On the Cognitive Effects of Abstraction and Fragmentation in Modularized Process Models*

Clemens Schreiber[1], Amine Abbad-Andaloussi[2], and Barbara Weber[2]

[1] Karlsruhe Institute of Technology, Karlsruhe, Germany
clemens.schreiber@kit.edu
[2] University of St. Gallen, St. Gallen, Switzerland
{amine.abbad-andaloussi,barbara.weber}@unisg.ch

Abstract. Process models support a variety of tasks, which can be organized differently. Notably one can discern local tasks focusing on a single part of a model and global tasks requiring an overview of several parts. These two task types are assumed to affect users' understanding of processes differently especially if the processes are decomposed into many interlinked and self-contained models through modularization. Local tasks can benefit from abstraction as they enable information hiding, while global tasks can be impeded by fragmentation caused by the split attention effect. Following a task-centric approach, we substantiate this hypothesis by investigating the cognitive effects of abstraction and fragmentation in modularization. Therein, we focus particularly on horizontal modularization and study users' cognitive load, comprehension and behavior when solving local and global tasks. Our findings confirm that, compared to abstraction, fragmentation hinders users' comprehension of the model and raises their cognitive load. Additionally, users exhibit different search and integration behaviors when performing local and global tasks. The outcome of this work motivates the shift from artifact-centric to task-centric empirical studies, raises the need for approaches to mitigate the effect of fragmentation and explores different alternatives to achieve this goal.

Keywords: Process model understandability · modularization · cognitive load theory · eye tracking.

1 Introduction

Process models serve as a mediator for digital transformation as they support process enactment but also facilitate the communication between stakeholders and promote change in business processes [31]. Different process modeling initiatives exist to support these different activities (e.g., SAP Signavio Process Transformation Suite[3]). Yet, most of the research and tool development in the

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

[3] See https://www.signavio.com/products/process-transformation-suite/

Business Process Management (BPM) field are focused on the model artifacts themselves (i.e., the graphical representation of processes), with limited interest in the task-perspective [25], i.e., what types of tasks a process model is supposed to support. This perspective is crucial for identifying what kind of information and support are needed in a particular context. In this paper we address this gap by investigating two recurrent tasks encountered in practice: *local* and *global* tasks. The former is typically focused on a local part of the model (e.g., to comprehend or maintain a particular sub-process or model fragment), while the latter covers global aspects requiring an overview of several parts of the model (e.g., cross-cutting concerns spanning over several sub-processes or process fragments). We study these tasks in the light of *horizontal modularization*, i.e., a particular branch of modularization where a system is decomposed into interlinked modules with no strict assumptions on their hierarchy [23,40]. This approach allows modeling flexible and knowledge intensive processes [3,17]. Moreover, it is commonly used to support case-based process execution [3,17] and has shown a positive impact on model comprehension [40].

While one should expect that modularization generally supports model comprehension, existing empirical studies suggest that it does not always make the execution of tasks easier [32,40,43]. Notably, it was hypothesized that modularization enables abstraction as it supports information hiding and pattern recognition, while it causes fragmentation since users are required to split their attention between different fragments to find relevant information [43]. Hence, local tasks, which are constrained to a single module might be supported by abstraction and become easier, whereas global tasks, which refer to several modules might be hindered by fragmentation and become more difficult to perform.

To investigate this hypothesis we conduct an empirical study. Therein, we first confirm the effects of abstraction and fragmentation on the understandability of modularized process models. Secondly, we delve into users' behavior when engaging with modularized models to compare the behavioral traits characterizing the solving of local tasks and global tasks.

Our empirical study is based on eye-tracking. Following the fragment-based modeling approach proposed in [17], we design a horizontally modularized process model and a set of local and global tasks. Then, we test the effects of abstraction and fragmentation on both model comprehension and users' cognitive load. Therein, we use a multi-modal approach covering the typical model comprehension and *subjective* cognitive load measures used in the literature [14,43] – but also advanced *objective* cognitive load measures derived from eye-tracking. Doing so, we provide a multi-perspective empirical account allowing us to draw more robust conclusions on the effects of abstraction and fragmentation. Our findings confirm that the task type is a crucial factor with a significant impact on model comprehension and users' cognitive load. As for the behavioral investigation, we look at eye-tracking measures reflecting the search complexity and cognitive integration effort exhibited by the users when solving local and global tasks. Our analysis shows that global tasks are associated with complex search and higher cognitive integration due to the effect of fragmentation. Based on these insights, we propose several solutions to compensate for this effect and thus support users when solving global tasks on modularized process models.

Overall, this work provides an empirical affirmation for the cognitive effects associated with the solving of global and local tasks. Moreover, our research model demonstrates how to investigate users' comprehension, cognitive load and behavior while reading process models using a wide array of multi-modal measures. From a more applied perspective, the outcome of this work raises the need for incorporating the task type in model comprehension frameworks, developing novel mechanisms to support global tasks and providing new foundations for managerial decisions on task distribution. In the remainder, Sect. 2 provides the background and related work. Sect. 3 describes our empirical study design. Sects. 4 and 5 present and discuss the findings of our study respectively. Sect. 6 concludes the paper.

2 Background and Related Work

This section describes the theoretical underpinnings related to this work from the process modeling (cf. Sects. 2.1-2.3) and cognitive (cf. Sect. 2.4) perspectives.

2.1 Modularization in Process Modeling

Modularization denotes the process of decomposing a system into interlinked modules with self-contained properties [29]. Authors in the process modeling literature [23,40] have discerned three types of modularization: *vertical*, *horizontal* and *orthogonal*. Vertical modularization decomposes the process into sub-processes using a hierarchical structure [23,40]. Horizontal modularization, in turn, aims at partitioning the process into several interconnected fragments with no strict assumptions on their hierarchy [23,40]. Lastly, orthogonal modularization divides the process along cross-cutting concerns (i.e., privacy, security) spanning over several elements of the model [23,40].

Since horizontal modularization allows separating process modules regardless of their hierarchical relationship and cross-cutting concerns, this approach appears to be more general and thus well suited to investigate the effects of abstraction and fragmentation on users' comprehension and cognitive load. While a general positive impact of horizontal modularization can be assumed based on [40], we investigate how horizontal modularization affects comprehension tasks focused on single modules (local) versus comprehension tasks focused on multiple modules (global). We use the fragment-based approach [17] (cf., Sect. 2.2) as a representative for this type of modularization.

2.2 Fragment-based Process Modeling Approach

The fragment-based modeling approach [17] is based on Business Process Modeling and Notation (BPMN). It aims at supporting the case-based modeling paradigm [3], which focuses on the dynamical adjustments of processes during execution and emphasizes the role of data as a driving force. At design time, process modules are modeled as separate process fragments, interlinked with data

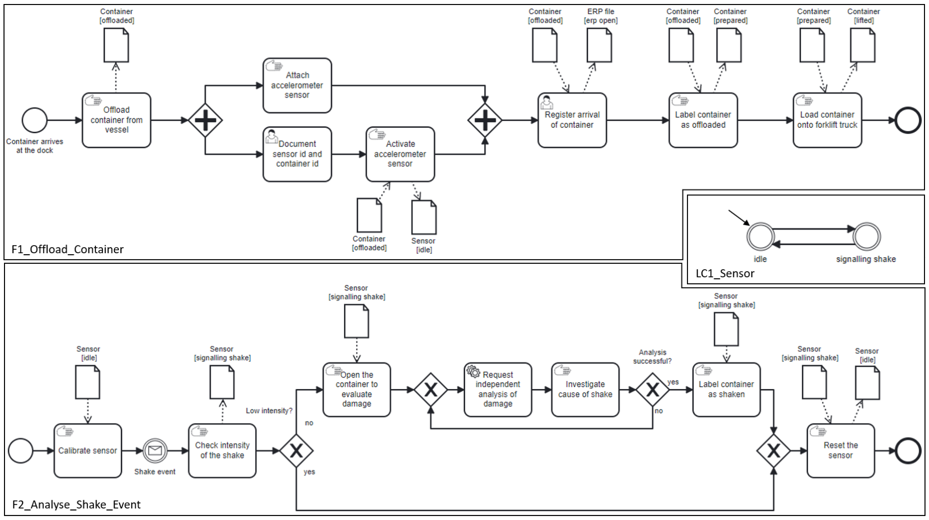

Fig. 1. Example of a horizontally modularized process model following the fragment-based approach. A better resolution is available at github.com/../figures [1].

objects. As the linkage between the fragments does not necessarily imply any hierarchical relationship among them, the fragment-based modeling approach can be seen as an instance of horizontal modularization (cf. Sect. 2.1). At run time, the fragments can be dynamically combined in any required order. The data objects can be defined as input conditions for the activities in the fragments and can thereby pose restrictions on the fragments. A data object belongs to a specific data class and is assigned to a state, which is mutable based on the execution of an activity. The set and order of possible states are depicted by a labeled transition system called lifecycle.

Figure 1 provides an example of the fragment-based modeling approach, as it was used in our experiment. The two fragments "F1_Offload_Container" and "F2_Analyse_Shake_Event" are connected based on the data object "Sensor". As soon as the activity "Activate accelerometer sensor" in the first fragment is executed the state of the sensor changes to "idle" and the second fragment can be executed. When the event "Shake event" occurs the activity "Check intensity of the shake" needs to be executed and the state of the sensor changes to "signalling shake". It is important to notice that the second fragment only depends on a single data-object, i.e., the sensor. This means that as soon as the sensor reaches the state "idle" the fragment can be executed multiple times, independently from any other fragment. The sensor lifecycle "LC1_Sensor" additionally shows that the state of the data-object can alternate between "idle" and "signalling shake".

Similar to the fragment-based approach, other techniques have been proposed in the literature. Therein, events were instead used to link the process fragments (cf. overview in [22]) or other modeling languages such as (colored) Petri nets were extended to achieve the same outcome [13]. We decided to focus on the fragment-based approach [17] since it is based on BPMN (i.e., the de-facto modeling language in the community) and provides in particular a clear execution semantic, based on data input and output conditions for the activities. In this way, the modeling approach provides a higher expressiveness and can depict real-world behavior more accurately.

2.3 The Effects of Abstraction and Fragmentation in Modularized Process Models

The use of modularization to support the comprehension of business processes has been subject to many empirical investigations (overview in [43]). Yet, based on Zugal's Literature review [43], there is no consensus on whether modularization supports or hinders model comprehension as existing empirical studies yielded inconsistent and therefore inconclusive results. The opposing effects of *abstraction* and *fragmentation* have been raised in the context of vertical modularization to explain these inconclusive results [32,36,43]. Similarly, they are likely to occur in other modularization types (e.g., horizental modularization), due to the spatial separation of modules.

Abstraction emerges from the ability of modularization to divide a large process into composable units (or fragments) and thus allowing users to focus their attention on the task-relevant fragments only while hiding the irrelevant ones. Following the insights in [30], users are presumably better at performing comprehension tasks when their attention is focused on the task-relevant parts in the model. Additionally, the recognition of patterns becomes much easier within the individual fragments compared to the overall process model, especially if their size remains within manageable limits. As pointed out in [27] large models are difficult to understand and thus should be decomposed into smaller fragments. Hence, with abstraction, modularization is expected to support model comprehension and lower users' cognitive load. Conversely, fragmentation is caused by the need to repeatedly switch attention between the different fragments, which is likely to cause the *split attention effect* [35,43]. This effect occurs when information is distributed across several locations, requiring the reader to repeatedly switch their attention between these locations, which can be distracting and more cognitively demanding. This effect is likely to impede model comprehension and raise users' cognitive load. Additionally, compared to abstraction, fragmentation is likely to affect users' behavior differently. Indeed, the attention switching and the underlying attention split effect would make the search for relevant information more complex and would raise the need to constantly integrate the information coming from the separated locations (i.e., fragments).

The effects of abstraction and fragmentation can be captured using *local* and *global* tasks. Therein, the local tasks requiring information within a single fragment can benefit from abstraction, whereas the global tasks, in which the relevant information is distributed over several fragments would be impeded by the effect of fragmentation. Accordingly, users are expected to get more challenged when solving global tasks compared to local ones. This hypothesis takes root in the cognitive fit theory [37], positing that the fit between the task at hand and the problem representation (i.e., the artifact presented to the user e.g., a modularized process model) affects the problem-solving procedure [12,37]. Since the fit is lower when solving global tasks (compared to the local ones), the problem-solving procedure would be more complex. Such complexity is likely to affect the model comprehension as well as users' cognitive load and behavior.

Our work extends existing research on the effects of abstraction and fragmentation [32,43] by (1) switching the focus from vertical to horizontal modularization, (2) using a wider array of measures capturing both users' comprehension and cognitive load to confirm these effects from different perspectives and (3) delving into users' behaviour to investigate the traits characterizing their engagement with local and global tasks. The measures used to confirm the effects of abstraction and fragmentation and to study the underlying users' behavior are introduced in the following section (cf. Sect 2.4).

2.4 Investigating Users' Comprehension, Cognitive Load and Behavior

The following paragraphs introduce the concepts and measures serving to (1) investigate the effects of abstraction and fragmentation on *model comprehension* and *cognitive load* as well as (2) assess users' behavior in terms of *search* and *cognitive integration* when solving local and global tasks.

Comprehension. Comprehension denotes the process of interpreting an artifact (e.g., text or process model) and building a mental representation of it [21]. In the literature, comprehension measures have been used as key dependent variables to study the effect of different factors (e.g., presentation medium, model complexity, the naming of activities, task type) on users' understanding of the model (cf. overview in [14]). Therein, low *comprehension accuracy* (i.e., the correctness of provided answers) and low *comprehension efficiency* (i.e., time needed to solve a task) indicate challenges in understanding the model at hand.

Cognitive Load. The cognitive load theory posits that when humans approach the full capacity of their working memory, they become more challenged with the task at hand, which affects their performance negatively and raises the risk of committing errors or taking wrong decisions [10,28]. In the context of process model comprehension, this would translate to difficulties understanding the given model and a higher risk of misinterpreting the encoded behavior.

Cognitive load can be captured *subjectively*, through the means of introspection, or *objectively* through changes in humans' behavior and cognitive states [10,18,39]. The subjective assessment of *perceived difficulty* is a measure that has been shown to differentiate different levels of cognitive load [34]. It is typically administrated as a questionnaire based on Likert scales following each task [34] (e.g., a *5-points Likert scale:* [0: "very easy", 4: "very difficult"]).

Eye-tracking provides substantial means to study humans' cognitive load more objectively [10,18,39]. Additionally, unlike the perceived difficulty measure which can be administrated only after a task is completed, the measures derived from eye-tracking can be collected continuously along the entire task, providing in turn, measurements with finer granularity in time and potentially more precise insights. In eye-tracking, the eye-mind hypothesis associates *fixations*[4] on the stimulus with mental processing under the assumption that what is fixated by the eyes is being simultaneously processed by the mind [20]. Accordingly, the *average fixation duration* has been used in several studies as an indicator of cognitive load such that the longer the fixation duration, the higher the user's cognitive load will be (e.g., [5,38]). Another prominent cognitive load measure is the *low/high index of pupil activity* (LHIPA) [11]. This measure is based on the pupillary features of the eyes. It separates the low pupil oscillations frequencies from the higher ones to derive a measure that correlates negatively with cognitive load, i.e., low value refers to high cognitive load [11].

Beside these cognitive load measures, the aforementioned comprehension measures have been also proposed to capture cognitive load in the literature [10].

Search. Search denotes the process of identifying task-relevant information among a set of irrelevant information [42,43]. In the process modeling literature, investigations of users' search behavior yield, for instance, interesting insights about their engagement with different hybrid process model representations [4], extending imperative process models with linked rules [38] or enriching declarative process models with textual annotations and guided simulations [6]. The duration of fixations has been used, in this vein, to discern different types of mental processes [15,38]. Notably, in [38], *fixations with duration* $\leq 250ms$ were associated with *information screening* behavior. Accordingly, increased numbers of this type of fixations imply a more pronounced search behavior. Another prominent search measure is the *scan-path precision* used in [30]. This measure calculates the ratio of fixations on the task-relevant parts of the model to the total number of fixations. Herein, a high ratio would indicate a simpler search, more focused on the task-relevant parts, whereas a low ratio would imply a more difficult search impeded by too many distracting fixations on non-relevant parts.

Cognitive Integration. Cognitive integration denotes the consolidation of information from different locations [8]. In the process modeling literature, visual associations derived from eye-tracking have been suggested to indicate cognitive integration [8]. In turn, the *average returns to relevant-regions* (or look-backs as defined in [18]) measures how often on average a relevant region (i.e., an area on the screen, which provides relevant information for a comprehension task) has been repeatedly fixated. It can therefore provide insights on visual associations made when recalling and integrating different pieces of task-relevant information located within a single fragment or spread over several fragments.

Overall, through our different measures of model comprehension and cognitive load, we aim at confirming that compared to the local tasks (benefiting from abstraction), the global tasks (impeded by fragmentation) yield lower comprehension accuracy and efficiency as well as higher perceived difficulty, higher average fixation duration, and lower LHIPA. Also, using our behavioral measures we aim at showing, that global tasks are associated with complex search

[4] A *fixation* refers to the time interval when the eye remains relatively stationary at a specific position within a visual field [18].

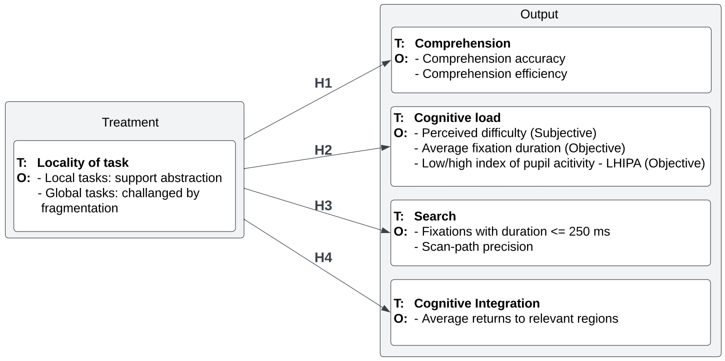

Fig. 2. Research model for hypothesis testing. T - Theoretical construct, O - Operationalization of construct.

and higher demand for cognitive integration than local tasks. Based the aferomentioned theoretical underpinnings, our measures are particularly suited to investigate the difference in local and global task solving. Their selection is also supported by existing cognitive theories and their successful application in previous experiments.

3 Research Method

To investigate the effects of abstraction and fragmentation, we have conducted an eye-tracking study following the guidelines in [2]. Sects. 3.1 – 3.3 provide an overview of the study design, the study execution, and the data analysis procedures respectively.

3.1 Study Design

Research Model. Based on the theoretical background introduced in Section 2, our empirical study aims at investigating the impact of abstraction and fragmentation on users' comprehension and cognitive load as well as users' behavior in terms of search and cognitive integration. Our research model is depicted in Figure 2. As *treatment*, we manipulate the *locality of task* factor, 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 Sect. 2.3, local tasks benefit from the effect of abstraction, whereas global tasks are challenged by the effect of fragmentation. The task locality factor is expected to impact *comprehension*, *cognitive load*, *search* and *cognitive integration*, which denote the theoretical constructs at the *output* side. As shown in Fig. 2, we operationalize these theoretical constructs using the measures introduced in Sect. 2.4. Following our research model, we formulate the following hypotheses:
• $\mathbf{H}_1$: Global tasks yield lower model comprehension than local tasks in horizontal modularization.
• $\mathbf{H}_2$: Global tasks yield a higher cognitive load than local tasks in horizontal

modularization.
• $\mathbf{H}_3$: Global tasks yield more complex search than local tasks in horizontal modularization.
• $\mathbf{H}_4$: Global tasks yield a higher cognitive integration than local tasks in horizontal modularization.

Material. The material used for the experiment comprises a process model, based on the fragment-based modeling approach proposed in [17], and a set of local and global tasks capturing abstraction and fragmentation effects respectively. The model depicts a logistic process and consists of six different fragments. Each fragment refers to a particular part of the onward carriage process of a container after it arrived at a port, which involves the unloading, scanning, temporally storing and loading of a container, as well as its registration and monitoring. The fragments are connected based on three different data objects, i.e., container, enterprise resource planning file, and sensor. The possible state changes of each data object are additionally visualized by three respective life cycles. To avoid confounding factors [41] the design of the layout and the modeling constructs are carefully aligned with the guidelines proposed by [7], which will be addressed in the following.

All process fragments are of similar size (six to eight unique activities) and complexity. Each fragment contains two to four gateways, which are either exclusive (XOR) or parallel (AND). Since domain knowledge can also be a confounding factor [41], the activities are labeled in layman's terms and specific details of a real logistic process are left out.

In total, the material includes eight different tasks, which have to be solved by every participant. Each task consists of a question, which is posed as a true or false statement. The statements address one of the following control-flow patterns: sequential, concurrent, exclusive, or repetition. Each statement refers to two particular activities, which can either be found in a single fragment (local task) or in two separate fragments (global task). For each control-flow pattern, a participant has to solve one local and on global task. Figure 3 provides an overview of all eight tasks, the control-flow patterns used in the statements and how they are modeled for the local and the global task type respectively. Additionally, Figure 3 shows that the used tasks were formulated following a predefined template. For each control-flow pattern, the associated local and global tasks were phrased in the same way. Also, each pair of local and global tasks have the same answer (either both true or both false). With this formulation, we ensure that no task-related confounding factors are affecting our design since the only difference between the local and global tasks is the locality of the activities, i.e., either within the same fragment or spanning over two fragments. The experiment material is available online at github.com/../material [1].

Figure 4 depicts a screenshot of the 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 one single file can be viewed at a time. The question is shown at the very top of the screen.

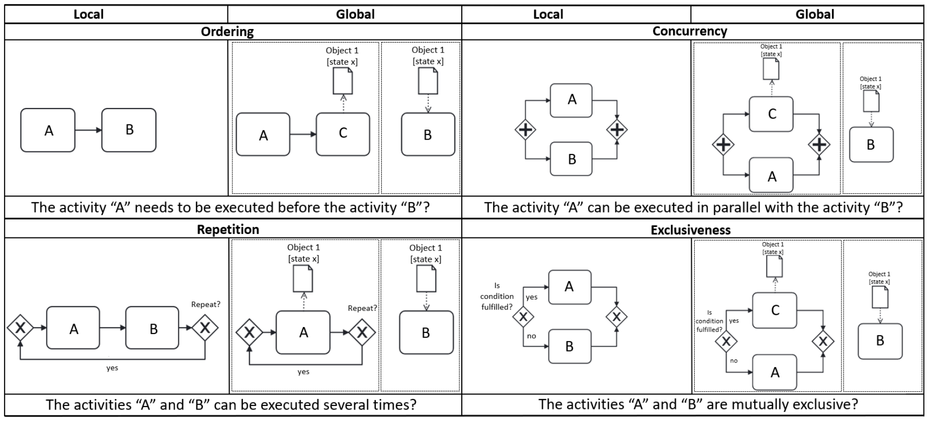

Fig. 3. Difference between local (left) and global (right) tasks and how they are modeled by the fragment-based modeling approach. A better resolution is available at github.com/../figures [1].

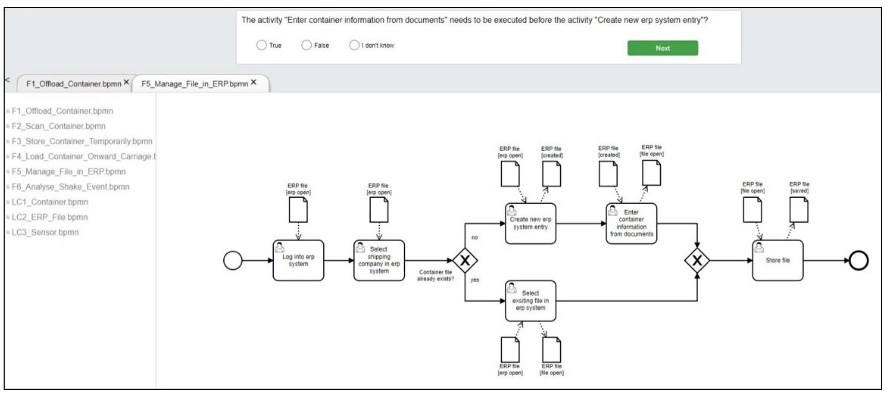

Fig. 4. User interface with an example of an experiment question.

Participants. The eye-tracking study covers 46 participants. The experiment sessions were conducted at four different locations: at the University of St. Gallen (22), at the Karlsruhe Institute of Technology (17), at the research institute Forschungszentrum für Informatik FZI in Karlsruhe (4), and at the IT consulting company Promatis in Karlsruhe (3), within a time frame of 4 weeks. The participants were between 20-50 years old with a majority in the age range of 20-30 years (63%). For the professional background, there were three main groups: researchers at PhD level or higher (22), students (17), and participants with an industry or IT-admin background (7). The participants had different levels of expertise in BPMN. To ensure that they could take part in the experiment, we have conducted a uniform familiarization and tested their knowledge with a quiz as it will be explained in Sect. 3.2. Additionally, to avoid confounding factors related to differences in the participants' expertise, we have used a within-subject design. Therein, every participant was exposed to each factor level multiple times, i.e., four local and four global tasks. Although the participants' expertise might

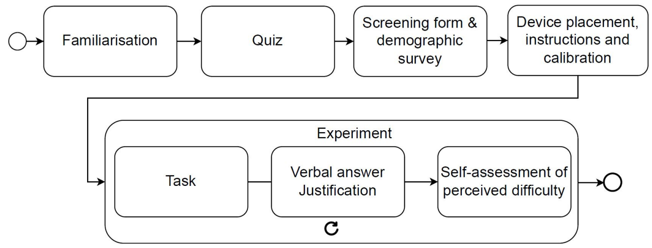

Fig. 5. The different steps during each experiment session.

still impact the outcome of each task, based on the within-subject design both treatments are equally impacted and it is therefore possible to draw valid conclusions regarding their distinctiveness. A complete overview of the participants' demographic data is available at github.com/../demograhics [1].

3.2 Experiment Procedure

Each participant was invited to an individual experiment session, which lasted about one hour. The experiment was conducted in a controlled lab environment. The procedure for each session is depicted in Figure 5. The session starts with familiarizing the participant with BPMN and the fragment-based modeling approach. The familiarization is followed by a short quiz, consisting of two comprehension tasks, to test whether the fragment-based modeling approach is understood. During the quiz, the participant could ask additional questions and receive feedback regarding the solution. The actual experiment starts with two test tasks, which allow the participants to familiarize themselves with the user interface and to get to know the procedure. These tasks are not used for the data analysis later on. This approach is chosen to avoid any bias in the data due to the unfamiliarity of the participants with the experiment setup. To compensate for a potential learning effect, the tasks are randomized and presented in different orders to the participants. For each task, the participant has to answer a comprehension question and two additional follow-up questions. In the first follow-up question, the participant is asked to verbally justify how the task was solved. In the second follow-up question, the participant is asked to provide a self-assessment of how difficult the task was.

3.3 Data Analysis

Based on the collected data, the measures introduced in Sect. 2.4 are calculated. 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 statistics in order to investigate our hypotheses (cf. Table 1). 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. Our analysis approach is documented through the used Python notebooks available online at github.com/../analysis [1].

4 Findings

This section presents the findings of our empirical study. Sect. 4.1 reports the results of the comprehension and cognitive load analysis, while Sect 4.2 reports the results of the behavioral analysis.

4.1 Comprehension and Cognitive Load Analysis

Based on the results in Table 1, *comprehension accuracy* (in the range [0:incorrect, 1:correct]) was significantly ($p< .001$) lower for global tasks (*M=0.739*) than for local tasks (*M=0.978*). Likewise, *comprehension efficiency* (measured in seconds) was significantly ($p< .001$) lower for global tasks (*M=117.593 seconds*) compared to the local ones (*M=54.097 seconds*). Based on the background presented in Sect. 2.4, these two measures indicate that global tasks yield reduced model comprehension compared to local tasks, which confirms *Hypothesis* H_1.

With regards to cognitive load, the subjective assessment of *perceived difficulty* (measured in the range [0: "very easy", 4:"very difficult"]) was significantly ($p< .001$) higher for global tasks (*M=1.891*) than for local tasks (*M=0.62*). Similarly, the *average fixation duration* was significantly ($p< .001$) higher for global tasks (*M=200.966*) compared to local tasks (*M=194.323*). Finally, the *LHIPA* score was significantly ($p< .001$) lower for global (*M=0.815*) than for local (*M=1.185*) tasks. The trends of these measures (cf. Sect. 2.4) show that when solving global tasks users experience increased cognitive load, compared to local tasks, which confirms *Hypothesis* H_2.

4.2 Behavioral Analysis

Based on Table 1, the *fixations with duration* $\leq 250ms$ are significantly higher ($p< .001$) for global *(M=298.883)* than for local tasks *(M=142.153)*. This indicates that users' exhibited more information-screening behavior when performing global tasks (cf. Sect. 2.4). Additionally, the *scan path precision* is significantly lower ($p< .001$) for global tasks *(M=0.077)* than for local tasks *(M=0.111)*. This indicates that for global tasks users had to look and scan through a higher number of irrelevant process model elements before finding the relevant ones (cf. Sect. 2.4). These insights hint towards a complex search when dealing with global tasks, which confirm *Hypothesis* H_3.

With regards to cognitive integration, the *average returns to relevant regions* is significantly higher ($p< .001$) for global tasks *(M=7.844)* compared to the local

Hypothesis/Construct	Measure	Descriptive			Inferential
		N	Local M	Global M	p-value
Comprehension and Cognitive Load Analysis					
H_1/Comprehension	Comprehension accuracy	46	0.978	0.739	$< .001$
	Comprehension efficiency	46	54.097	117.593	$< .001$
H_2/Cognitive load	Perceived difficulty	46	0.62	1.891	$< .001$
	Average fixation duration	44	194.323	200.966	$< .001$
	LHIPA	43	1.185	0.815	$< .001$
Behavioral Analysis					
H_3/Search	Fixations with duration ≤ 250 ms	44	142.153	298.883	$< .001$
	Scan path precision	44	0.111	0.077	$< .001$
H_4/Cognitive Integration	Av. returns to relevant regions	44	4.108	7.844	$< .001$

Table 1. Comprehension, cognitive load and behavioral analyses. N: number of observations (cf. Sect. 3.3), M: calculated mean. A *p-value*$< .05$ means that the pairwise comparison results are significant.

ones *(M=4.108)*, showing, in turn, more demand for recalling and integrating information, spread over different fragments, when engaging with global tasks (cf. Sect. 2.4). This finding confirms *Hypothesis* H_4.

5 Discussion

Based on the inferential statistics reported in Sect. 4, we can confirm that the task type has an impact on users' comprehension and cognitive load when engaging with horizontally modularized process models. This validates the existing suppositions in the literature [32,43] , i.e., while local tasks require less cognitive effort, due to the effect of abstraction, global tasks require more cognitive effort, due to the effect of fragmentation. Additionally, our behavioral analysis suggests that the cognitive processes underlying global and local tasks differ, as global tasks are associated with a more complex search and require more cognitive integration than local tasks. Since horizontal modularization allows for a very flexible decomposition of process models into modules, regardless of their hierarchical relationships and cross-cutting concerns, it can be assumed that the gained insights generally apply to other modularization approaches as well, which are more restrictive in terms of decomposition.

Following these insights, we suggest that researchers should consider the task type as an integral part of their research models aiming at explaining the comprehension of process models. This proposal finds also support in other literature, demonstrating the impacts of different task types on the understandability of software artifacts [12,33]. The framework of Mandelburger and Mendling [25] can be seen as a prominent example in this direction as it motivates the shift from artifact-centric to task-centric research. This way, the comprehension of process models would not only be justified by the model proprieties (e.g., size, density, connectivity) but also through the characteristics of the task at hand.

From a practitioner's perspective, it should be recognized that not only design choices impact the comprehensibility of a modular system, but also how specific tasks are defined and distributed among development teams. Tasks should, if possible, be confined to specific self-contained modules instead of multiple modules at once. This is indeed in line with existing software engineering practice and the general idea of microservice architecture [19]. At the same time it should be clear that in general, tasks cannot always be confined to a single module. This holds, in particular, true for cross-cutting concerns, such as security and privacy. In this case, the restriction of tasks to specific services or modules hinders the comprehensibility of the overall system and therefore the ability to solve cross-cutting concerns [19]. To overcome this issue our findings motivate the need to develop support for search and cognitive integration to compensate for some of the negative effects of fragmentation.

One possibility to support the search as well as the cognitive integration of information is to use modeling tools interlinking modules to improve their navigation. Yet, empirical evidence is missing, on how such a navigation affects model comprehension [24]. Another approach to avoid the searching and integration issues is to use simulators to evaluate the executions allowed under different scenarios as proposed in [26]. Therein, for certain global tasks users do not necessarily need to search and cognitively integrate information across several modules, since the simulation can provide information regarding global process behavior. To further facilitate the integration of information, one could also provide a general overview or mapping of the modules, e.g., in the form of a model landscape. This approach has also been applied for fragment-based process models and empirical tests show indeed a benefit in terms of process model comprehensibility [16]. Additionally, to particularly facilitate the search for information, one could use natural language processing based on the lexical similarity between task description and the textual corpus of the software artifact. A similar approach has been proposed for feature location in software systems [9].

All in all, future BPM tools should incorporate the characteristics of the user task to augment their support. Based on our empirical findings, for global tasks, this support should particularly focus on reducing the search and cognitive integration effort underlying this type of tasks.

Threats to Validity. Our work can be subject to some validity threats. Although the experiment sessions are conducted in a unified manner based on an experiment protocol and in a controlled environment, one cannot eliminate the possibility of the existence of confounding factors, which might have influenced the outcome of the study and therefore its internal validity. However, several measures are taken to mitigate this risk. To avoid learning effects, the comprehension tasks were presented in a random order to each participant. These tasks also refer to different process fragments or different combinations of fragments. Moreover, each of the used six fragments is only relevant for two out of eight tasks. The relatively high number of subjects for an eye-tracking study (46) further supports the internal validity of the experiment. Each participant received a uniform introduction to BPMN and to the fragment-based modeling approach, thus ensuring that all participants have a similar basic knowledge to participate in the study. A further concern is whether the identified findings could be applied to other process modeling languages. This risk to external validity is again mitigated to some extent by the design of our comprehension tasks addressing different control-flow patterns. Hence, the results are less dependent on particular model structures. Also, it is worthwhile to mention that these control-flow patterns are typically found in process models [43].

6 Conclusion and Future Work

The conducted eye-tracking study confirms that, in horizontal modularization, global tasks impede the model comprehension and raise users' cognitive load in comparison to local tasks. This can be explained by the two opposing effects of abstraction and fragmentation, caused by modularization. Additionally, based on our behavior analysis, it becomes, in particular, apparent that global tasks require complex search and a higher cognitive integration effort, due to fragmentation. To overcome this issue we propose several possible tool augmentations and motivate the need to attend more attention to the task perspective.

In the future, we are planning to investigate tasks addressing the data-flow perspective in fragmented process models. Additionally, we will extend our behavioral investigation to identify other traits characterizing the solving of the covered tasks. Moreover, we can move our analysis of cognitive load to a more fine-grained level to obtain more detailed insights into the particular modeling constructs challenging users in different tasks. Lastly, we can also investigate the effects of the proposed augmented tooling on solving global tasks.

Data Availability. As pointed out in the relevant sections of the paper, the experiment and analysis material are available on GitHub [1].

References

1. Data repository, `https://github.com/aminobest/BPM2023TaskType`
2. Emperical standard guidelines for experiments, `https://github.com/acmsigsoft/EmpiricalStandards/blob/master/docs/Experiments.md`
3. Van der Aalst, W.M., Weske, M., Grünbauer, D.: Case handling: a new paradigm for business process support. Data & knowledge engineering **53**(2), 129–162 (2005)
4. Abbad-Andaloussi, A., Burattin, A., Slaats, T., Kindler, E., Weber, B.: On the declarative paradigm in hybrid business process representations: A conceptual framework and a systematic literature study. Information Systems **91**, 101505 (2020)
5. Abbad-Andaloussi, A., Sorg, T., Weber, B.: Estimating developers' cognitive load at a fine-grained level using eye-tracking measures. In: Proceedings of IEEE/ACM International Conference on Program Comprehension. pp. 111–121 (2022)
6. Abbad-Andaloussi, A., Zerbato, F., Burattin, A., Slaats, T., Hildebrandt, T.T., Weber, B.: Exploring how users engage with hybrid process artifacts based on declarative process models: a behavioral analysis based on eye-tracking and think-aloud. Software and Systems Modeling **20**, 1437–1464 (2021)
7. Becker, J., Rosemann, M., Von Uthmann, C.: Guidelines of business process modeling. In: Business Process Management: Models, Techniques, and Empirical Studies, pp. 30–49. Springer (2002)

8. Bera, P., Soffer, P., Parsons, J.: Using eye tracking to expose cognitive processes in understanding conceptual models. MIS Quarterly **43**(4), 1105–1126 (2019)
9. Biggers, L.R., Bocovich, C., Capshaw, R., Eddy, B.P., Etzkorn, L.H., Kraft, N.A.: Configuring latent dirichlet allocation based feature location. Empirical Software Engineering **19**, 465–500 (2014)
10. Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S.Z., Khawaji, A., Conway, D.: Robust multimodal cognitive load measurement. Springer (2016)
11. Duchowski, A.T., Krejtz, K., Gehrer, N.A., Bafna, T., Bækgaard, P.: The low/high index of pupillary activity. In: CHI Conference on Human Factors in Computing Systems. pp. 1–12 (2020)
12. Dunn, C., Grabski, S.: An investigation of localization as an element of cognitive fit in accounting model representations. Decision Sciences **32**(1), 55–94 (2001)
13. Fettke, P., Reisig, W.: 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. pp. 77–89. Springer (2021)
14. Figl, K.: Comprehension of procedural visual business process models: a literature review. Business & Information Systems Engineering **59**, 41–67 (2017)
15. Glöckner, A., Herbold, A.K.: Information processing in decisions under risk: Evidence for compensatory strategies based on automatic processes. MPI collective goods preprint (2008/42) (2008)
16. Gonzalez-Lopez, F., Pufahl, L., Munoz-Gama, J., Herskovic, V., Sepúlveda, M.: Case model landscapes: toward an improved representation of knowledge-intensive processes using the fcm-language. Software and Systems Modeling (2021)
17. Hewelt, M., Weske, M.: 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. pp. 38–54. Springer (2016)
18. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., van de Weijer, J.: Eye Tracking: A comprehensive guide to methods and measures. OUP Oxford (2011)
19. Jamshidi, P., Pahl, C., Mendonça, N.C., Lewis, J., Tilkov, S.: Microservices: The journey so far and challenges ahead. IEEE Software **35**(3), 24–35 (2018)
20. Just, M.A., Carpenter, P.A.: A theory of reading: From eye fixations to comprehension. Psychological review **87**(4), 329 (1980)
21. Keselman, A., Slaughter, L., Patel, V.L.: Toward a framework for understanding lay public's comprehension of disaster and bioterrorism information. Journal of Biomedical Informatics **38**(4), 331–344 (2005)
22. Krumeich, J., Weis, B., Werth, D., Loos, P.: Event-driven business process management: where are we now?: A comprehensive synthesis and analysis of literature. Business Process Management Journal **20**(4), 615–633 (2014)
23. La Rosa, M., Wohed, P., Mendling, J., Ter Hofstede, A.H., Reijers, H.A., van der Aalst, W.M.: Managing process model complexity via abstract syntax modifications. IEEE Transactions on Industrial Informatics **7**(4), 614–629 (2011)
24. Lübke, D., Ahrens, M.: Towards an experiment for analyzing subprocess navigation in bpmn tooling (2022)
25. Mandelburger, M.M., Mendling, J.: Cognitive diagram understanding and task performance in systems analysis and design. MIS Quarterly **45**(4), 2101–2157 (2021)
26. Marquard, M., Shahzad, M., Slaats, T.: 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. pp. 209–225. Springer (2015)

27. Mendling, J., Reijers, H.A., van der Aalst, W.M.: Seven process modeling guidelines (7pmg). Information and software technology **52**(2), 127–136 (2010)
28. Paas, F., Tuovinen, J.E., Tabbers, H., Van Gerven, P.W.: Cognitive load measurement as a means to advance cognitive load theory. Educational psychologist (2003)
29. Parnas, D.L.: On the criteria to be used in decomposing systems into modules. Communications of the ACM **15**(12), 1053–1058 (1972)
30. Petrusel, R., Mendling, J., Reijers, H.A.: How visual cognition influences process model comprehension. Decision Support Systems **96**, 1–16 (2017)
31. Recker, J.C., Lukyanenko, R., Jabbari Sabegh, M., Samuel, B., Castellanos, A.: From representation to mediation: a new agenda for conceptual modeling research in a digital world. MIS Quarterly: Management Information Systems **45**(1) (2021)
32. Reijers, H.A., Mendling, J.: Modularity in process models: Review and effects. In: BPM. vol. 8, pp. 20–35. Springer (2008)
33. Ritchi, H., Jans, M.J., Mendling, J., Reijers, H.: The influence of business process representation on performance of different task types. Journal of Information Systems (2019)
34. Sweller, J., Ayres, P., Slava, K.: Cognitive load theory. Springer (2011)
35. Sweller, J., Chandler, P.: Why some material is difficult to learn. Cognition and instruction **12**(3), 185–233 (1994)
36. Turetken, O., Rompen, T., Vanderfeesten, I., Dikici, A., van Moll, J.: The effect of modularity representation and presentation medium on the understandability of business process models in bpmn. In: Business Process Management: 14th International Conference, BPM 2016, Rio de Janeiro, Brazil, September 18-22, 2016. Proceedings 14. pp. 289–307. Springer (2016)
37. Vessey, I.: Cognitive fit: A theory-based analysis of the graphs versus tables literature. Decision sciences **22**(2), 219–240 (1991)
38. Wang, W., Chen, T., Indulska, M., Sadiq, S., Weber, B.: Business process and rule integration approaches—an empirical analysis of model understanding. Information Systems **104**, 101901 (2022)
39. Weber, B., Fischer, T., Riedl, R.: Brain and autonomic nervous system activity measurement in software engineering: A systematic literature review. Journal of Systems and Software **178** (2021)
40. Winter, M., Pryss, R., Probst, T., Baß, J., Reichert, M.: Measuring the cognitive complexity in the comprehension of modular process models. IEEE Transactions on Cognitive and Developmental Systems **14**(1), 164–180 (2022)
41. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: Experimentation in software engineering. Springer Science & Business Media (2012)
42. Wolfe, J.M.: Guided search 2.0 a revised model of visual search. Psychonomic bulletin & review **1**, 202–238 (1994)
43. Zugal, S.: Applying cognitive psychology for improving the creation, understanding and maintenance of business process models. Ph.D. thesis, University of Innsbruck (2013)