the entropy measure was higher in global tasks than in local tasks. These order relations between the task types are consistent among a vast majority of the participants, independent from their demographic background (an overview on the number of expected and deviating order relations can be found in Appendix B).

For the *inference* phase, the phase duration in global tasks is also significantly higher than in local tasks. When it comes to scan-path variability, the entropy measure was higher for global tasks than for local ones.

Based on the background presented in Section 3.4, our findings support the two formulated hypotheses H_3 and H_4, indicating that global tasks yield more complex search behavior (H_3) and more complex inference behavior (H_4) than local tasks, therefore indicating a significant effect of the task locality on search and inference complexity at phase level. These findings are further discussed in Section 6.

To complement the inferential statistics, we also conduct some qualitative analysis based on process maps and time series. The process maps show the differences in the scan-paths of a reader, during search (Fig. 10) and inference phase (Fig. 11), when solving a local and a global task. The circles depict the process start and process end, and rectangles refer to the visited AOIs within the process model, following a numerical encoding. The edges indicate the respective transitions between the AOIs. The number in brackets within the rectangles, as well as the one next to the transitions, respectively indicate the frequency of the visits to the AOI and of the transitions, which is also supported by the coloring of the rectangles and the edge thickness. For both phases the process maps show that in the local task the scan-path involves a lower number of visits to process model tasks, as well as less intricate reading patterns, when compared to the global task.

Similarly, Fig. 12 depicts differences between the search and inference phases for the local and global task type, when comparing the evolution of AOI first-time visits and entropy relative to the total number of fixations over time. The comparison of the phases of the local task (left diagrams) with the ones of the global task (right diagrams), shows that search and inference in the global task require a higher number of overall fixations, as well as a higher number of AOI first-time visits. Thus, indicating that both phases involve higher cognitive complexity in global tasks, similar to the findings of the statistical analysis. One can further observe differences between the search and inference phases in terms of increase in the number of first-time visits and change in entropy, relative to the overall number of fixations over time. The entropy seems to particularly decrease after the search phase terminates, therefore indicating behavioral changes. These observations also support the statistical analysis conducted in Section 4.3.2. Hence, providing additional validation, that the two distinct phases can be associated with search and inference. Finally, one can also observe alternating short intervals of high and low entropy during the search, and during the inference phases. This might indicate additional distinctive behavior during these phases, such as information validation or reconciliation [32].

6. Discussion

In this section we discuss the results derived from the coarse-grained as well as fine-grained analysis, followed by several implications from a research and application perspective. Finally, we discuss the main limitations of this study concerning its internal and external validity.

Drawing on the findings from the coarse-grained analysis detailed in Section 5, our study shows that the nature of the task significantly influences users' comprehension and cognitive load when interacting with horizontally modularized process models. This substantiates prevailing hypotheses in the literature [2,9], suggesting that global tasks, impeded by fragmentation, require more cognitive load than local tasks, which benefit from abstraction. Since horizontal modularization facilitates a highly adaptable decomposition of process models into modules, independent of hierarchical dependencies and cross-cutting concerns, it is reasonable to extrapolate that the findings are generally applicable to other modularization approaches, even though they involve more stringent decomposition constraints.

Our fine-grained analysis based on search and inference phases further suggests, that there also exist significant differences regarding phase duration and scan-path entropy when comparing local and global tasks. Inference phases in global tasks seem to particularly increase in complexity due to fragmentation. This is indicated by the entropy, i.e., the uncertainty regarding the occurrence of a fixation (cf. 3.5), which is higher for the inference phase compared to the search phase, when solving global tasks, while for local tasks it is the opposite. Additionally, the inference phase duration is three times higher in global tasks, when compared to local ones.

The gained insights carry several implications for future research and for process comprehension in practice. First, we suggest that future

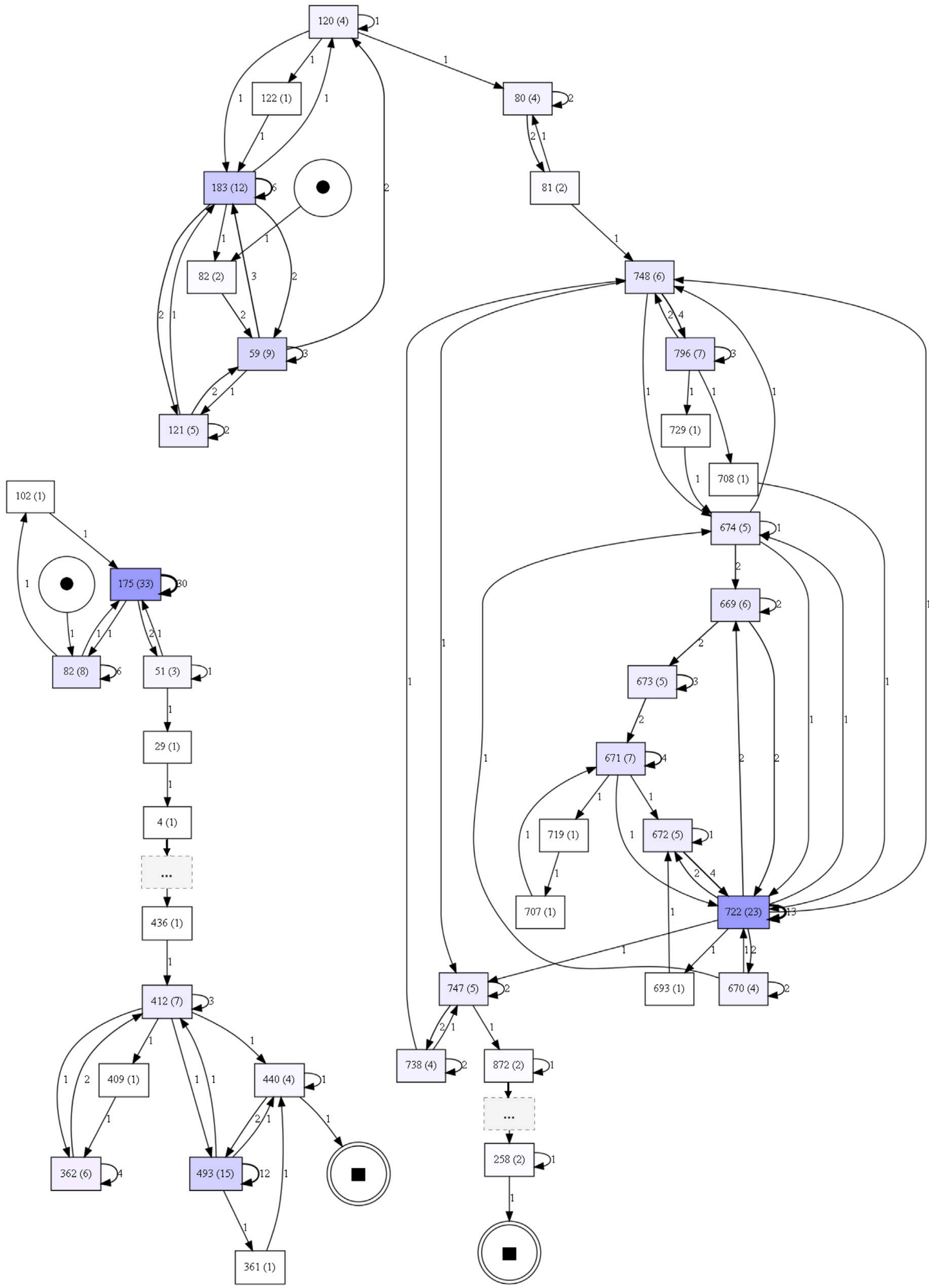


Fig. 10. Two process maps depicting the search phase for a local task (left) and for a global task (right). To improve readability, a similar proportion of visits to AOIs are omitted in both maps, indicated by three dots. The complete maps are available at https://github.com/promilab/InfoSysTaskType.git.

research should consider task type, as well as information processing phases as essential components for the analysis of process model comprehension. While the relationship between task type and cognitive processing has already been investigated in other research areas [15–17], this is not the case in the literature on process model comprehension. The potential for this kind of research was already identified by Mandelburger and Mendling [6] based on their proposition for a theoretical cognitive framework of task performance with diagrams. Our

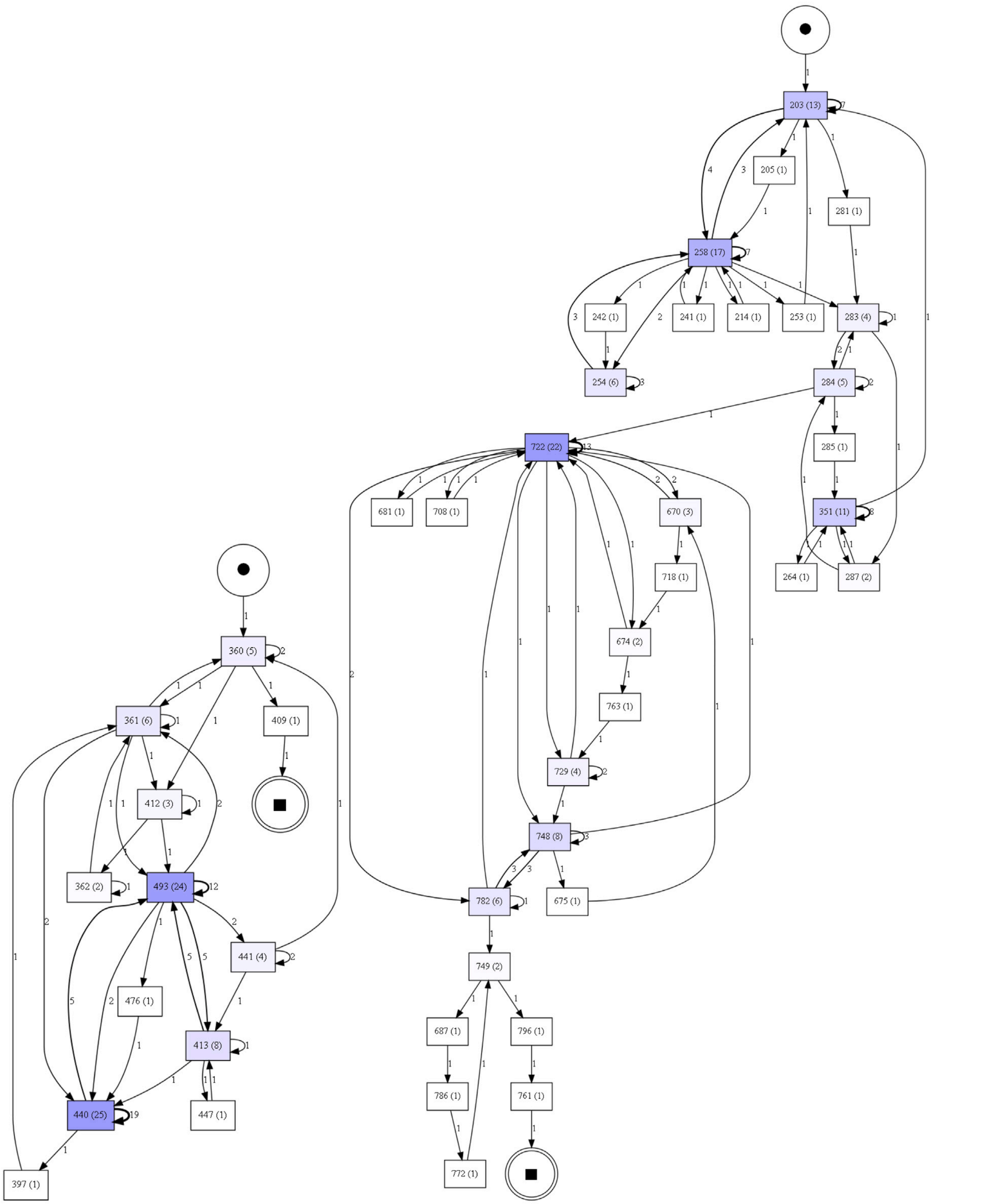


Fig. 11. Two process maps depicting the inference phase for a local task (left) and for a global task (right). A better resolution is available at https://github.com/promilab/InfoSysTaskType.git.

coarse-grained and fine-grained analysis provides some initial empirical work in this direction.

For future empirical work we present a novel research method to investigate information processing during process model comprehension. In Section 4.3.2 we could show that based on our experiment design it is possible to separate information search from information inference phases, thereby allowing for a more fine-grained analysis of the cognitive aspects in process model comprehension. The proposed method could, for example, be used in future research to further investigate the variability of scan-path entropy during both search and inference phases, as shown by the time series in Fig. 12. Here, a promising approach would be to extend the eye-tracking analysis by the think-out-loud method [75] to make the cognitive processes during the phases even more explicit.

From a more practical point of view, our gained insights could help to better evaluate the difficulties an individual is facing when solving a

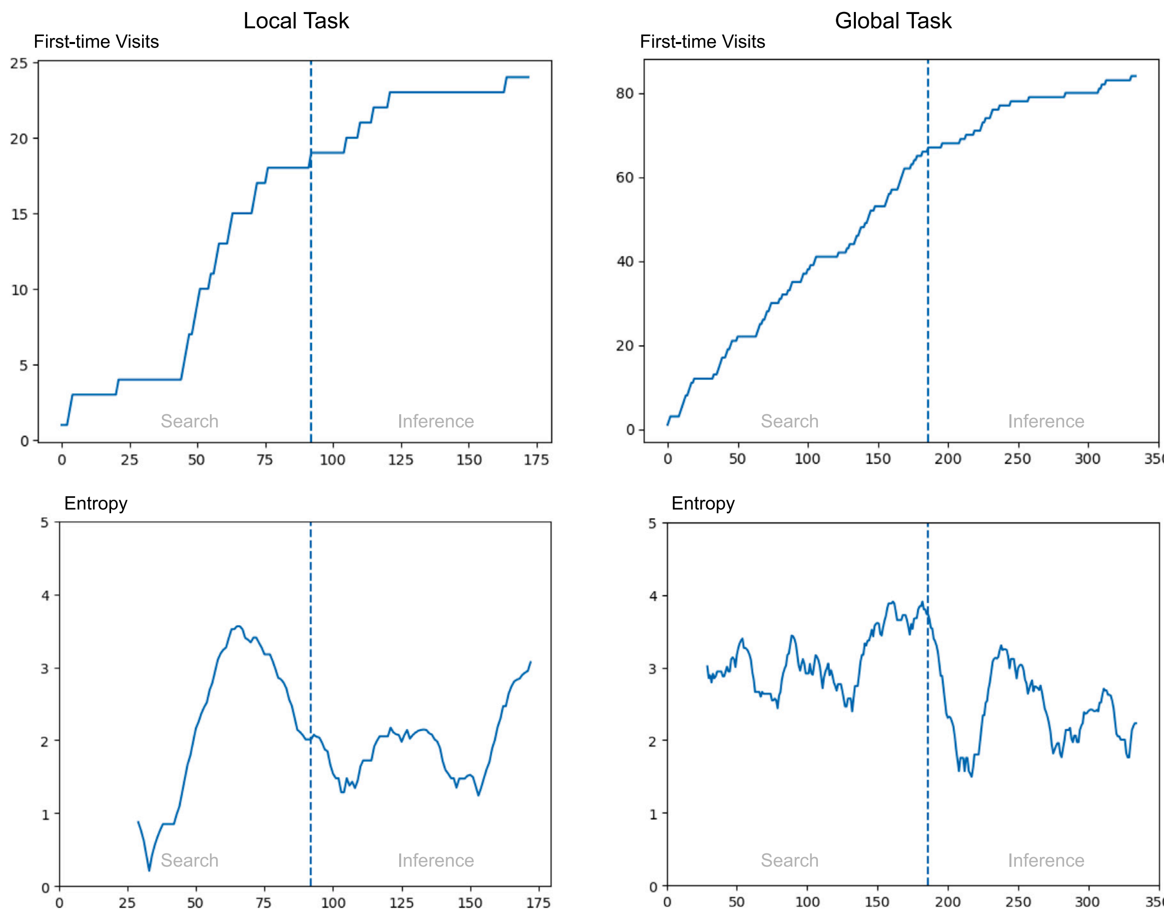


Fig. 12. Qualitative analysis based on time series. The two diagrams on the left side depict the evolution of the first-time visits and the entropy for a selected participant solving a local task, the two diagrams on the right side depict the same metrics for a selected participant solving a global task. The time scale for all diagrams are the total number of fixations over time.

task, i.e., whether the difficulty is related to finding relevant information, inferring the right conclusions, or both. Based on eye-tracking it is possible to provide customized and real-time support, e.g., by providing additional information when an individual faces comprehension issues at a phase level. This could for example be helpful in an e-learning setting [76], teaching about modularized models. Similarly, such support could be provided to practitioners working with modularized models, e.g., in software development [7], or in logistics [8].

The identification of task-specific model comprehension issues allows to support a user in two different ways. The comprehension task itself can be adjusted in such a way that it becomes easier to solve, e.g., by splitting a large (global) task covering multiple modules into smaller (local) tasks. While this approach is not always necessarily applicable in practice, especially in the case of cross-cutting concerns, another approach to facilitate comprehension is to provide better user support in process modeling tools. Global tasks could for example be better supported by interlinking modules to enhance navigation. Even though, empirical evidence is missing to which extent the interlinking supports model comprehension [77]. Alternatively, information search and inference across multiple modules can also be supported based on chatbots, which are capable of communicating dependencies between different parts of a model upon request [78]. In a similar way it can already help to provide a comprehensive overview of modules, i.e., model landscape, to support process model comprehension [79]. Finally, a further possibility to omit comprehension issues based on fragmentation, is the application of simulation to analyze interdependencies between modules [80]. Based on simulation it is for example possible to test the executability of a process model, or the reachability of different process modules, while potentially avoiding the need for information search and inference [2].

Threats to Validity. The following potential threats to the validity of our work should be considered. While all experiment sessions are conducted according to a thoroughly defined protocol and in a controlled environment, one cannot entirely rule out the existence of confounding factors. Several measures are taken to ensure the internal validity of our study. To avoid learning effects, each participant received the different comprehension tasks in a different order. Additionally, all tasks refer to a unique combination of two activities, which are equally distributed among six fragments. In this way, each of the fragments is only relevant for two out of eight tasks. Furthermore, the relatively high number of participants (46) further strengthens the validity of our study. To avoid misinterpretations of the depicted process models and to ensure a similar basic knowledge among the participants, everybody received a uniform introduction to BPMN with a detailed explanation of the fragment-based modeling approach.

To avoid any confounding factors regarding the experimental setting, all local and global tasks were phrased in an identical manner, always referring to exactly two activities, and covering the same set of control-flow patterns (cf. Section 4.1). Additionally, each task referred to the same modularized process model, consisting of six equally complex fragments (cf. Appendix A). However, there remains some risk, that the fragments contain some unforeseen cognitive complexity for the participants, which would bias the comparison of the task types. Besides the thorough design of the process model and the tasks, this risk is additionally mitigated due to the high number of collected data points (184) per factor level (cf. Section 4.1).

Regarding our applied segmentation method to distinguish between search and inference, there remains some uncertainty, whether a participant repeats to search for information after conducting some inference.

This risk is mitigated by the implied stopping rule in the comprehension tasks of our experiment, i.e., the instruction to search for the two stated activities in each task and then to infer their relationship. Furthermore, our statistical analysis in Section 4.3.2 provides comprehensive validation of the segmentation. The statistical analysis based on average fixation duration, saccade amplitude, scan-path precision, and proportion of short and long fixations, provides strong evidence that the search and inference phases are clearly distinct from one another. Thereby leading to the conclusion that the participants indeed followed the stopping rule.

The validity of the applied constructs is supported by existing research studies (cf. Section 3). Comprehension accuracy and comprehension efficiency are widely used to capture comprehension [1]. Also, perceived difficulty, average fixation duration, and LHIPA are well-established measures to capture cognitive load [54,56]. Furthermore, also the applied eye-tracking measures to investigate search and inference are well established in the literature to support these concepts [36,56,65]. Applying all these different measures helps in particularly to avoid a mono-method bias. Our results are further supported by the qualitative data analysis based on the process maps and the depicted time series (cf. Section 5.2). All our observations are in line with similar research experiments [2,9,38], even though these experiments do not provide such detailed insights on the effects of abstraction and fragmentation as in our study.

Finally, there remains some uncertainty, regarding the external validity of our findings. To ensure the applicability of the findings to other process modeling languages, we addressed different control-flow patterns in the comprehension tasks to avoid the dependency on a particular model structure (cf. Fig. 6). These control-flow patterns are commonly found in process models, independent of the modeling language.

It can further be questioned, to which extent the designed comprehension tasks in our research setting resemble comprehension tasks in the real world. There are three main aspects to consider in this regard. Firstly, the contained control-flow patterns in the comprehension questions are commonly used in process models [69]. Secondly, local and global tasks are also commonly encountered in practice [7,8]. Finally, the implied stopping rules in the comprehension tasks imitate search strategies, which might also occur in a different context [61,62]. Even though, in a more complex task setting it might not always be the case, that a clear distinction between search and inference is possible. Although we are not aware of any studies, which have investigated search strategies in the context of process model comprehension.

Furthermore, the heterogeneity of the participants in our study, regarding their process model experience (cf. Section 4.1.3), indicates that our results are somewhat valid, independently from a person's background. To ensure that every participant had the same minimal necessary knowledge of the fragment-based modeling approach, as well as the relevant BPMN components, everyone first completed a familiarization phase and a quiz before the experiment (cf. Section 4.2). In this way, we could ensure, that every participant was able to solve the comprehension tasks, independent from the personal background. While not all participants were familiar with process models, all of them had an educational background in information systems. This reflects well the target group of our study since process models are commonly applied to represent information systems [1–3]. Furthermore, the experiment follows a within-subject design, which implies that every participant is equally exposed to every treatment, providing additional robustness toward heterogeneous subjects.

7. Conclusion and future work

This study provides an analysis of the effects of abstraction and fragmentation in modularized process models at task (coarse-grained analysis) and at phase (fine-grained analysis) levels. At both levels, the study confirms the negative impact of fragmentation in modularized process models. This results in lower task comprehension, higher cognitive load, as well as more complex search and inference phases in global tasks, when compared to local ones. To gain these insights we proposed a novel research method, which allows to separate search from inference phases. We additionally conducted an extensive validation of this method, providing future research opportunities.

Based on the gained insights on process model comprehension, we plan to extend our research at the coarse-grained, as well as the fine-grained level. One important aspect, we have not yet considered at the coarse-grained level, is the effect of process model complexity. In this regard, it will be interesting to analyze how the size and connectedness [81] of a process model influence the effects of abstraction and fragmentation. Generally, we would expect that search and inference become more complex and time-consuming with increasing size and connectedness, while global tasks are more negatively affected than local tasks. Although this remains to be investigated.

In a similar vein, we would further like to investigate, how different process model perspectives might be affected by abstraction and fragmentation. In the context of fragment-based modeling, it would be especially interesting to compare our results with comprehension questions regarding the data perspective.

A further research direction, we intend to follow, is the exploration of the information processing during process model comprehension, focusing on even more fine-grained comprehension phases. The gained insights can further be used to develop predictive models based on machine learning, allowing for online support during process model comprehension. Lastly, we can additionally extend our fine-grained analysis by also considering differences in terms of control-flow patterns.

CRediT authorship contribution statement

Clemens Schreiber: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Conceptualization. **Amine Abbad-Andaloussi:** Writing – review & editing, Supervision, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Barbara Weber:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Clemens Schreiber reports financial support was provided by Karlsruhe Institute of Technology. Amine Abbad-Andaloussi reports financial support was provided by University of St Gallen. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

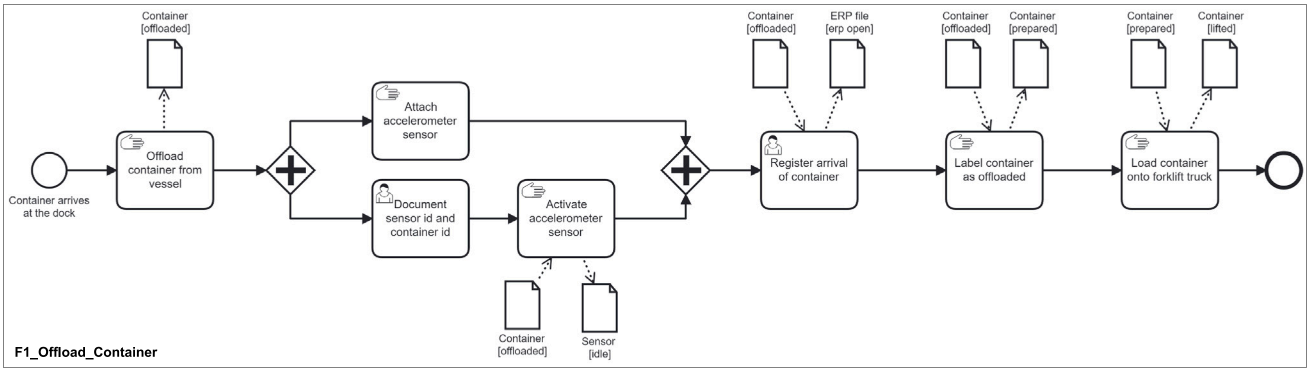
Declaration of Generative AI and AI-assisted technologies in the writing process

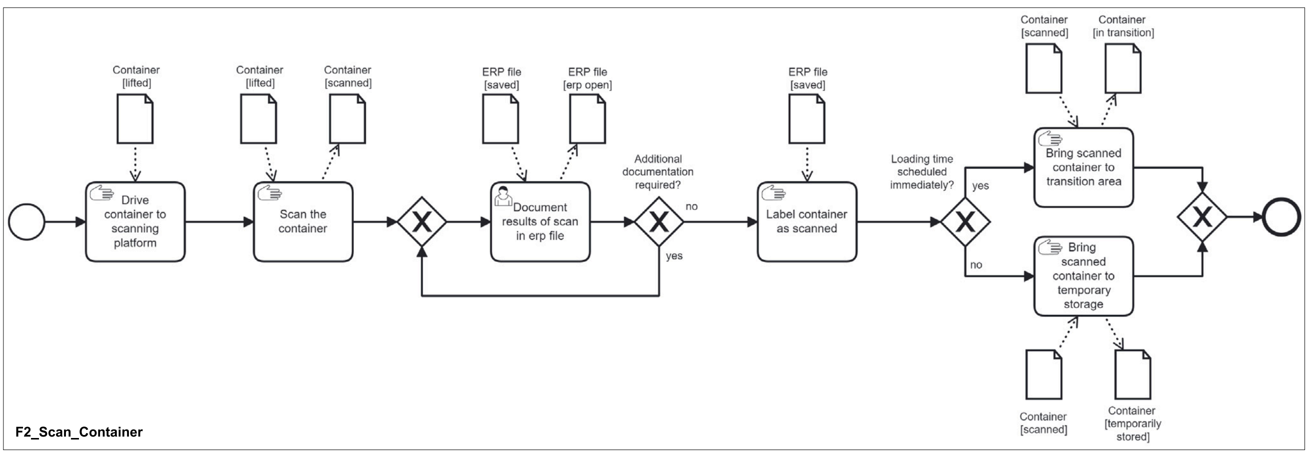
During the preparation of this work the authors used ChatGPT 3.5 in order to improve formulations. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

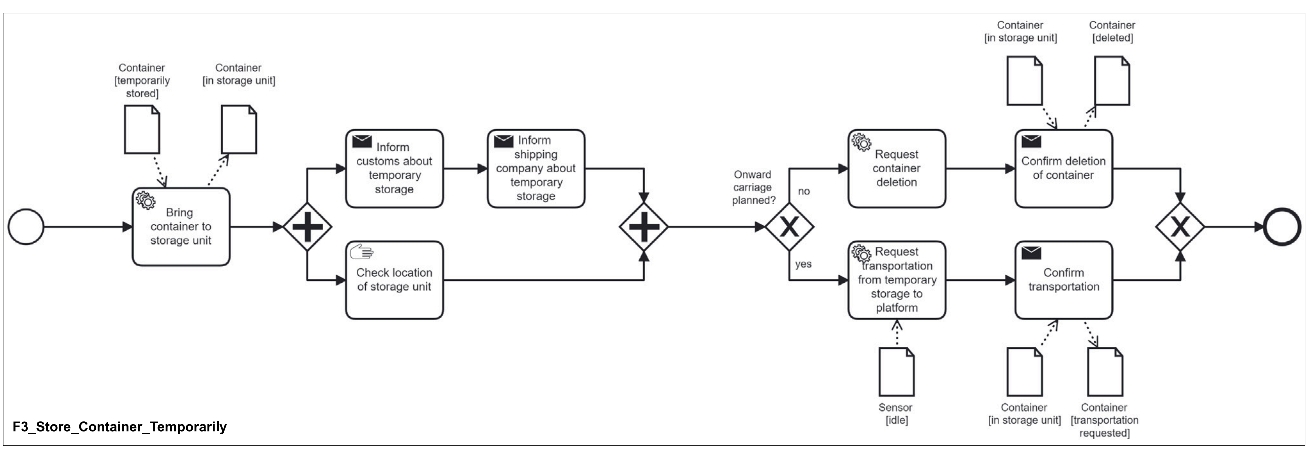
Acknowledgments

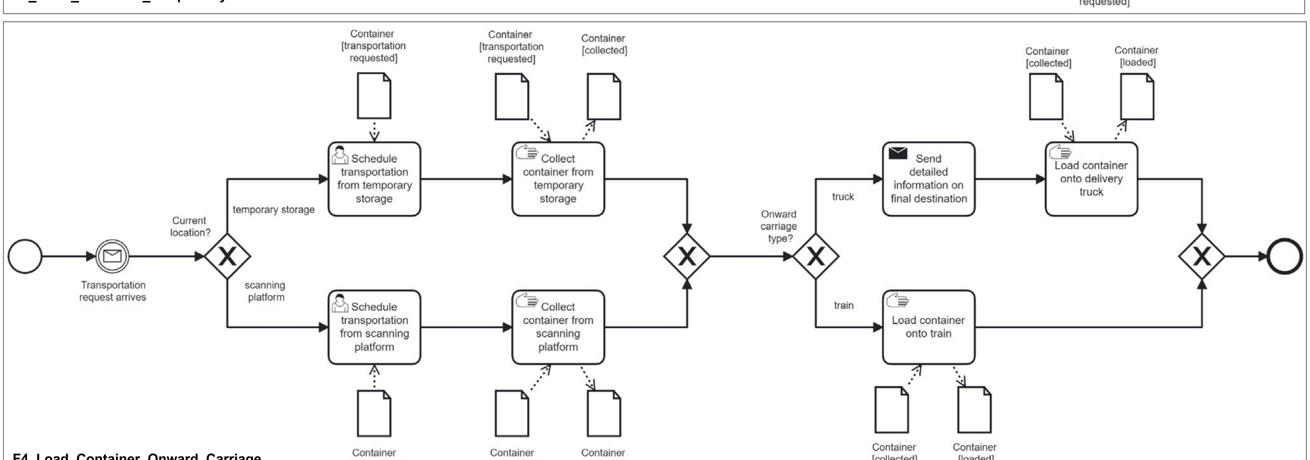
First author supported by the Karlsruhe House of Young Scientists Research Travel Grant, Germany. Second author supported by the International Postdoctoral Fellowship (IPF) Grant (Number: 1031574) from the University of St. Gallen, Switzerland.

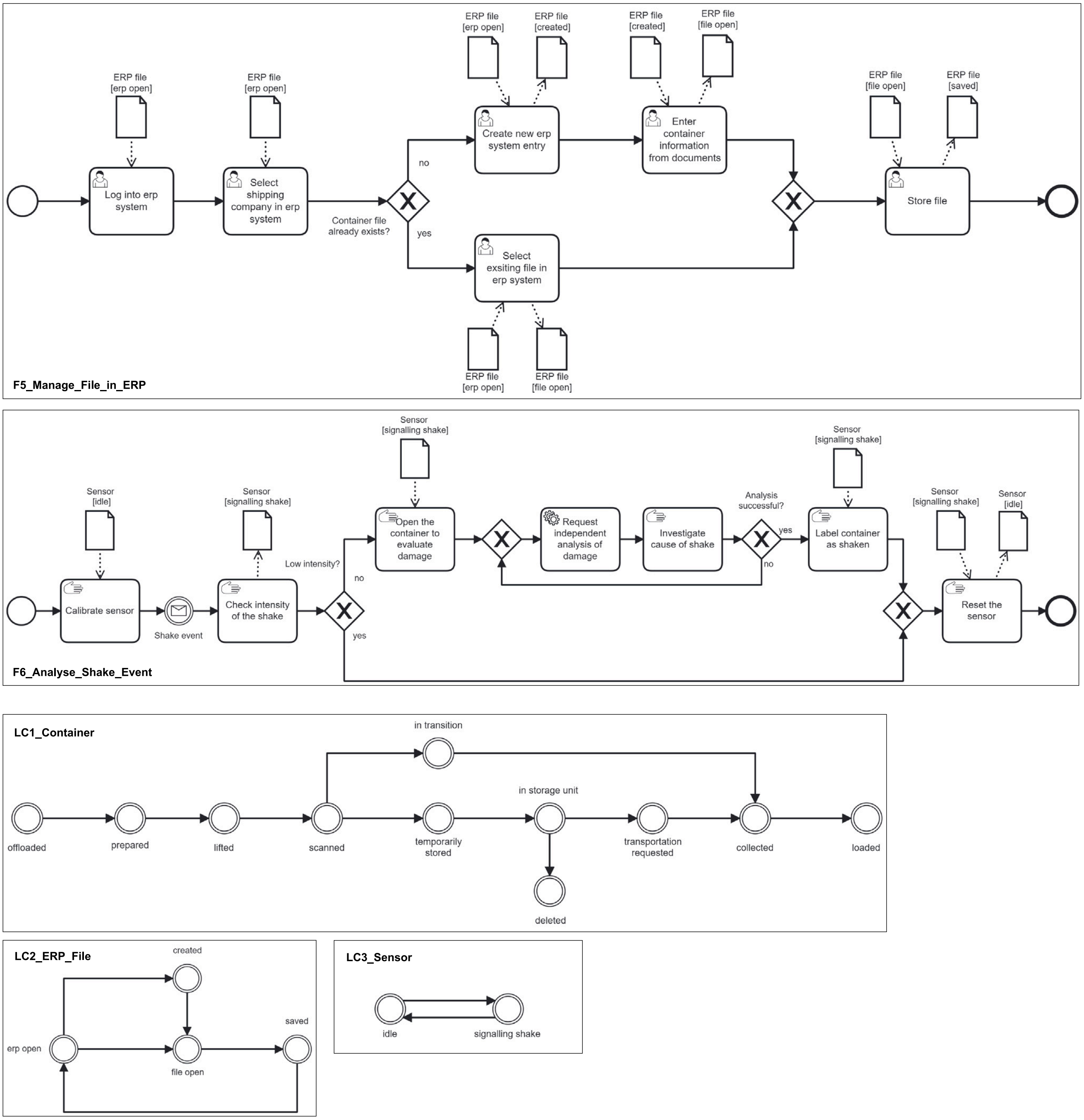
Appendix A. Experiment material: process model and comprehension tasks










F5_Manage_File_in_ERP
ERP file [erp open]
Log into erp system
ERP file [erp open]
Select shipping company in erp system
Container file already exists?
no
yes
ERP file [erp open]
ERP file [created]
Create new erp system entry
ERP file [created]
ERP file [file open]
Enter container information from documents
Select exsiting file in erp system
ERP file [erp open]
ERP file [file open]
ERP file [file open]
ERP file [saved]
Store file
F6_Analyse_Shake_Event
Sensor [idle]
Calibrate sensor
Shake event
Sensor [signalling shake]
Check intensity of the shake
Low intensity?
no
yes
Sensor [signalling shake]
Open the container to evaluate damage
Request independent analysis of damage
Investigate cause of shake
Analysis successful?
yes
no
Sensor [signalling shake]
Label container as shaken
Sensor [signalling shake]
Sensor [idle]
Reset the sensor
LC1_Container
offloaded
prepared
lifted
scanned
in transition
temporarily stored
in storage unit
deleted
transportation requested
collected
loaded
LC2_ERP_File
created
erp open
file open
saved
LC3_Sensor
idle
signalling shake

Type	Control-flow	Task
Local	Ordering	The activity "Enter container information from documents" needs to be executed before the activity "Create new erp system entry"?
	Concurrency	The activity "Inform customs about temporary storage" can be executed in parallel with the activity "Inform shipping company about temporary storage"?
	Exclusiveness	The activities "Send detailed information on final destination" and "Load container onto train" are mutually exclusive?
	Repetition	The activities "Request independent analysis of damage" and "Investigate cause of shake" can be executed several times?
Global	Ordering	The activity "Drive container to scanning platform" needs to be executed before the activity "Attach accelerometer sensor"?
	Concurrency	The activity "Document sensor id and container id" can be executed in parallel with the activity "Calibrate sensor"?
	Exclusiveness	The activities "Request container deletion" and "Schedule transportation from temporary storage" are mutually exclusive?
	Repetition	The activities "Document results of scan in erp file" and "Log into erp system" can be executed several times?

Appendix B. Descriptive statistics on order relations

Table B.4
An overview of the number of participants, where the measures on the left reflect the expected order relation between the local and global tasks, equality between the two task types, or the opposite of the expected order relation. Based on the formulated hypotheses in Section 4.1.1, in comparison to local tasks, global tasks are expected to yield lower accuracy, longer response time, higher perceived difficulty, longer average fixation duration, lower LHIPA, longer search duration, longer inference duration, higher search entropy, and higher inference entropy.

	Expected order relation	Equality order relation	Opposite order relation
Accuracy	29	17	0
Response time	44	0	2
Perceived difficulty	43	3	0
Average fixation duration	33	0	11
LHIPA	38	0	5
Search duration	41	0	3
Inference duration	43	0	1
Search entropy	37	0	7
Inference entropy	43	0	1

References

[1] K. Figl, Comprehension of procedural visual business process models: A literature review, Bus. Inf. Syst. Eng. 59 (2017) 41–67.

[2] S. Zugal, Applying Cognitive Psychology for Improving the Creation, Understanding and Maintenance of Business Process Models (Ph.D. thesis), University of Innsbruck, 2013.

[3] M. Indulska, P. Green, J. Recker, M. Rosemann, Business process modeling: Perceived benefits, in: Conceptual Modeling-ER 2009: 28th International Conference on Conceptual Modeling, Gramado, Brazil, November 9-12, 2009. Proceedings 28, Springer, 2009, pp. 458–471.

[4] W. Van Der Aalst, Process Mining: Data Science in Action, vol. 2, Springer, 2016.

[5] C. van Dun, L. Moder, W. Kratsch, M. Röglinger, ProcessGAN: Supporting the creation of business process improvement ideas through generative machine learning, Decis. Support Syst. 165 (2023) 113880.

[6] M.M. Mandelburger, J. Mendling, Cognitive diagram understanding and task performance in systems analysis and design, MIS Q. 45 (4) (2021) 2101–2157.

[7] O. Manglaras, A. Farkas, P. Fule, C. Treude, M. Wagner, Problems in microservice development: Supporting visualisation, in: 2023 IEEE Working Conference on Software Visualization, VISSOFT, IEEE, 2023, pp. 62–72.

[8] A. Bucchiarone, A. Marconi, M. Pistore, H. Raik, Dynamic adaptation of fragment-based and context-aware business processes, in: 2012 IEEE 19th International Conference on Web Services, IEEE, 2012, pp. 33–41.

[9] H.A. Reijers, J. Mendling, Modularity in process models: Review and effects, in: BPM, vol. 8, Springer, 2008, pp. 20–35.

[10] M. Winter, R. Pryss, T. Probst, J. Baß, M. Reichert, Measuring the cognitive complexity in the comprehension of modular process models, IEEE Trans. Cogn. Dev. Syst. 14 (1) (2022) 164–180.

[11] K. Figl, R. Laue, Influence factors for local comprehensibility of process models, Int. J. Hum.-Comput. Stud. 82 (2015) 96–110.

[12] R. Petrusel, J. Mendling, H.A. Reijers, Task-specific visual cues for improving process model understanding, Inf. Softw. Technol. 79 (2016) 63–78.

[13] P. Bera, P. Soffer, J. Parsons, Using eye tracking to expose cognitive processes in understanding conceptual models, MIS Q. 43 (4) (2019) 1105–1126.

[14] J. Recker, H.A. Reijers, S.G. van de Wouw, Process model comprehension: the effects of cognitive abilities, learning style, and strategy, Commun. Assoc. Inf. Syst. 34 (1) (2014) 9.

[15] P. Vakkari, Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval, Inf. Process. Manag. 35 (6) (1999) 819–837.

[16] A. Haji-Abolhassani, J.J. Clark, A computational model for task inference in visual search, J. Vis. 13 (3) (2013) 29–29.

[17] B. Choi, A. Ward, Y. Li, J. Arguello, R. Capra, The effects of task complexity on the use of different types of information in a search assistance tool, ACM Trans. Inf. Syst. (TOIS) 38 (1) (2019) 1–28.

[18] J.H. Larkin, H.A. Simon, Why a diagram is (sometimes) worth ten thousand words, Cogn. Sci. 11 (1) (1987) 65–100.

[19] J.M. Wolfe, Guided search 2.0 a revised model of visual search, Psychon. Bull. Rev. 1 (1994) 202–238.

[20] C. Schreiber, A. Abbad-Andaloussi, B. Weber, On the cognitive effects of abstraction and fragmentation in modularized process models, in: International Conference on Business Process Management, Springer, 2023, pp. 359–376.

[21] M. Hewelt, M. Weske, A hybrid approach for flexible case modeling and execution, in: Business Process Management Forum: BPM Forum 2016, Rio de Janeiro, Brazil, September 18-22, 2016, Proceedings 14, Springer, 2016, pp. 38–54.

[22] W. Wang, T. Chen, M. Indulska, S. Sadiq, B. Weber, Business process and rule integration approaches—An empirical analysis of model understanding, Inf. Syst. 104 (2022) 101901.

[23] I. Vessey, Cognitive fit: A theory-based analysis of the graphs versus tables literature, Decis. Sci. 22 (2) (1991) 219–240.

[24] C. Dunn, S. Grabski, An investigation of localization as an element of cognitive fit in accounting model representations, Decis. Sci. 32 (1) (2001) 55–94.

[25] I. Vessey, D. Galletta, Cognitive fit: An empirical study of information acquisition, Inf. Syst. Res. 2 (1) (1991) 63–84.

[26] H. Ritchi, M.J. Jans, J. Mendling, H. Reijers, The influence of business process representation on performance of different task types, J. Inf. Syst. (2019).

[27] A. Abbad Andaloussi, P. Soffer, T. Slaats, A. Burattin, B. Weber, The impact of modularization on the understandability of declarative process models: A research model, in: Information Systems and Neuroscience: NeuroIS Retreat 2020, Springer, 2020, pp. 133–144.

[28] P. Soffer, M. Kaner, Y. Wand, Towards understanding the process of process modeling: theoretical and empirical considerations, in: Business Process Management Workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I 9, Springer, 2012, pp. 357–369.
[29] J. Pinggera, P. Soffer, D. Fahland, M. Weidlich, S. Zugal, B. Weber, H.A. Reijers, J. Mendling, Styles in business process modeling: an exploration and a model, Softw. Syst. Model. 14 (2015) 1055–1080.
[30] J. Claes, I. Vanderfeesten, F. Gailly, P. Grefen, G. Poels, The structured process modeling theory (SPMT) a cognitive view on why and how modelers benefit from structuring the process of process modeling, Inf. Syst. Front. 17 (2015) 1401–1425.
[31] B. Weber, J. Pinggera, M. Neurauter, S. Zugal, M. Martini, M. Furtner, P. Sachse, D. Schnitzer, Fixation patterns during process model creation: Initial steps toward neuro-adaptive process modeling environments, in: 2016 49th Hawaii International Conference on System Sciences, HICSS, IEEE, 2016, pp. 600–609.
[32] A. Burattin, M. Kaiser, M. Neurauter, B. Weber, Learning process modeling phases from modeling interactions and eye tracking data, Data Knowl. Eng. 121 (2019) 1–17.
[33] K. Rosenthal, J. Wagner, S. Strecker, Modeling styles in conceptual data modeling: Reflecting observations in a series of multimodal studies, in: ECIS 2022 Research Papers, 2022.
[34] K. Rosenthal, S. Strecker, M. Snoeck, Modeling difficulties in creating conceptual data models: Multimodal studies on individual modeling processes, Softw. Syst. Model. 22 (3) (2023) 1005–1030.
[35] W. Philip, R. Kristina, S. Stefan, Exploring conceptual data modeling processes: Insights from clustering and visualizing modeling sequences, in: Modellierung 2024, Gesellschaft für Informatik eV, 2024, pp. 25–42.
[36] R. Petrusel, J. Mendling, Eye-tracking the factors of process model comprehension tasks, in: C. Salinesi, M.C. Norrie, O. Pastor (Eds.), Advanced Information Systems Engineering, Springer, Berlin, Heidelberg, 2013, pp. 224–239.
[37] J. Kim, J. Hahn, H. Hahn, How do we understand a system with (so) many diagrams? Cognitive integration processes in diagrammatic reasoning, Inf. Syst. Res. 11 (3) (2000) 284–303.
[38] M. Winter, H. Neumann, R. Pryss, T. Probst, M. Reichert, Defining gaze patterns for process model literacy–exploring visual routines in process models with diverse mappings, Expert Syst. Appl. 213 (2023) 119217.
[39] G. Shanks, D. Moody, J. Nuredini, D. Tobin, R. Weber, Representing classes of things and properties in general in conceptual modelling: An empirical evaluation, J. Database Manag. (JDM) 21 (2) (2010) 1–25.
[40] C. Haisjackl, I. Barba, S. Zugal, P. Soffer, I. Hadar, M. Reichert, J. Pinggera, B. Weber, Understanding declare models: strategies, pitfalls, empirical results, Softw. Syst. Model. 15 (2) (2016) 325–352.
[41] A. Abbad Andaloussi, F. Zerbato, A. Burattin, T. Slaats, T.T. Hildebrandt, B. Weber, Exploring how users engage with hybrid process artifacts based on declarative process models: A behavioral analysis based on eye-tracking and think-aloud, Softw. Syst. Model. 20 (2021) 1437–1464.
[42] M. Pesic, H. Schonenberg, W.M. Van Der Aalst, Declare: Full support for loosely-structured processes, in: 11th IEEE International Enterprise Distributed Object Computing Conference, (EDOC 2007), IEEE, 2007, 287–287.
[43] T.T. Hildebrandt, R.R. Mukkamala, Declarative Event-Based Workflow as Distributed Dynamic Condition Response Graphs, Electron. Proc. Theor. Comput. Sci. 69 (2011) 59–73.
[44] O. Turetken, T. Rompen, I. Vanderfeesten, A. Dikici, J. van Moll, The effect of modularity representation and presentation medium on the understandability of business process models in BPMN, in: Business Process Management: 14th International Conference, BPM 2016, Rio de Janeiro, Brazil, September 18-22, 2016. Proceedings 14, Springer, 2016, pp. 289–307.
[45] R. Petrusel, J. Mendling, H.A. Reijers, How visual cognition influences process model comprehension, Decis. Support Syst. 96 (2017) 1–16.
[46] J. Mendling, H.A. Reijers, W.M. Van Der Aalst, Seven process modeling guidelines (7PMG), Inf. Softw. Technol. 52 (2) (2010) 127–136.
[47] J. Sweller, P. Chandler, Why some material is difficult to learn, Cogn. Instr. 12 (3) (1994) 185–233.
[48] D.L. Parnas, On the criteria to be used in decomposing systems into modules, Commun. ACM 15 (12) (1972) 1053–1058.
[49] M. La Rosa, P. Wohed, J. Mendling, A.H. Ter Hofstede, H.A. Reijers, W.M. Van Der Aalst, Managing process model complexity via abstract syntax modifications, IEEE Trans. Ind. Inform. 7 (4) (2011) 614–629.
[50] W.M. Van Der Aalst, M. Weske, D. Grünbauer, Case handling: A new paradigm for business process support, Data Knowl. Eng. 53 (2) (2005) 129–162.
[51] J. Krumeich, B. Weis, D. Werth, P. Loos, Event-driven business process management: where are we now?: A comprehensive synthesis and analysis of literature, Bus. Process Manag. J. 20 (4) (2014) 615–633.
[52] P. Fettke, W. Reisig, Modelling service-oriented systems and cloud services with heraklit, in: Advances in Service-Oriented and Cloud Computing: International Workshops of ESOCC 2020, Heraklion, Crete, Greece, September 28–30, 2020, Revised Selected Papers 8, Springer, 2021, pp. 77–89.
[53] A. Keselman, L. Slaughter, V.L. Patel, Toward a framework for understanding lay public's comprehension of disaster and bioterrorism information, J. Biomed. Inform. 38 (4) (2005) 331–344.
[54] F. Chen, J. Zhou, Y. Wang, K. Yu, S.Z. Arshad, A. Khawaji, D. Conway, Robust Multimodal Cognitive Load Measurement, Springer, 2016.
[55] F. Paas, J.E. Tuovinen, H. Tabbers, P.W. Van Gerven, Cognitive load measurement as a means to advance cognitive load theory, Educ. Psychol. (2003).
[56] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, J. van de Weijer, Eye Tracking: A Comprehensive Guide to Methods and Measures, OUP Oxford, 2011.
[57] B. Weber, T. Fischer, R. Riedl, Brain and autonomic nervous system activity measurement in software engineering: A systematic literature review, J. Syst. Softw. 178 (2021).
[58] J. Sweller, P. Ayres, K. Slava, Cognitive Load Theory, Springer, 2011.
[59] R. Riedl, P. Léger, Fundamentals of neurois: Information systems and the brain, in: Studies in Neuroscience, Psychology and Behavioral Economics, Springer, Berlin, Heidelberg, 2015.
[60] A.T. Duchowski, K. Krejtz, N.A. Gehrer, T. Bafna, P. Bækgaard, The low/high index of pupillary activity, in: CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–12.
[61] G.J. Browne, M.G. Pitts, Stopping rule use during information search in design problems, Organ. Behav. Hum. Decis. Processes 95 (2) (2004) 208–224.
[62] G.J. Browne, M.G. Pitts, J.C. Wetherbe, Cognitive stopping rules for terminating information search in online tasks, MIS Q. (2007) 89–104.
[63] A. Glöckner, A.-K. Herbold, An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes, J. Behav. Decis. Mak. 24 (1) (2011) 71–98.
[64] R.E. Mayer, Human nonadversary problem solving, Hum. Mach. Probl. Solving (1989) 39–56.
[65] A. Glöckner, A.-K. Herbold, Information processing in decisions under risk: Evidence for compensatory strategies based on automatic processes, MPI Collect. Goods Preprint 2008 (42) (2008).
[66] A. Abbad Andaloussi, A. Burattin, T. Slaats, A.C.M. Petersen, T.T. Hildebrandt, B. Weber, Exploring the understandability of a hybrid process design artifact based on DCR graphs, in: Enterprise, Business-Process and Information Systems Modeling: 20th International Conference, BPMDS 2019, 24th International Conference, EMMSAD 2019, Held At CAiSE 2019, Rome, Italy, June 3–4, 2019, Proceedings 20, Springer, 2019, pp. 69–84.
[67] J. Gulden, A. Burattin, A.A. Andaloussi, B. Weber, From analytical purposes to data visualizations: A decision process guided by a conceptual framework and eye tracking, Softw. Syst. Model. 19 (2020) 531–554.
[68] J. Becker, M. Rosemann, C. Von Uthmann, Guidelines of business process modeling, in: Business Process Management: Models, Techniques, and Empirical Studies, Springer, 2002, pp. 30–49.
[69] M.z. Muehlen, J. Recker, How much language is enough? Theoretical and practical use of the business process modeling notation, in: Seminal Contributions to Information Systems Engineering: 25 Years of CAiSE, Springer, 2013, pp. 429–443.
[70] H.A. Reijers, J. Mendling, R.M. Dijkman, Human and automatic modularizations of process models to enhance their comprehension, Inf. Syst. 36 (5) (2011) 881–897.
[71] A. Abbad-Andaloussi, D. Lübke, B. Weber, Conducting eye-tracking studies on large and interactive process models using EyeMind, SoftwareX 24 (2023) 101564.
[72] C.O. Fritz, P.E. Morris, J.J. Richler, Effect size estimates: current use, calculations, and interpretation., Journal of experimental psychology: General 141 (1) (2012) 2.
[73] A. Berti, S. van Zelst, D. Schuster, PM4py: A process mining library for python, Softw. Impacts 17 (2023) 100556.
[74] R.J.C. Bose, W.M. Van Der Aalst, I. Žliobaitė, M. Pechenizkiy, Handling concept drift in process mining, in: Advanced Information Systems Engineering: 23rd International Conference, CAiSE 2011, London, UK, June 20-24, 2011. Proceedings 23, Springer, 2011, pp. 391–405.
[75] G.M. Olson, S.A. Duffy, R.L. Mack, Thinking-out-loud as a method for studying real-time comprehension processes, in: New Methods in Reading Comprehension Research, Routledge, 2018, pp. 253–286.
[76] M. Porta, S. Ricotti, C.J. Perez, Emotional e-learning through eye tracking, in: Proceedings of the 2012 IEEE Global Engineering Education Conference, EDUCON, IEEE, 2012, pp. 1–6.
[77] D. Lübke, M. Ahrens, Towards an experiment for analyzing subprocess navigation in bpmn tooling (short paper), in: ZEUS, 2022, pp. 1–5.
[78] A. López, J. Sànchez-Ferreres, J. Carmona, L. Padró, From process models to chatbots, in: Advanced Information Systems Engineering: 31st International Conference, CAiSE 2019, Rome, Italy, June 3–7, 2019, Proceedings 31, Springer, 2019, pp. 383–398.
[79] F. Gonzalez-Lopez, L. Pufahl, J. Munoz-Gama, V. Herskovic, M. Sepúlveda, Case model landscapes: toward an improved representation of knowledge-intensive processes using the fCM-language, Softw. Syst. Model. (2021).
[80] M. Marquard, M. Shahzad, T. Slaats, Web-based modelling and collaborative simulation of declarative processes, in: Business Process Management: 13th International Conference, BPM 2015, Innsbruck, Austria, August 31–September 3, 2015, Proceedings, Springer, 2015, pp. 209–225.
[81] J. Mendling, Metrics for Process Models: Empirical Foundations of Verification, Error Prediction, and Guidelines for Correctness, vol. 6, Springer Science & Business Media, 2008.