Software & Systems Modeling (2020) 19:531–554
https://doi.org/10.1007/s10270-019-00742-z

SPECIAL SECTION PAPER

From analytical purposes to data visualizations: a decision process guided by a conceptual framework and eye tracking

Jens Gulden[1] · Andrea Burattin[2] · Amine A. Andaloussi[2] · Barbara Weber[2,3]

Received: 30 October 2017 / Revised: 24 May 2019 / Accepted: 1 July 2019 / Published online: 10 July 2019

Abstract

Data visualizations are versatile tools for gaining cognitive access to large amounts of data and for making complex relationships in data understandable. This paper proposes a method for assessing data visualizations according to the purposes they fulfill in domain-specific data analysis settings. We introduce a framework that gets configured for a given analysis domain and allows to choose data visualizations in a methodically justified way, based on analysis questions that address different aspects of data to be analyzed. Based on the concepts addressed by the analysis questions, the framework provides systematic guidance for determining which data visualizations are able to serve which conceptual analysis interests. In a second step of the method, we propose to follow a data-driven approach and to experimentally compare alternative data visualizations for a particular analytical purpose. More specifically, we propose to use eye tracking to support justified decisions about which of the data visualizations selected with the help of the framework are most suitable for assessing the analysis domain in a cognitively efficient way. We demonstrate our approach of how to come from analytical purposes to data visualizations using the example domain of Process Modeling Behavior Analysis. The analyses are performed on the background of representative analysis questions from this domain.

Keywords Data visualization · Process execution data · Process Modeling Behavior Analysis · Eye tracking · Reading patterns · Process mining

1 Introduction

Visual analysis tools allow to leverage capabilities of the human cognitive apparatus which is capable of pattern-based processing of perceived stimuli on multiple levels of granularity in parallel [9,10,21,22]. These kinds of analyses allow to gain insights by projecting data into appropriate perceptual spaces and thus offer a complementary perspective on existing statistical approaches [39,42].

In order to perform visual analyses of data, it is necessary to be aware of the capabilities of data visualizations to fulfill the information needs that arise in a specific analysis setting. Information needs are depending on the domain from which data are analyzed and on the purposes that underlie the analysis.

Up to now, there is, however, little methodical guidance in how to reasonably justify the choice of visualizations that are used during an analysis. Visualizations are often chosen in an ad hoc manner, possibly in trial-and-error iterations, until a suitable one is available. It is an important research goal to establish a systematic link between purposes and conceptual analysis questions on the one hand, and the expressiveness and appropriateness of visualizations for fulfilling these purposes on the other hand [22]. This article contributes to this goal.

Communicated by Selmin Nurcan and Rainer Schmidt.

✉ Jens Gulden
jensgulden@acm.org

Andrea Burattin
andbur@dtu.dk

Amine A. Andaloussi
amab@dtu.dk

Barbara Weber
barbara.weber@unisg.ch

1 University of Duisburg-Essen, Essen, Germany

2 Technical University of Denmark, Kongens Lyngby, Denmark

3 University of St. Gallen, St. Gallen, Switzerland

In the upcoming elaboration, we establish a theoretical conceptualization of intended purposes for visualizations used in analysis scenarios, which results in a multidimensional framework in which aspects of visual analysis scenarios can be mapped both onto questions that reflect analysis purposes on the conceptual level and on visualizations that are capable of answering these questions. The *Vi*sualization-*P*urposes (ViP) framework allows to match information demands that are characterized by purposes of an analysis setting with visualizations used for the analysis.

Using the ViP framework, we identify candidates of visualization types satisfying the information needs specific to the analysis domain. We then propose to follow a systematic and data-driven approach and to experimentally compare the identified candidates regarding their cognitive suitability for our analytical purposes using eye tracking. Both phases together, the application of the mapping framework followed by performing a psychophysiological experimental analysis, form a method for deriving appropriate analytical visualizations from analysis purposes in a justified and systematic manner.

To demonstrate the usefulness of our approach, we use the domain of Process Modeling Behavior Analysis as an example [8,32], which is concerned with analyzing the behavior of human modelers while creating process models. In order to gain an analytical understanding from process modeling behavior data, it is important to apply analysis techniques which allow to navigate through the available data in an exploratory manner, rather than to perform statistical analyses that presuppose an underlying structure of the data. In previous examinations [44], we deployed the Rhythm-Eye visualization type [20] (cf. Sect. 2.2.3) to test its applicability for answering three particular analysis questions. This article extends that work and adds the Modeling Phase Diagram (MPD) [8] (cf. Sect. 2.2.1) and PPM Chart [12] (cf. Sect. 2.2.2) visualization types to the set of examined candidates, to provide a comprehensive coverage of existing visualization types for Process Modeling Behavior Analysis.

Section 2 introduces the backgrounds needed for the further understanding of the paper including the related work we build our approach upon. The methodical approach including the multidimensional ViP framework for matching analysis purposes and visualization types is explained in Sect. 3. In Sect. 4, the design of the eye tracking study to compare visualization types in terms of effectiveness and reading patterns is laid out. The data analysis and the results are presented in Sect. 5, whereas Sect. 6 reports some reflections on the limitations and the impact of the research. In Sect. 7, we draw a conclusion and relate our findings to prospective future work.

2 Backgrounds and related work

This section introduces backgrounds needed for the further understanding of the paper including related work we build our approach upon. We first discuss different frameworks for categorizing purposes of visualizations (cf. Sect. 2.1) and then introduce the Process Modeling Behavior Analysis as the domain we chose to demonstrate our approach, including existing visualization types from this domain (cf. Sect. 2.2). Finally, we provide backgrounds and related work on applying eye tracking evaluations to visual analytics (cf. Sect. 2.3).

2.1 Frameworks for categorizing purposes of visualizations

The demand for systematically describing visualizations that are used in analysis scenarios has been recognized by a number of representatives in the existing body of literature. In the domain of Enterprise Architecture Management (EAM), the EAM Pattern Catalog [7,28] introduces, among others, the notion of so-called methodology patterns and viewpoint patterns for EAM. Methodology patterns describe analytical tasks that occur in EAM, e.g., "identify organizational units where a lot of changes take place" or "find organizational units with an exceptionally high amount of (not) standardized business applications" [28]. These tasks define the purpose of the analysis. Viewpoint patterns describe visual representations that can be used to support these tasks. They reflect the notion of visualization types in this article and cover a range from classical data visualizations, e.g., diagrams, charts or tables, up to domain-specific visualization types for EAM such as cluster maps to visualize different logical and physical views on an enterprise architecture.

The EAM Pattern Catalog lists selected methodology patterns and suggests corresponding viewpoint patterns as best-practice recommendations to perform analysis tasks. This way, a domain-specific mapping between analysis tasks and corresponding visualizations is established. However, the EAM Pattern Catalog does not operate with generalizable abstractions that would allow to transfer parts of the results to other domains. All descriptions are in natural language, and there is no formalization that allows to answer questions about why some visualizations better serve the purpose of answering specific domain analysis questions than others. In contrast to this, the categorization framework developed in this article provides a conceptual link between described purposes and visual means of expression that are made available by different types of visualizations. It could in principle serve as an overarching formalization approach for the EAM Pattern Catalog, based on an in-depth analysis of the analysis patterns and visualization types with which the framework would have to be configured (cf. Sect. 3.2).

Other approaches for relating intended analysis purposes to visualizations come in the form of guidelines for ensuring correct, nonmisleading visualizations in business scenarios. As one example, the International Business Communication Standards (IBCS) [23] offer a set of prescriptive hints that support in creating presentation graphics such as charts, diagrams, and tables for the domain of business reports and presentations. The set of rules is divided into three subsets, which are conceptual rules, perceptual rules, and semantic rules. Conceptual rules refer to the contents that are to be visualized and give hints on what elements should be part of the information that is to be conveyed, e.g., the rule that there should be an unambiguous title for each element in a presentation. Perceptual rules refer to best practices and scientifically justified design guidelines for visual representations of information, e.g., the rule that distances should be preferred over areas to express magnitudes in diagrams. Semantics rules, as they are called in the IBCS approach, describe design conventions that ensure a unified appearance of multiple different visualizations. Although this collection provides a valuable source of best practices for using visualizations in a defined domain, there is no theoretical reflection about the characteristics of domain-specific analysis purposes in relation to the rules that are proposed.

A number of publications exclusively deal with visual characteristics of diagrams, charts, and other forms of information visualization. These works partially originate from times where information visualization was not yet related to computer-generated graphics, such as the initial works of Bertin [5] and Tufte [43] on (manually drawn) diagrams. Beginning with the time that computers played a increasingly important role in visually displaying information, works such as investigations about information dashboard design [16,17] and user interaction [3,15,38] provide insight into relevant characteristics of graphical displays and their interactive features. A wide variety of publications deal with design principles for data visualization and information graphics [6,9–11,29,41]. All these contributions, however, exclusively argue about the expressiveness of visual characteristics and general categories of meaning that can be assigned to them. None of them incorporate a systematic mapping between domain-specific analysis purposes and corresponding visual representations. It is the very aim of the work proposed in this article to provide a justified methodological link of this kind. This is achieved by leveraging the ViP framework, which on the one hand acts as a conceptual framework that structures analysis purposes expressed in words, and on the other hand makes use of locations in space to express relationships between purposes and visualizations in a nonverbal, spatially embodied [18,26] way. This double-faced nature of the framework is the key approach to bridging between verbally expressed purposes and characteristics of visualizations. Together with empirical eye tracking analysis, the use of the framework constitutes the methodical approach for justified selection of visualizations for analytical purposes introduced in this article.

2.2 Process Modeling Behavior Analysis

Process Modeling Behavior Analysis (PMBA) refers to the analysis of process traces with the aim of identifying patterns of behavior. A process traces thereby describes the behavior of an individual (e.g., a single modeler), a group (e.g., a software development team), an organization (e.g., in the case of business processes), or society (e.g., network analysis) as a sequence of events. Process traces can originate, for example, from the interactions of a user with a development platform (e.g., creating a digital artifact like a process model or a piece of source code). A process trace can also describe a user's fixations on the screen when interacting, for example, with a digital artifact like a process model, a piece of software, or a Web site. Other examples of a process trace could describe how a user moves around in a smart home or how a business process is executed.

These events can be aggregated into phases (e.g., by aggregating events temporarily or spatially). For example, when creating a process model, all user interactions that constitute structural changes to a model could be aggregated into modeling phases. In turn, when analyzing eye tracking data, all fixations that belong to a specific area of interest on the screen could be combined into a phase.

The ViP framework provides systematic guidance to identify visualization types for analyzing process traces as well as abstractions thereof with the goal to discover behavioral patterns. To exemplify the framework, we decided to apply it to one domain instance: the process of creating a process model, also denoted as process of process modeling (PPM). This is an iterative process during which a modeler communicates with a modeling platform through model interactions to gradually evolve a process model. Model interactions include, for example, the creation or deletion of activities, edges, or gateways, and the movement of elements on the modeling canvas. The sequence of model interactions that led to a particular model is denoted as PPM instance. At a more abstract level, the process of creating a process model involves different phases [32], i.e., comprehension, modeling, and reconciliation, that can be combined in a flexible way (i.e., phases can occur repeatably and phases can be skipped) [33]. During *comprehension* phases, the modeler understands the problem at hand and builds an internal representation of it, i.e., a mental model. During *modeling* phases, the modeler interacts with the modeling tool in order to externalize the mental model and to create an actual representation. Finally, *reconciliation* phases represent actions aiming at improving the understandability of the model by changing the layout of the modeling. In general, creation and deletion interactions (e.g., creating a new task on the modeling canvas, or deleting an edge between two tasks) can be classified as modeling phase. Interactions to rename modeling elements and move elements usually characterize reconciliation phases. Comprehension phases, in turn, are phases without any interaction between the user and the modeling tool. Techniques for the automatic identification of modeling phases are reported in the literature [35].

As just described, the nature of the data generated from the process of process modeling is clear: it is possible to investigate it at the "event level" (i.e., interactions with the modeling platform) or at the "phase level." These data provide the empirical underpinning for PMBA. PMBA was initially developed to deal with the analysis of data concerning model creation and editing. It supports the exploration of data with the goal to discover behavioral patterns which, e.g., is relevant for the development of context-aware modeling support that guides users during the creation of model artifacts.

Gained insights allow to identify modeling styles [34] or build personalized modeling environments or tailored methods [13]. In the literature, various visualization types for Process Modeling Behavior Analysis applied to the process of process modeling have been proposed, including Modeling Phase Diagrams [35] (cf. Sect. 2.2.1), PPM Chart visualizations [12] (cf. Sect. 2.2.2), and Rhythm-Eye visualizations [20] (cf. Sect. 2.2.3).

2.2.1 Modeling phase diagrams

The first visualization type for PMBA in process of process modeling we introduce is called Modeling Phase Diagrams (MPD) [35]. An example of such diagram is depicted in Fig. 1a[1], and it is a line chart where the x axis refers to the modeling time and the y axis reports the number of activities observed in the modeling canvas at each point in time. Additionally, the line is divided into different segments of varying colors encoding the specific modeling phase that the given time period refers to (i.e., modeling, reconciliation, comprehension) abstracting from the underlying interactions with the modeling environment (i.e., modeling events) and grouping model interactions into phases (based by an algorithm proposed in [35]). Moreover, Modeling Phase Diagrams provide a representation that makes it easy to see the pace at which elements were added or removed. The Modeling Phase Diagram depicted in Fig. 1a shows an example of a process instance with only short comprehension phases, numerous modeling phases followed by reconciliation phases throughout the entire process, and a very long reconciliation phase at the end.

[1] Larger versions of the figures are available at https://doi.org/10.5281/zenodo.1419594.

2.2.2 PPM charts

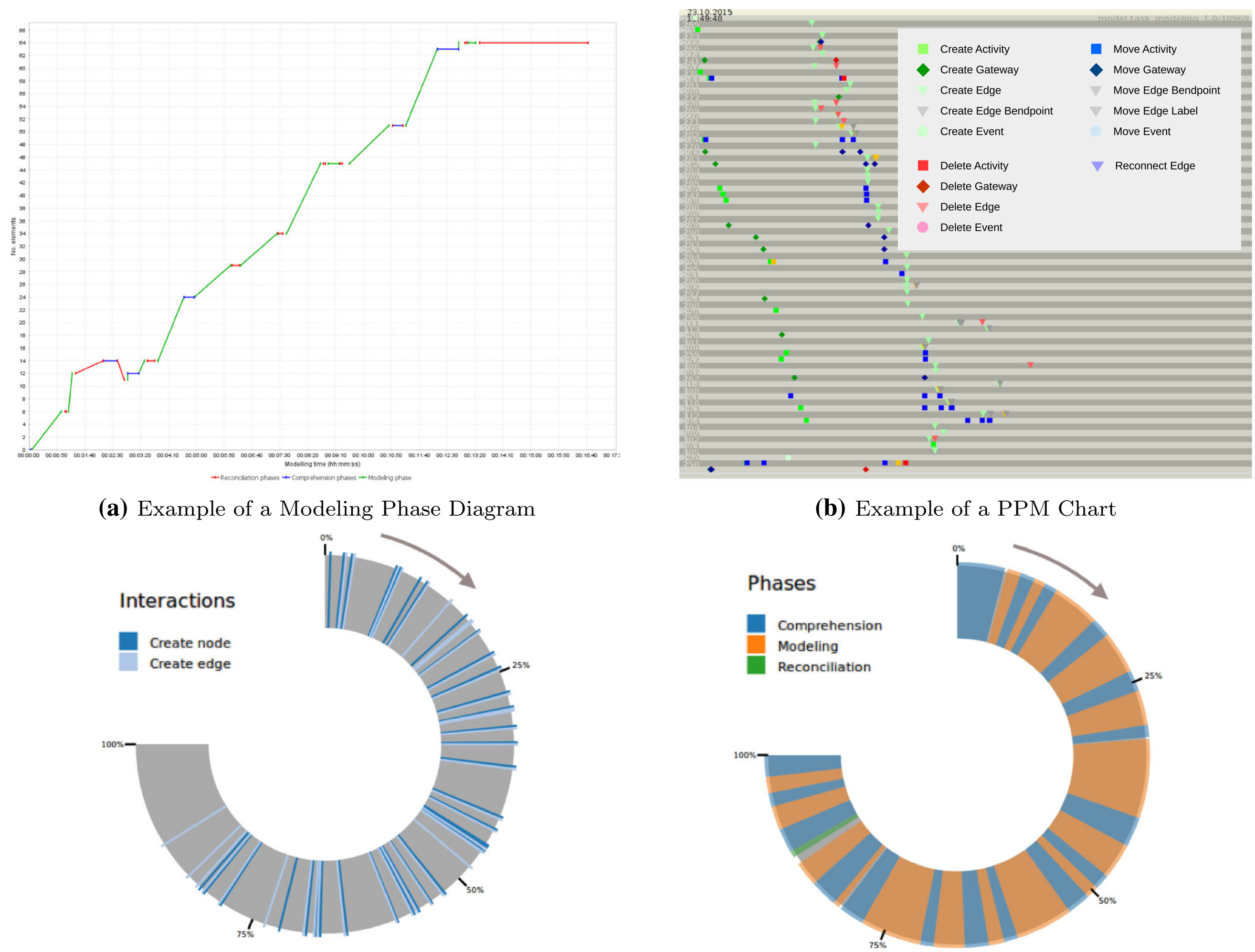
PPM Charts [12] display individual events that represent interactions a modeler has performed during modeling. The set of interactions that are represented consists of create, delete, and move operations applied to any elements of type activity, gateway, event, edge, and edge-bendpoints.

The PPM Charts visualization type conveys information about modeling instances using a two-dimensional placement of symbols. Information about when an event occurred is expressed by the placement of events along the x-axis, which represents the flow of time from left to right. In contrast to this, the y-axis of a PPM Chart separates interactions with different model elements from each other. All interactions that relate to the same model element are placed on the same horizontal line in the chart. The first events on each line thus always represent creation interactions of the new model element. Any later events displayed on the same line, i.e., to the right of the creation event, are interactions performed with the same element. By making use of the two-dimensional diagram space in this way, PPM Charts manage to incorporate both structural and dynamic information about a modeling instance in one diagram.

Figure 1b[1] shows an example PPM Chart diagram together with a legend of the different event symbols that may occur in the diagram. The PPM Chart depicts an example of a process instance where the creation of activities, gateways, and edges is not interwoven. Instead, first activities and gateways are created, and only afterward edges. This exemplifies a modeling behavior which is referred to as aspect-oriented modeling [13].

2.2.3 The Rhythm-Eye visualization

The Rhythm-Eye visualization [20] uses a circular representation to display the temporal progress of a process instance. Events and phases are projected onto a ring structure, rather than onto a linear timeline, according to their time of occurrence during process execution. Figure 1c[1] shows an example where events have been projected onto the ring, more specifically "create activity" and "create edge" interactions. The visualization depicts a process instance in which activities and edges are created in an interwoven manner. Figure 1d,[1] in turn shows a projection of the three modeling phases onto the ring. The visualization illustrates a process instance where comprehension phases are regularly followed by modeling phases with only a brief reconciliation phase toward the end. In order to differentiate between the starting point of a process instance and its end, not the whole 360° is used, but a gap between the start and the end is inserted to distinguish both sides of the displayed process. With this circular projection, it can be expected that rhythmic patterns in process data can be made visible at a glance [20]. This assumption is



(a) Example of a Modeling Phase Diagram

(b) Example of a PPM Chart

(c) Example of a Rhythm-Eye showing Events

(d) Example of a Rhythm-Eye showing Phases

Fig. 1 Existing visualization types for Process Modeling Behavior Analysis

based on the consideration that a ring structure, other than a linear projection, avoids the impression of lesser important periphery areas at the very start and end of the projection space. Hence, a more homogeneous perception of the distribution of events and phases over time would be achieved. The ring projection can also use space on a display device in a more compact and efficient way than a timeline projection. In the same way as multiple lanes of a timeline projection can be placed below each other, the Rhythm-Eye visualization allows to nest multiple rings inside each other (which is not used in our examples).

Our implementation of the Rhythm-Eye visualization is embedded into a software environment which, among others, allows to assign data from different sources to be projected onto the rings, and configure their size and color parameters. The software environment is shown in Fig. 2. The configuration allows for a free choice of combinations of phase data from one process instance (*single instance* analysis) or different ones (*multi-instance* analysis). Types of events and phases can be filtered individually per ring. As a consequence, a variety of configurations becomes possible, in which multiple rings may be used to differentiate between different instances or events, and phases of the same instance are projected onto multiple rings for comparison. Since the configuration of the visualization is performed dynamically and the resulting rendering is immediately shown, the visualization environment also allows for a seamless navigation between these different analytical perspectives. For example, an exploration can begin with a multi-instance analysis that compares phases of one particular kind with each other and then the analyst decides to drill down into the details of one specific instance to compare the individual modeling phases of this instance with each other and later widens the focus again by re-incorporating other instances to investigate a particular constellation discovered. This type of explorative navigation shares similarities with “slicing & dicing” techniques from

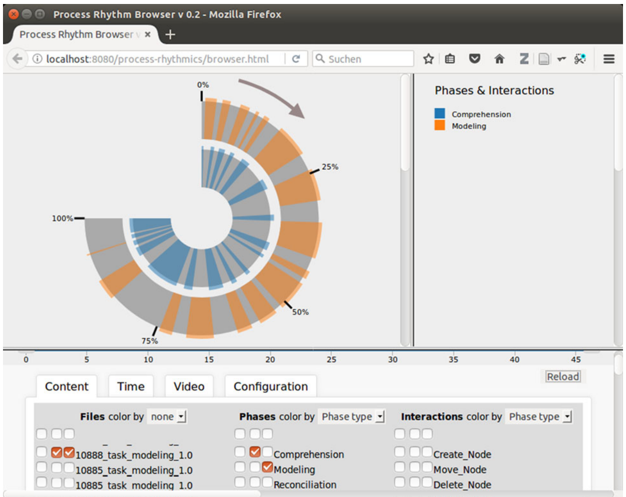


Fig. 2 Visualization environment for configuring and displaying Rhythm-Eyes

Online Analytical Processing (OLAP) approaches in the field of data warehouse analyses [4,14].

In our analysis, we use two configurations of the Rhythm-Eye visualization, one that shows modeling events, and one that shows modeling phases. Dynamic configurations of Rhythm-Eyes over time are not used.

2.3 Eye tracking evaluation of visual analytics

While eye tracking has been widely used to measure the distribution of visual dimensions, e.g., in marketing or psychology, it only recently gained popularity in visualization research [31].

The analysis of the spatiotemporal eye tracking data can be performed following a visual analysis or a statistical approach [31]. The *visual analysis* of eye tracking data is usually based on visualizations like attention maps (i.e., heat maps) and gaze plots illustrating the user's eyepath [24]. The *statistical analysis* of eye tracking data typically starts from the raw gaze data obtained from an eye tracker. The gaze data are then typically preprocessed to detect *fixations* and *saccades*. A fixation is the maintaining of the gaze on a single location, while saccades are the movements between fixations. Typically, fixations are further mapped to areas of interest (AOIs), i.e., specific regions on the stimulus that are of special interest and which allow for a spatial segmentation of the collected data.

According to [36], existing eye tracking metrics can be subdivided into:

- **Fixation-derived metrics** including metrics like fixation duration or fixation count calculated for an area of interest. They can be used to analyze the distribution of attention over different areas of interest.
- **Saccade-derived metrics** including saccade amplitude can be used to analyze the quality of visual clues in a stimulus and the extent of visual searching [31]. Movement direction measures of saccades like saccadic direction, in turn, can be used to analyze in which direction the saccade takes the eye [24].
- **Scanpath-derived metrics** A scanpath refers to the entire sequence of fixations and saccades and can be used to analyze reading strategies [31]. Transition matrices are commonly used to analyze transition patterns between areas of interest [24]. According to a recently conducted study by Kurzhals et al. [31], existing studies in the visualization community mainly focused on the spatial aspect of the recorded gaze data. Temporal aspects of the data, such as AOI sequences, were often completely missing or only partially covered through transition matrices. The analyses conducted as part of this paper clearly go beyond the existing state- of the art and propose a novel way of using process mining technology to analyze AOI sequences.

The first part of this article will continue by introducing a framework for matching visualization types with analysis purposes. The focus in the second part will be on the statistical analysis of eye tracking data using fixation-derived metrics and the visual analysis using saccade-derived measures, as well as scanpath-derived metrics.

3 A framework for matching visualization types with analysis purposes

As a first part of the methodical examination, in this section the *V*isualization *P*urposes (ViP) framework is developed for categorizing data visualizations according to the purposes they fulfill in domain-specific data analysis settings. Once the framework is configured for a given analysis domain, data visualizations can be chosen based on analysis questions that address different aspects of data to be analyzed.

Based on the concepts addressed by the analysis questions, the framework provides systematic guidance for determining which data visualizations are able to serve conceptual analysis interests. This allows to identify candidates of visualization types that satisfy the information needs specific to the analysis domain.

3.1 A method for configuring the ViP framework

The procedure for setting up the ViP framework for a given analysis domain consists of five essential configuration steps. These steps serve to configure a multidimensional categorization framework, which subsequently allows to map from analysis questions to visualization candidates. In the following, the general configuration procedure is described. It will be applied to the domain of Process Modeling Behavior Analysis (cf. Sect. 2.2) in the upcoming subsections.

1. *Collect domain-relevant questions by performing a domain analysis*
   As with every domain-specific setting, a systematic analysis of the concepts, actors, goals, and procedures that underlie the domain needs to be initially performed. For the purpose of setting up the ViP framework, this analysis has to focus on questions that are to be answered in the domain's analysis scenarios. As a result, the questions should be articulated together with the identification of stakeholders who have an interest in answering the question, and they should unambiguously refer to specific objects of interest, i. e., characteristics of relevant domain concepts, which are addressed by each question. In general, the questions can be written in natural language. However, depending on the complexity of the domain, it may make sense to further refine this step by prescribing a specific format in which the questions should be stated, in order to better point out the involved stakeholders and objects of interest. For our purposes, since we operate with a manageable set of eight questions, we refrain from prescribing a fixed format and instead adhere to natural language which points out all relevant aspects of analysis questions. Section 3.2.1 contains the questions we collected for the Process Modeling Behavior Analysis domain.
2. *Identify independent dimensions that are addressed by the questions*
   Once a set of multiple questions is available, the total set of involved stakeholders, as well as characteristics of objects of interest in the domain addressed by all questions, can be extracted. This allows to identify common aspects addressed by the questions, along which they can be grouped, as well as varying aspects, which potentially can serve as categories that distinguish between the questions. E.g., it may turn out that questions can be grouped according to the same domain elements they address, but different characteristics of these elements are put in focus by different questions. In this case, dimensions describing the domain element's characteristics would qualify as a candidate for categorizing the analysis questions of the domain. On the contrary, other sets of analysis questions may rather address diverse kinds of domain elements, but each one dealing with identical characteristics of these elements. In this case, the distinction between different domain element types could serve as a categorization dimension for the questions.
   As a result of this step, a number of aspects addressed by analysis questions are identified, which can serve as category dimensions along which the different analysis questions of the domain can be either grouped together or be differentiated. These dimensions can then be used as axes of a categorization space that is spanned by the ViP framework, by naming them appropriately, and label the axis intercepts with the category values that are part of this dimension. In case of a large number of identified dimensions, it may make sense to refine this configuration step with a phase of systematically assessing the relevance of each identified dimension, so it can be decided whether they are to be included in the framework or not. For the Process Modeling Behavior Analysis domain, we identified three dimensions in Sect. 3.2.2 which were used for the framework configuration.
3. *Locate analysis questions in the framework*
   Subsequently, the questions identified in step 1 are placed inside the categorization space created in step 2. A question is understood as being present at a location in the framework, when it addresses those aspects that are represented by the axis intercepts of a categorization dimension. This means that the analysis questions get associated with symbolic, nonnumerical, coordinates, such as (cat1="A," cat2="B," cat3="C"). Assuming a number of dimensions not larger than three, and a given ordering of the axis intercepts values of each dimension, coordinates of this kind can be projected into a human-perceivable Euclidean space, as it is exemplified in Fig. 3.
   The result of this step does not necessarily have to be presented visually as shown in the figure, but could be manifested, e.g., as a table which lists all coordinates and associated questions. Such a representation of the locations of question is applicable to spaces with any number of dimensions and is in principle not limited to a maximum number of questions and aspects along which they are categorized. It would also easily allow for questions to be associated with multiple coordinates at the same time, in cases where some of the identified analysis questions match multiple category aspects simultaneously.
   We locate the analysis questions of the Process Modeling Behavior Analysis domain in Sect. 3.2.3.
4. *Locate visualization types in the framework*
   In analogy to the previous step in which questions have been located in the categorization space, also visualization types can be associated with places in the space. In order to achieve this, each visualization type is examined whether it is capable of visualizing information that matches each of the different category characteristics represented by the coordinates of the space. Like with the positioning of questions, visualization types can be located at more than one single point in the space.

They may cover lines, planes, or hyperplanes in the categorization space, which is the case when visualization types are capable of visualizing any of the category characteristics that belong to one category dimension of the space.
To go into more details, general visualization types can be distinguished with regard to their configurations, i.e., the same visualization type may be configured to map different kinds of information onto its visual elements. For example, a bar chart can be configured to display diverse sets of figures from different categories. The step of locating visualization types in the category space may distinguish between configurations of this kind, if the visualization types can be configured in multiple ways to address the analysis questions from the configured domain.
The localization of visualization types according to the Process Modeling Behavior Analysis domain is done in Sect. 3.2.4.

5. *Apply the framework by finding matching visualizations at the respective spatial position of a question*
The configured framework allows for finding matches between analysis questions and visualization types in the category space, by looking at the colocations where both individually have been placed. Beginning with an analysis question to find visualization types for, all coordinates in the space at which the question is located are listed, and subsequently all visualization types that are located at these coordinates can be derived. The result is the set of those visualization types which may be applicable for creating insightful visualizations that answer the analysis question. For the Process Modeling Behavior Analysis domain, the framework is applied in Sect. 3.2.5.

3.2 Configuring the ViP framework for the process modeling behavior analysis domain

In the following, the general procedure for configuring the ViP framework (cf. Sect. 3.1) will be applied to the Process Modeling Behavior Analysis domain. For each of the five described steps, the concrete application to the domain is explained in one separate section.

3.2.1 Collect domain-specific analysis questions

In order to select suitable visualization types for the analysis domain of Process Modeling Behavior Analysis (PMBA) (cf. Sect. 2.2), we associate analysis questions of the domain with a selection of visualization types that are suitable to answer these questions. Objects of interests in the PMBA domain in general are characterized by the nature of time-related data that are collected during process modeling activities. As discussed in Sect. 2.2, in the first place these consist of a stream of raw action events that are recorded as representations of the actions modelers perform when creating or modifying process models, e.g., when new activity nodes are added to a process model, or when edges that connect activities are modified or deleted from a process model. These individual events can be aggregated to phase data. Both aggregation levels of data characterize the main objects of interests examined in PMBA, and analyses in general revolve around questions about how different event and phase types are interrelated, or how they are distributed over time.

As typical representatives from the PMBA domain that address these epistemological interests, we selected eight analysis questions which are listed in the following:

- Are activities and edges created in an intertwined way, or does the user first create activities and then add edges?
- Are comprehension phases regularly followed by modeling phases?
- Is there a long reconciliation phase at the end?
- Are there a lot of move operations at the end of a modeling process?
- Which participant creates and immediately deletes elements more frequently?
- Which participant alternates most often between modeling and reconciliation phases?
- Which participant does most reconciliation phases continuously throughout the session?
- Which participant performs most delete operations?

By later locating the analysis questions inside the multidimensional categorization framework in Sect. 3.2.3, it becomes possible to systematically differentiate between purposes that are addressed by each question.

3.2.2 Identify categorization dimensions in the Process Modeling Behavior Analysis domain

Resulting from the methodological considerations in the previous section, a concrete instantiation of a ViP framework for analysis questions in the PMBA domain can now be explicated. This is done in the form of a three-dimensional cube which allows to locate each concrete domain-specific question of the PMBA domain inside the ViP framework. To form the cube, three independent dimensions that are addressed by the identified questions (cf. Sect. 3.2.1) are selected from the total set of questions. They are discussed in the following.

Action dimension (*What?*) Depending on the focus of an analysis, individual events that occur during the modeling process, or entire phases of modeling activities, can be in focus. It is also possible to address both together by an analysis question. A question that addresses both aspects is, e.g., “Can a calm versus a hectic modeling style be distinguished?”. Looking exclusively at event information alone is done, e.g., by asking “Does the participant often create and immediately afterward delete elements again?”.

Timing dimension (*When?*) Analysis questions can additionally be distinguished with respect to whether they address characteristics of entire PPM instances, or whether they rather focus locally on the patterns of interplay among modeling activities. A question that relates to an entire PPM instance is, e.g., “Does the participant spend more time with modeling in the second half of the experiment time than in the first?”. In contrast, the interplay of phases is addressed by the question “Are comprehension phases regularly followed by modeling phases?”.

Instance dimension (*How many?*) Some of the analysis questions focus exclusively on data referring to a single PPM instance. These are, e.g., questions starting with “Did the participant...”. Other questions ask for comparing behavior of multiple instances with each other, e.g., “Can different modeling styles of participants be recognized?”. It is thus one fundamental criterion to distinguish between analysis scenarios that operate on collected data from single sessions, and scenarios that work on data from multiple sessions.

Based on the identified dimensions, a three-dimensional space is spanned, inside which both analysis questions and available visualization types will be located in the following steps. By doing so, the framework becomes a conceptual glue between questions that represent analysis purposes, and visualizations.

3.2.3 Locate analysis questions inside the categorization framework

To exemplify the use of the ViP framework, we list one example question for each combination of dimension characteristics and localize it in the three-dimensional space that is spanned by the previously configured ViP framework.

We keep this rather informal by first describing the locations of questions verbally and then deriving the visual representation as shown in Fig. 3 from this. The questions identified in Sect. 3.2.1 are numbered `Q1` to `Q8`, and their location inside the framework is indicated by the labels shown inside the cube at the approximate locations which resemble the Instance, Axis, and Timing characteristics addressed by each question.

In the following, the example questions identified in Sect. 3.2.1 are listed with reference to their locations in the ViP framework, written in italics as a list of three dimension characteristics that are addressed by each question:

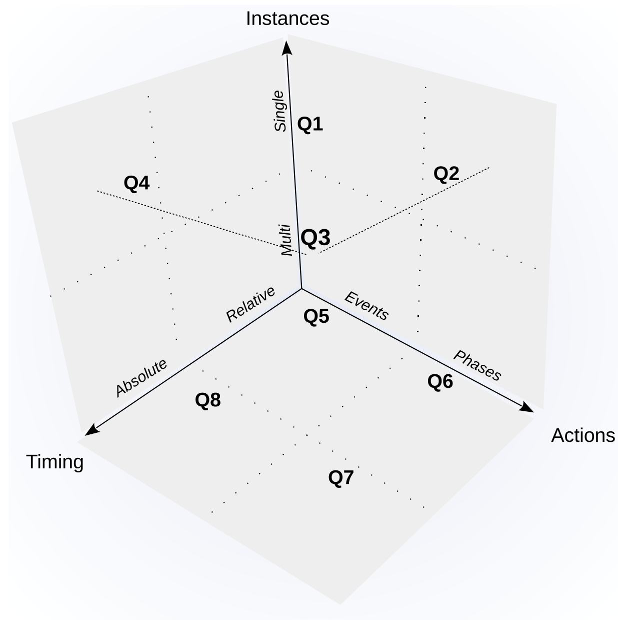


Fig. 3 Multidimensional categorization framework for analysis questions regarding Process Modeling Behavior Analysis

`Q1` Are activities and edges created in an intertwined way, or does the user first create activities and then add edges? *(Single instance, Event actions, Relative timing)*

`Q2` Are comprehension phases regularly followed by modeling phases? *(Single instance, Phase actions, Relative timing)*

`Q3` Is there a long reconciliation phase at the end? *(Single instance, Phase actions, Absolute timing)*

`Q4` Are there a lot of move operations at the end of a modeling process? *(Single instance, Event actions, Absolute timing)*

`Q5` Which participant creates and immediately deletes elements more frequently? *(Multiple instances, Event actions, Relative timing)*

`Q6` Which participant alternates most often between modeling and reconciliation phases? *(Multiple instances, Phase actions, Relative timing)*

`Q7` Which participant does most reconciliation phases continuously throughout the session? *(Multiple instances, Phase actions, Absolute timing)*

`Q8` Which participant performs most delete operations? *(Multiple instances, Event actions, Absolute timing)*

3.2.4 Locate visualization types inside the categorization framework

The categorization dimensions identified in Sect. 3.2.2 allow to make justified statements about the capabilities of different visualization types to be supportive in answering analysis questions associated with the PMBA domain. Specifics of the analysis purposes addressed by the questions can now be represented conceptually by means of dimension characteristics of the ViP framework.

At the same time, the visual expressiveness of different available visualization types can be systematically expressed with respect to the question whether a visualization type is capable of answering analysis questions from the domain. This means the ViP framework allows to map from conceptually expressed purposes to visualizations. In order to achieve this mapping, visualization types can be described in terms of subspaces of the framework cube. A visualization type may or may not be able to visually express some of the dimension characteristics of the framework. Those areas of the cube space which represent combinations of dimension characteristics that a visualization type is able to express, can be considered as the subspace of the cube which contains all those analysis questions that can be addressed by the respective visualization type.

In addition to previous examinations on the Rhythm-Eye visualization type [20,44] (cf. Sect. 2.2.3), we now add the Modeling Phase Diagram (MPD) [8] (cf. Sect. 2.2.1) and PPM Chart [12] (cf. Sect. 2.2.2) to the set of available visualization types, to provide a comprehensive coverage of existing visualization types for PMBA. We then examine which dimension characteristics of the ViP framework can be represented visually by each of these visualization types.

The results of matching visualization types to dimension characteristics are summarized in Table 1. All dimension characteristics are flattened onto the vertical axis of the table. The Rhythm-Eye visualization covers all dimension characteristics and consequently can be used to answer all analysis questions, i.e., `Q1`–`Q8`. This is possible because the Rhythm-Eye visualization type offers a wide range of configuration options by which it can be more flexibly configured than the other two. MPDs support all timing characteristics, but are only suitable to analyze single instances and questions concerning phases (i.e., `Q2` and `Q3`). PPM Charts cover all timing characteristics and are suitable to analyze single instances in respect to questions concerning events (i.e., `Q1`, `Q4`). In summary, the mapping reveals that for `Q5`–`Q8` only one visualization type, i.e., the Rhythm-Eye, exists, while `Q1`–`Q4` can be answered by two alternatives each.

Table 1 Results of matching visualization types to dimension characteristics

Characteristic	MPD	PPM C.	Rhythm-E.
Instance			
Single	✓	✓	✓
Multi			✓
Action			
Events		✓	✓
Phases	✓		✓
Timing			
Relative	✓	✓	✓
Absolute	✓	✓	✓

3.2.5 Apply the ViP framework

The described approach makes it possible to reflect about the purposes of visualizations on a conceptual level purely in words. The analysis questions are describing specific analysis purposes, and overlaps between visualization types and questions located in the framework provide a formal mapping between described purposes and visualization capabilities. Methodologically, this avoids the need to talk about individual graphical properties of visualizations in combination with conceptually high-level analysis purposes, which could become utterly complex and would not provide an appropriate level of explanatory abstraction.

To apply the configured ViP framework to any of the Process Modeling Behavior Analysis questions, the location of the respective question is to be determined as the dimension characteristics of the question's conceptual coordinates in the framework space. Any visualization types which have been placed at the same position in step 4 (cf. Sect. 3.2.4) are possible candidates for using them as analytical tools to answer the question.

The corresponding visualization types from the PMBA domain that are candidates to answer each of the analysis questions are listed in the following.

- `Q1`: *PPM Chart, Rhythm-Eye*
- `Q2`: *Modeling Phase Diagram, Rhythm-Eye*
- `Q3`: *Modeling Phase Diagram, Rhythm-Eye*
- `Q4`: *PPM Chart, Rhythm-Eye*
- `Q5`–`Q8`: *Rhythm-Eye*

4 Eye tracking for evaluating the suitability of alternative visualization types

This section outlines our method for testing alternative visualization types (cf. Sect. 4.1) and explains how we applied the proposed procedure in the context of Process Modeling Behavior Analysis (cf. Sect. 4.2).

4.1 Procedure for testing competing visualization types

The application of the ViP framework outlined in Sect. 3.2 is able to identify alternative visualization types for different spatial positions in the framework. If only one visualization at a certain spatial position exists, then this visualization can be selected for further analysis. For example, the recommendation would be to use the Rhythm-Eye visualization for questions related to multi-instances.

If more than one visualization is available, we suggest a systematic and data-driven way to support the selection of visualization types as well as their improvement. Similar to a setting where A/B testing is conducted to compare variants, our setting aims to find the variant that is most effective for a specific analysis purpose. For example, in the context of Process Modeling Behavior Analysis, the question is whether to use Rhythm-Eye for Events or PPM Charts for questions related to events and Rhythm-Eye Phases versus MPD for questions related to phases. This can be more generally translated into the following research question **RQ1: Which visualization type is most effective for a specific analysis purpose?** To answer this research question, we suggest to conduct an experiment which allows for group comparisons between alternatives covering all relevant dimension characteristics identified in the framework. Such an approach is rather data-driven than theory-driven and therefore refrains from formulating hypotheses.

The literature about software experiments provides various design guidelines for setting up an experiment [2,27,30,40,45]. Our setting requires the experiment to be designed as a two-factor experiment investigating the effects of two factors (i.e., visualization type and analysis purpose) in terms of their effectiveness (i.e., response variable) and operationalized, e.g., as answering correctness, answering time, total fixation time on graph, total fixation count. Such an experiment design allows us to conduct group comparisons between variations of a factor called *factor levels* for different analysis purposes (e.g., Rhythm-Eye for Events vs. PPM Chart for the action dimension event characteristics; Rhythm-Eye Phases vs. MPD for the action dimension phase characteristics). The experiment is then rolled out to different participants, who are asked to conduct different comprehension tasks (answering questions for different experimental objects, i.e., concrete visualizations) using the different visualization types for different analysis purposes. Furthermore, we suggest to design the experiment as a balanced experiment with repeated measurement. This design is particularly suitable for comparing design artifacts [30] (i.e., in our case different visualization types) for different purposes. An experiment design is denoted as *balanced* if all factor levels are used by all participants of the experiment. This enables repeated measurements and thus the collection of more precise data, since every subject generates data for every treated factor level. Moreover, this choice of design has the advantage of being better able to deal with heterogeneous backgrounds of participants. Choosing a design with repeated measurements involves the risk for *learning effects*. Randomizing the ordering in which the comprehension tasks are presented can mitigate the risk of learning effects and systematic biases favoring a particular factor level. As another measure to minimize learning effects, we suggest to base the comprehension tasks referring to the same question, but different factor levels on different data (e.g., different modeling sessions in our context). This ensures that participants cannot answer question by memory, but have to engage in the comprehension task to answer correctly.

To identify the most suitable visualization types, group comparisons are conducted comparing competing visualization types for each of the analysis purposes in terms of their effectiveness (e.g., Rhythm-Eye for events vs. PPM Chart for the action dimension event characteristics; Rhythm-Eye for events vs. PPM Chart for the action dimension phase characteristics).

If the group comparison between different competing visualizations does not reveal significant differences, then the user can choose either visualization types. To minimize the number of used visualization types, one might choose the visualization type with the highest number of analysis purposes supported.

If the framework is used by a visualization provider, then an analysis of reading patterns can be additionally conducted to better understand why a particular visualization type is better than another or why a certain visualization type is not effective. Here, we expect to be able to identify reading strategies distinct to each visualization type. By comparing reading patterns, the quantitative results of the group comparisons can be complemented with explanations which provide important input for improving a particular visualization type. This results in research question **RQ2: What reading strategies can be identified for a given visualization type?**

4.2 Testing competing visualization types for process modeling behavior analysis

To analyze the suitability of the three visualization types in the context of Process Modeling Behavior Analysis, we conduced an eye tracking session where we asked 15 novice participants (i.e., students from the Technical University of Denmark as well as academics) to answer the analysis questions introduced in Sect. 3.2.1, after being introduced to Process Modeling Behavior Analysis.

The matching of visualization types to dimension characteristics (cf. Table 1) revealed that for both the action dimension and the timing dimension alternative visualization types could be identified. For visualizing data related to multi-instances, in turn, only a single visualization, i.e., the Rhythm-Eye, exists. In the empirical study, we therefore focused on single-instance characteristics only.

Design We designed the experiment as a balanced two-factor experiment with repeated measurements (cf. Sect. 4.1).

For each of the dimension characteristics of the framework where more than one alternative visualization type exists, we presented participants with the corresponding analysis questions, i.e., `Q1`–`Q4` (cf. Sect. 3.2.1). To make the experiment balanced, participants had to perform several sense-making tasks for each analysis question (one for each visualization type).

By designing the experiment as balanced experiment with repeated measurement where each participant is exposed to each factor level (Rhythm-Eye for events and PPM Chart), we could reach an N high enough for conducting a quantitative analysis (i.e., it is usually well accepted to conduct a quantitative group comparison with an $N \geq 25$ for each group).

The stimulus presented to the participants included the question text on the top of the screen, the graph depicting the visualization itself, and the legend.[2] At the end of the session, we asked the participants for feedback and their perception regarding the different visualization types.

Operationalization For answering research question RQ1, we operationalized the effectiveness of a visualization using traditional performance measures like answer correctness and answering time along with commonly used fixation-derived eye tracking metrics like fixation duration and fixation count [36]. Using fixation-derived metrics, the attention needed for answering analysis questions using different visualization types can be compared. In the light of existing literature, we expect that visualizations that are cognitively more effective yield a lower fixation count [19] and a lower overall fixation duration [37]. For answering research question RQ2, we will use a combination of saccade-derived measures and scanpath-derived measures as basis.

Instrumentation To collect the data for answering the two above-mentioned research questions (i.e., RQ1 and RQ2), the participants' eye movements were tracked while answering the analysis questions using a Tobii Pro TX300 eye tracker (cf. Fig. 4). For designing the eye tracking session, for preprocessing the eye tracking data and defining areas of interest, Tobii Pro Studio 3.4 was used.

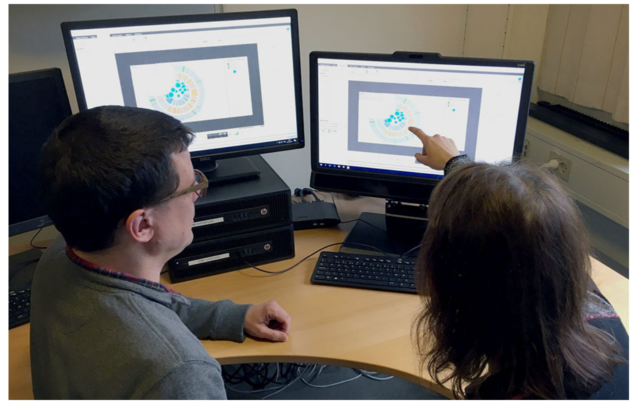


Fig. 4 Eye tracking machine during the analysis of the results

5 Results

5.1 Comparison of visualization types

This section aims to answer the research question RQ1 (i.e., which visualization type is more effective for answering the posed analytical question). Please note that the matching of visualization types to dimension characteristics (cf. Table 1) revealed that for the multi-instance characteristic of the instance dimension, only a single visualization, i.e., the Rhythm-Eye, exists. Thus, our analysis will focus on the *action* dimension and the *timing* dimensions where several alternative visualizations exist.

5.1.1 Analysis procedure

To answer RQ1, we conducted group comparisons using the nonparametric Mann–Whitney U test using SPSS version 19.

For analyzing the action dimension, we need to make the distinction between events and phases. The analysis of event characteristics is based on questions Q1 and Q3 which both posed questions concerning the event dimension and compares Rhythm-Eye and PPM Chart. The analysis of phases characteristics, in turn, is based on questions Q2 and Q4, which posed questions concerning the phase dimension. The analysis compares Rhythm-Eye and MPD. With 15 participants[3] this gives us 30 data points for each of the visualizations and action characteristic (i.e., 30 data points for Rhythm-Eye for events and PPM Chart as well as Rhythm-Eye for phases and MPD). For analyzing the time dimension, however, we further needed to split the data between relative and absolute time. More specifically, we compared Rhythm-Eye and PPM Chart based on the data of Q1 (relative) and Q3 (absolute) and Rhythm-Eye and MPD for Q2 (relative) and Q4 (absolute). After such split, we ended up with 15 data points per configuration, as commented in Sect. 5.1.3.

As outlined in Sect. 4, we considered answer correctness, answering time, fixation duration, and fixation count as dependent measures. We operationalized answer correctness[3] as a Boolean variable that could either be true (i.e., correct) or false (i.e., incorrect). Answering time was measured as the time needed to answer a question. To calcu-

[2] All presented stimuli are available at https://doi.org/10.5281/zenodo.1419598.

[3] All data about participants' background, expected tasks' answers, and answers accuracy are available at https://doi.org/10.5281/zenodo.1419598.

late fixation duration and fixation count, we defined an area of interest including the graph (excluding the question text and the legend) and used Tobii Studio 3.4 to obtain these metrics.

5.1.2 Results for the *Action* dimension

Tables 2 and 3 show the descriptive statistics for the phase characteristics and the event characteristics, respectively. Tables 4 and 5, in turn, show the results of the Mann–Whitney U tests for event characteristics and phases characteristics, respectively.

Events The results of our statistical analysis show that the Rhythm-Eye visualization outperformed the PPM Charts for questions concerning the event dimension in terms of answering correctness ($U = 360$, $p = 0.01$, < 0.05 two-tailed), total fixation duration on the graph ($U = 231$, $p = 0.001$, < 0.05 two-tailed), and fixation count on the graph ($U = 185.5$, $p = 0.000$, < 0.05 two-tailed). Differences in terms of answering time were not significant ($U = 320$, $p = 0.055$, > 0.05 two-tailed).

Phases Moreover, our results show that the Rhythm-Eye visualization outperformed the MPD visualization for questions concerning the phase dimension in terms of answering correctness ($U = 315$, $p = 0.001$, < 0.05 two-tailed), total fixation duration on the graph ($U = 299$, $p = 0.026$, < 0.05 two-tailed), and fixation count on the graph ($U = 280$, $p = 0.012$, < 0.05 two-tailed). Differences in terms of answering time were not significant ($U = 409$, $p = 0.544$, > 0.05 two-tailed).

5.1.3 Results for the *Timing* dimension

As previously mentioned, for this investigation we could rely only on 15 data points per configuration. With such number of observations (below the usually accepted $N \geq 25$), we cannot draw absolute conclusions. We report this investigation nonetheless, in order to better illustrate the followed methodology, which is the main contribution of the paper.

Relative timing Tables 6 and 7 show the results of the Mann–Whitney U tests of relative timing questions for event characteristics and phases characteristics, respectively. The corresponding descriptive statistics are available in Tables 10 and 11 in “Appendix.” Our results show that the Rhythm-Eye visualization outperformed the PPM Charts for relative timing questions in terms of answering time ($U = 63$, $p = 0.041$, < 0.05 two-tailed), total fixation duration on the graph ($U = 55$, $p = 0.016$, < 0.05 two-tailed), and fixation count on the graph ($U = 44$, $p = 0.004$, < 0.05 two-tailed). Differences in terms of answering correctness were not significant ($U = 82.5$, $p = 0.217$, > 0.05 two-tailed). Moreover, our results show that the Rhythm-Eye visualization outperformed the MPD visualization for relative timing questions in terms of answering correctness ($U = 52.5$, $p = 0.011$, < 0.05 two-tailed), total fixation duration on the graph ($U = 48$, $p = 0.007$, < 0.05 two-tailed), and fixation count on the graph ($U = 42$, $p = 0.003$, < 0.05 two-tailed). Differences in terms of answering time were not significant ($U = 93$, $p = 0.436$, > 0.05 two-tailed).

Table 2 Descriptive statistics for event characteristics

	N	Rhythm-Eye for events				PPM chart			
		Min	Max	Median	SD	Min	Max	Median	SD
Answer correctness	30	1.000	1.000	1.000	0.000	0.000	1.000	1.000	0.407
Answering time	30	6.580	87.630	20.705	17.894	7.880	105.190	27.015	27.309
Tot. fixation duration graph	30	1.160	27.910	6.455	6.623	3.850	55.360	13.055	15.232
Fixation count graph	30	5.000	108.000	30.000	23.891	20.000	266.000	59.500	66.165

Table 3 Descriptive statistics for phase characteristics

	N	Rhythm-Eye for phases				Modeling phase diagram			
		Min	Max	Median	SD	Min	Max	Median	SD
Answer correctness	30	1.000	1.000	1.000	0.000	0.000	1.000	1.000	0.466
Answering time	30	7.500	81.640	15.940	13.878	3.060	100.110	15.355	21.221
Tot. fixation duration graph	30	1.910	26.490	4.930	4.944	2.430	60.410	8.190	12.068
Fixation count graph	30	8.000	117.000	23.500	21.693	9.000	259.000	43.500	48.742

Table 4 Mann–Whitney *U* test for event characteristics

Rhythm-Eye for events versus PPM		
	Mann–Whitney *U*	Asymp. sig. (2-tailed)
Answer correctness	360.000	**0.010**
Answering time	320.000	0.055
Tot. fixation duration graph	231.000	**0.001**
Fixation count graph	185.500	**0.000**

Bold values are below the significance threshold of 0.05. They support the assumption that the results from two different experiments indeed originate from different statistical distributions

Table 5 Mann–Whitney *U* test for phase characteristics

Rhythm-Eye for phases versus modeling phase diagram		
	Mann–Whitney *U*	Asymp. sig. (2-tailed)
Answer correctness	315.000	**0.001**
Answering time	409.000	0.544
Tot. fixation duration graph	299.000	**0.026**
Fixation count graph	280.000	**0.012**

Bold values are below the significance threshold of 0.05. They support the assumption that the results from two different experiments indeed originate from different statistical distributions

Table 6 Mann–Whitney *U* test for event characteristics and relative timing

Rhythm-Eye for events versus PPM chart		
	Mann–Whitney *U*	Asymp. sig. (2-tailed)
Answer correctness	82.500	0.217
Answering time	63.000	**0.041**
Tot. fixation duration graph	55.000	**0.016**
Fixation count graph	44.000	**0.004**

Bold values are below the significance threshold of 0.05. They support the assumption that the results from two different experiments indeed originate from different statistical distributions

Table 7 Mann–Whitney *U* test for phase characteristics and relative timing

Rhythm-Eye for phases versus modeling phase diagram		
	Mann–Whitney *U*	Asymp. sig. (2-tailed)
Answer correctness	52.500	**0.011**
Answering time	93.000	0.436
Tot. fixation duration graph	48.000	**0.007**
Fixation count graph	42.000	**0.003**

Bold values are below the significance threshold of 0.05. They support the assumption that the results from two different experiments indeed originate from different statistical distributions

Table 8 Mann–Whitney *U* test for event characteristics and absolute timing

Rhythm-Eye for events versus PPM chart		
	Mann–Whitney *U*	Asymp. sig. (2-tailed)
Answer correctness	97.500	0.539
Answering time	86.000	0.285
Tot. fixation duration graph	49.000	**0.008**
Fixation count graph	40.500	**0.002**

Bold values are below the significance threshold of 0.05. They support the assumption that the results from two different experiments indeed originate from different statistical distributions

Table 9 Mann–Whitney U test for phase characteristics and absolute timing

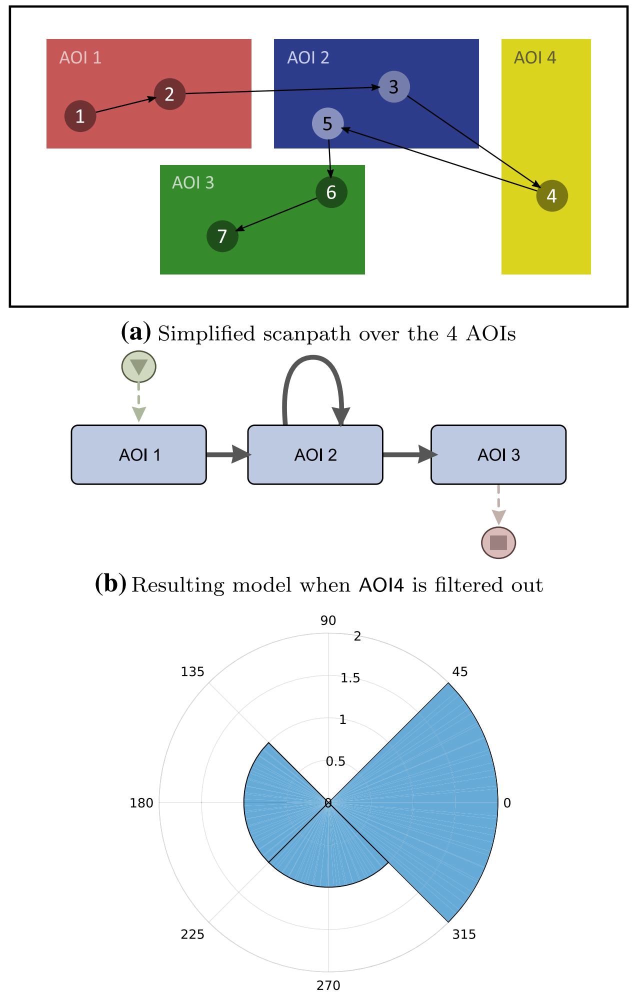
Rhythm-Eye for phases versus modeling phase diagram		
	Mann–Whitney U	Asymp. sig. (2-tailed)
Answer correctness	105.000	0.775
Answering time	69.000	0.074
Tot. fixation duration graph	103.000	0.713
Fixation count graph	104.500	0.744

Absolute timing Tables 8 and 9 show the results of the Mann–Whitney U tests of absolute timing questions for event characteristics and phases characteristics, respectively. The corresponding descriptive statistics can be found in Tables 12 and 13 in "Appendix." Our results show that the Rhythm-Eye visualization outperformed the PPM Charts for absolute timing questions in terms of total fixation duration on the graph ($U = 49$, $p = 0.008$, < 0.05 two-tailed), and fixation count on the graph ($U = 40.5$, $p = 0.002$, < 0.05 two-tailed). Differences in terms of answering correctness ($U = 97.5$, $p = 0.539$, > 0.05 two-tailed) and answering time ($U = 86$, $p = 0.285$, > 0.05 two-tailed) were not significant.

Our results show no significant differences between the Rhythm-Eye visualization and the MPD visualization regarding any of the dependent variables.

5.1.4 Discussion

Based on our results, it can be concluded that for questions concerning both events and phases, the Rhythm-Eye visualization outperforms its alternatives like PPM Charts and Modeling Phase Diagrams. Though results concerning the time are not completely reliable, the Rhythm-Eye visualization also outperforms its alternatives for questions on relative timing for at least some of the dependent variables; for questions concerning absolute timing and phase characteristics, Modeling Phase Diagrams might constitute a viable option to the Rhythm-Eye visualization.



(a) Simplified scanpath over the 4 AOIs

(b) Resulting model when AOI4 is filtered out

(c) Rose Plot of the scanpath in Fig. 5a

Fig. 5 Example of AOIs, with a scanpath over them and the minded model after filtering out AOI4. The Rose Plot of the scanpath is reported as well

5.2 Analysis of reading patterns

This section addresses RQ2, i.e., the identification of reading patterns for the different visualization types.

5.2.1 Analysis procedure

In order to analyze the reading patterns, we started from a scanpath of fixations and saccades (cf. Sect. 2.3). Additionally, we identified relevant areas of interest on top of the different visualizations (details regarding the defined AOIs are explained in Sect. 5.2.2). Figure 5a depicts an example visualization with four areas of interest and a scanpath plotted on top. Each filled circle represents a fixation. The number contained in the circle refers to the temporal ordering of the fixations. Edges refer to saccades connecting contiguous fixations.

As described in Sect. 2.3, scanpath-derived measures can be used to analyze reading strategies. In particular, transition matrices are commonly used to analyze transition patterns between areas of interest [24]. However, respective techniques focus on the spatial aspect of the recorded gaze data, while temporal aspects can only be partially covered [31]. In particular, transition matrices are not suitable to depict the starting/ending points of the reading. Additionally, reading such matrices can be difficult and it is not easy to elicit the main reading patterns. Instead, we propose a novel process mining-based technique [1] that is able to overcome these limitations of transition matrices.

Similar to transition matrices, our focus is on transitions between areas of interest. Therefore, in the first step, starting from the sequence of fixations, we mapped each fixation to the corresponding AOI. This results into the sequence of AOIs where the subject focused.[4] In the example of Fig. 5a, such sequence is:

$$S = \langle \mathsf{AOI1}, \mathsf{AOI1}, \mathsf{AOI2}, \mathsf{AOI4}, \mathsf{AOI2}, \mathsf{AOI3}, \mathsf{AOI3} \rangle.$$

In the second step, we need to merge contiguous elements which are referring to the same AOI to remove fixations within the same area of interest. Starting from S, we obtain:

$$S' = \langle \mathsf{AOI1}, \mathsf{AOI2}, \mathsf{AOI4}, \mathsf{AOI2}, \mathsf{AOI3} \rangle.$$

This sequence can be interpreted as an event log [1,25], where each element refers to the "activity" performed by the user while focusing on the given area of interest. Using state-of-the art process mining [1] techniques, the typical flow of the activities can be mined, which in our context can be interpreted as the typical way of reading a given visualization. Several control-flow discovery algorithms are available in the literature, and several implementations (both academic, open source, or commercial tools) can be used. Due to the exploratory nature of our investigation, we decided to use the tool Disco[5] which allows quick filtering and parameters tuning operations.

Using process mining tools like Disco, it is possible to filter activities, i.e., to only depict the transitions between a subset of the AOIs, e.g., AOI1–AOI3 while ignoring AOI4. In our context, for example, the area of interest containing the question text should be filtered, since the question text has to be read independently of the visualization and including this area of interest would just lead to unnecessarily complex models. To achieve that, we need to filter S' to discard events referring to areas of interest we are not interest in, i.e., AOI4. The new sequence of events is:

$$S'' = \langle \mathsf{AOI1}, \mathsf{AOI2}, \mathsf{AOI2}, \mathsf{AOI3} \rangle$$

We have now two contiguous events referring to the same AOI (i.e., AOI2). In this case, however, we do not want to merge them, otherwise we would alter the meaning of the transitions. If we enter this sequence of events in Disco we obtain the model reported in Fig. 5b. It is possible to see the starting point (depicted as light green circle), the end point (as light red circle) and the 3 AOIs. Additionally, we can see the transitions from AOI1 to AOI2 and then to AOI3. We can identify a self-loop in AOI2. In this context, the interpretation of a self-loop is different from the standard interpretation in process mining: in this case it indicates that the focus on the given area of interest was interrupted by fixations on some AOI, not represented in the model (AOI4 in our example).

We complement this process mining-based approach to analyze transitions between areas of interest with saccade-derived measures, more specifically, the saccadic direction, which allows to analyze in which direction a saccade takes the eye. Directions for all the saccades in selected AOIs are then depicted as angular histograms also denoted as Rose Plot [24]. In contrast to the process models obtained from Disco, the Rose Plots not only depict saccades that are transitions between different AOIs, but also transitions within the same AOI. Figure 5c shows a Rose Plot of the saccadic directions for the example depicted in Fig. 5a considering four cardinal directions (north, south, east, west). Note that saccades involving the filtered out AOI4 are not shown. It can be seen from the Rose Plot that 0 saccades point to the "north" (45°–135°), 2 saccades point to the "east" (315°–45°), 1 saccade points to the "south" (225°–315°), and 1 saccade points to the "west" (135°–225°).

5.2.2 Definition of areas of interest

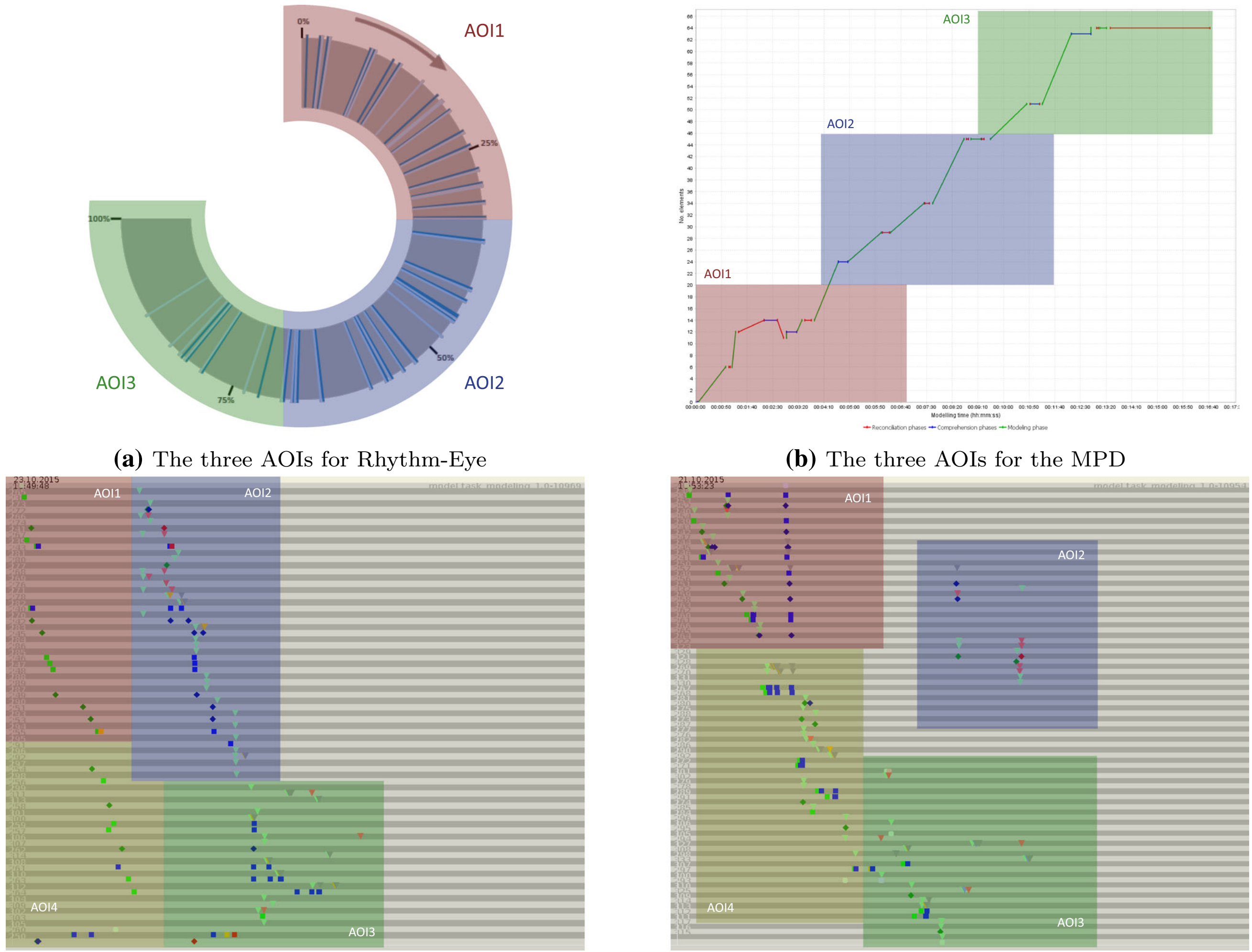
The process mining-based technique described previously relies on the definition of areas of interest. As suggested by [24], we defined the AOIs prior to the analysis considering our exploratory hypotheses, the composition of the stimulus, the quality of the data, and the method of analysis.

The AOIs of the Rhythm-Eye visualization are reported in Fig. 6a. Based on the initial findings reported in [44], we expect subjects to start reading the visualization in AOI1 and read the visualization in a clockwise direction, potentially with several iterations. Therefore, we defined three areas suitable to capture the time dimension of the capture data, i.e., the three AOIs correspond to the beginning, the middle, and the end of the temporal distribution of events/phases. The rationale behind using three AOIs is to keep the number of AOIs rather small, since we are interested in identifying reading patterns abstracting from details.

The AOIs for Modeling Phase Diagrams are depicted in Fig. 6b. Based on our experience in using MPDs, we expect subjects to start reading in the center (AOI2) followed by sequential reading of the visualization along the temporal dimension (i.e., AOI1 => AOI2 => AOI3). Therefore, we

[4] The complete source code for converting eye tracking data into event logs is available at https://github.com/DTU-SE/tsv2xes.

[5] See http://www.fluxicon.com/disco/.



(a) The three AOIs for Rhythm-Eye **(b)** The three AOIs for the MPD

(c) The four AOIs for the 2 PPM Charts we used for Q1 and Q4 respectively

Fig. 6 AOIs for the different visualizations used in this paper

identified three AOIs, which capture the temporal evolution of the modeling process.

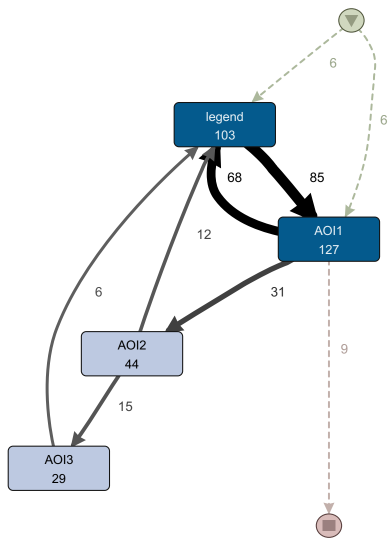
In the case of PPM Charts, the definition of AOIs is less obvious, since such visualization can be read horizontally in a temporal manner, vertically element by element, or in a combined manner. Moreover, the modeling elements are spread apart and difficult to group in homogeneous areas. To capture both the time dimension and the distribution of elements, we identified four AOIs, disposed in a sort of "flexible 2×2 grid" (cf. Fig. 6c). Moreover, the AOIs have been designed to roughly contain the same number of interactions.

In addition to the AOIs just mentioned, all models were decorated with two additional AOIs referring to clearly identifiable semantical components of the presented stimulus: one referring to the text with the analytical question and another one with the legend of the given visualization. We always filtered out fixations on the question text, since fixations on the text are independent of the actual visualization type. The legend has been included into all analyses where it was meaningful, as described in the result section.

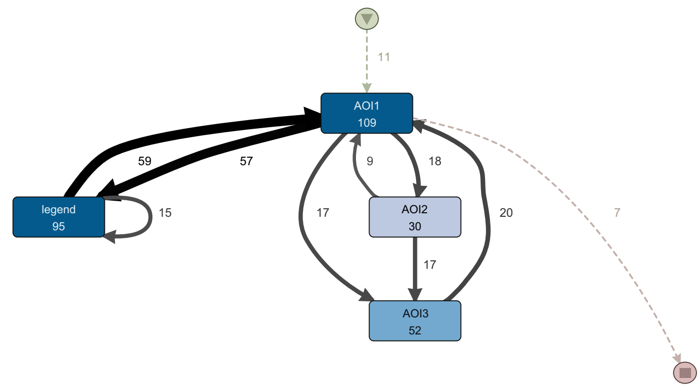
5.2.3 Observed reading patterns

Figures 7, 8, and 9 report the reading patterns identified by mining the event log and performing control-flow discovery. To extract these process models, we loaded the event logs (cf. Sect. 5.2.1) into the process mining tool Disco and performed the mining procedure.

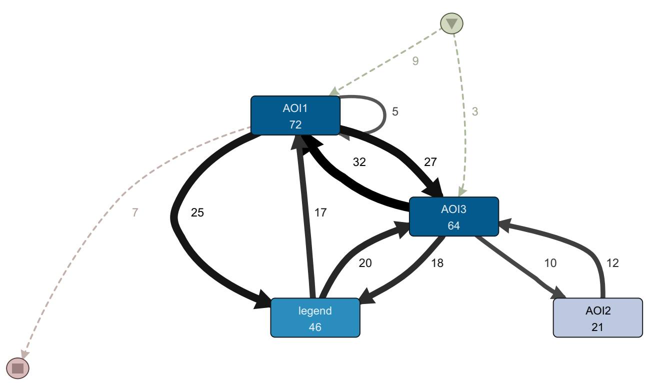
Rhythm-Eye The maps reported in Fig. 7 refer to the most frequent behavior observed of the participating subjects while answering `Q1-Q4` using the Rhythm-Eye visualization. From these charts, we can observe that area AOI1 was typically the first the subjects looked at. This is reasonable, since AOI1 coincides with the temporal beginning of the



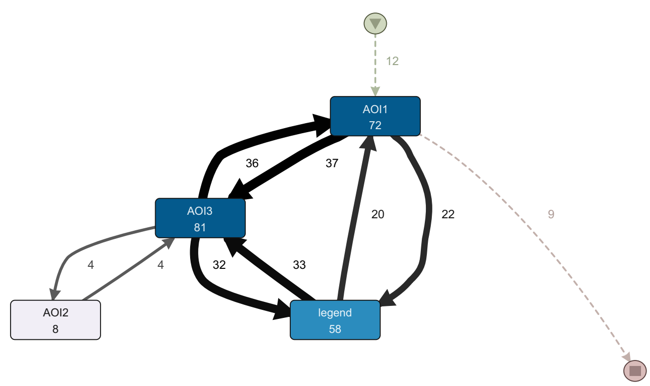
(a) Reading pattern of Rhythm-Eye answering Q1.



(b) Reading pattern of Rhythm-Eye answering Q2.



(c) Reading pattern of Rhythm-Eye answering Q3.

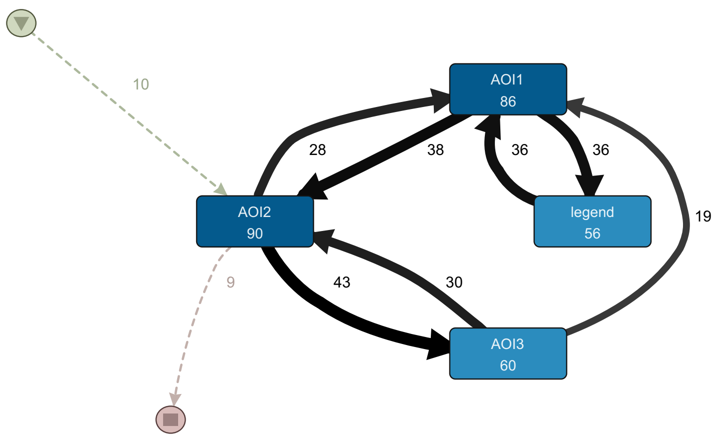


(d) Reading pattern of Rhythm-Eye answering Q4.

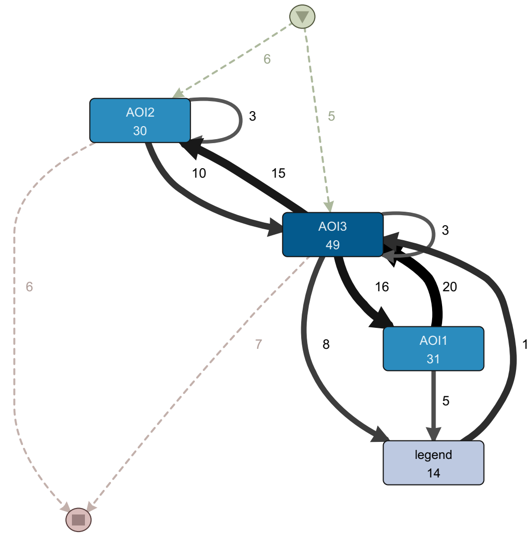
Fig. 7 Observed reading patterns for the Rhythm-Eye visualization

plotted data. In all cases, we see a notable amount of interactions between “AOI1” and “legend” which suggests that the training and learning take place at the very beginning of the analysis. As far as reading patterns are concerned, a distinction between relative timing (i.e., Q1 and Q2) and absolute timing (i.e., Q3 and Q4) can be observed. In case of absolute timing, the areas are typically visited in a clockwise direction, i.e., the flow goes from AOI1 to AOI2 and then AOI3. Transition frequencies associated with AOI1 are substantially higher than for AOI2 and AOI3 due to the interactions with the legend. In case of absolute questions, in turn, subjects typically started in AOI1 (which still represents the entry point) and then they quickly went to the area which is relevant for answering the question (cf. Fig. 7c, d). Both questions Q3 and Q4 asked about the ending of the process, i.e., the relevant area is AOI3. In both these cases, the patterns identified suggest that subjects who entered in AOI1 immediately moved to AOI3. Additionally, in both these cases, negligible number of fixations were on AOI2 (which does not provide information relevant for answering the questions). This suggests that the Rhythm-Eye visualization supports users in directing their attention to those parts of the visualization that is relevant for answering the respective questions.

Modeling phase diagram The maps reported in Fig. 8 refer to the reading patterns followed by subjects when answering Q2 and Q3 using Modeling Phase Diagrams. In general, most of the subjects started their reading in AOI2, which coincides with the center of the graph and not with its temporal beginning. When compared to the Rhythm-Eye visualization, the participants seem to be less guided in finding the starting point. Like with the Rhythm-Eye visualization, differences in the reading patterns can be identified depending on the timing dimension of the question. Question Q2 (cf. Fig. 8a) is a question with relative timing that requires to scan the entire



(a) Reading pattern of MPD answering Q2.



(b) Reading pattern of MPD answering Q3.

Fig. 8 Observed reading patterns for MPDs

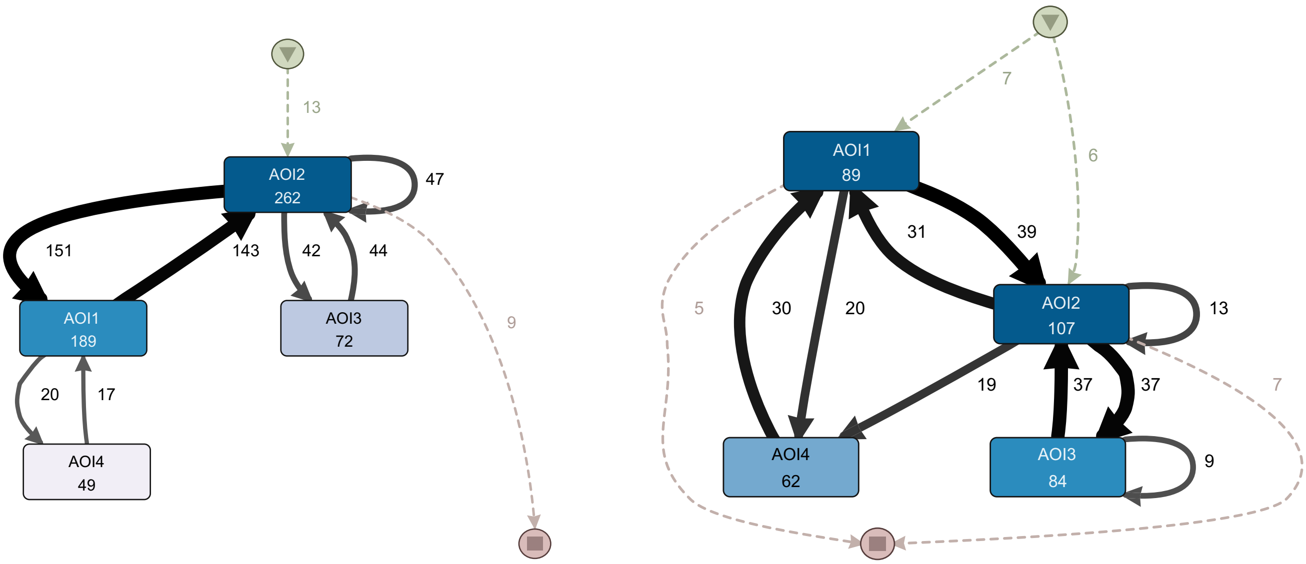
chart. Here, the users started from the center (i.e., AOI2) and then they went to AOI1, i.e., the temporal beginning of the graph. Here, subjects "learnt" how to read the chart by fixating the legend, and then they continued to AOI2 and then to AOI3, i.e., they followed the line from beginning to end. Some of the subjects repeated the whole procedure more than once (cf. the loop connecting AOI3 to AOI1). Question `Q3` (cf. Fig. 8b), in turn, is a question with absolute timing that asked to observe a certain phenomenon at the end of the modeling process. Again the majority of the subjects started their reading in AOI2; however, five subjects started their reading process immediately in AOI3, which is the area containing the information relevant for answering the question. Additionally, several interactions between AOI3 and the legend can be observed, suggesting that the learning took place in this area of interest.

PPM chart The last set of maps is depicted in Fig. 9. These maps refer to questions `Q1` and `Q4` answered using the PPM Charts. In both cases, users started reading the chart from the top (i.e., AOI1 and AOI2) which does not represent the logical beginning of the data, but just a subset of the elements represented. These results suggest that the PPM Chart visualization provides the least amount of guidance on where to start reading it. Moreover, it seems that PPM Charts to a lesser extent support users in focusing on the parts relevant for answering a particular question. For example, for `Q4` the relevant information is contained in AOI2 and AOI4, and still numerous visits of AOI1 and AOI3 can be observed. Irrespective of the timing dimension, we note horizontal movements on the top part of the chart (i.e., between AOI1 and AOI2) and vertical movements between top and bottom. Interestingly, horizontal movements are concentrated onto AOI1 and AOI2, while no horizontal movements between AOI3 and AOI4 can be observed. This might suggest that it takes subjects a while to understand how to best read a PPM Chart and that they transition from a horizontal reading pattern toward a more vertical reading pattern.

Since the area of interest definition for PPM Charts was not obvious, we decided to conduct further analysis to ensure that the results and their interpretations are not biased by the way the AOIs were defined. In particular, we wanted to ensure that the AOI definition was not too coarse-grained, potentially missing small vertical movements resulting from a line-by-line reading of the PPM Charts. We therefore decided to take a more fine-grained perspective and analyzed the saccadic direction (cf. Sect. 2.3) and visualized all saccades belonging to AOI1-4 using Rose Plots (cf. Fig. 10a, b). Note that these plots include both between and within AOI transitions and therefore provide more fine-grained insights into the predominant reading direction than the reading patterns obtained from Disco that only consider transitions between different AOIs. The Rose Plots are in line with the results obtained from Disco and confirm the predominantly horizontal reading pattern.

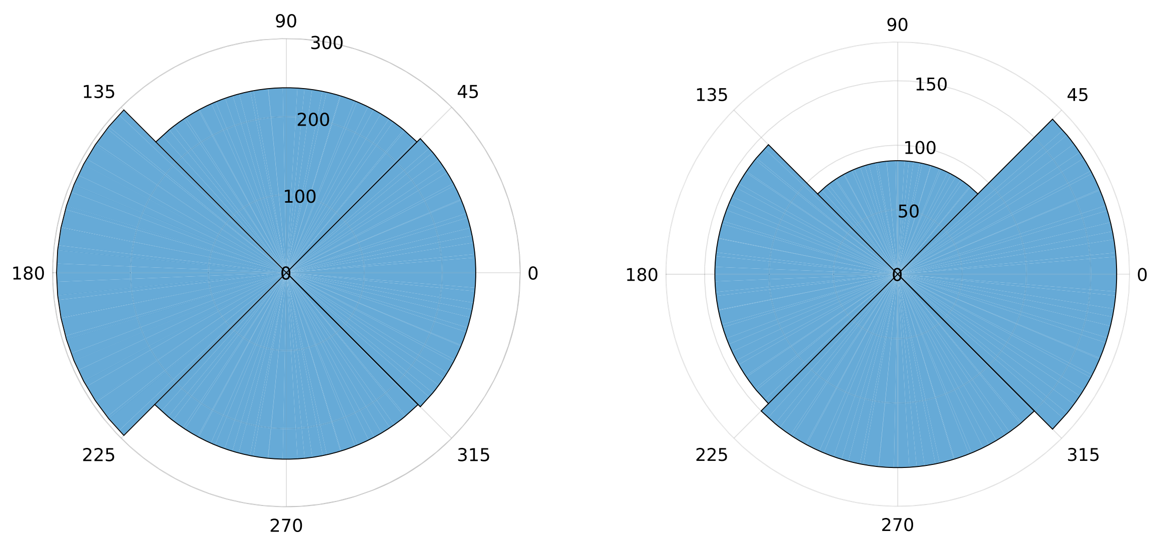
6 Discussion

In this section, authors propose some reflections concerning the presented technique and the results. Additionally, we point out possible limitations and describe the potential impact of the research.



(a) Reading pattern of PPM Chart answering Q1. **(b)** Reading pattern of PPM Chart answering Q4.

Fig. 9 Observed reading patterns for PPM charts



(a) Rose Plot of saccadic directions for Q1. **(b)** Rose Plot of saccadic directions for Q4.

Fig. 10 Rose Plots of saccadic directions for the PPM charts

6.1 Limitations

The findings reported in this paper are subject to several limitations: up to now we demonstrated the feasibility of the proposed approach considering one domain of Process Modeling Behavior Analysis. The application of our approach to other domains is needed to demonstrate its general nature. In order to achieve this, the use of purely categorical dimensions in the ViP framework will have to be revisited, and it should be considered how noncategorical dimensions, which allow to express continua or nonlinear hierarchies, can be integrated into the ViP framework approach.

There are also limitations concerning the evaluation of the different visualization types. The evaluation study is based on only 15 participants. This was mitigated by performing repeated measurements. It should be further pointed out that the core contribution of this paper, however, is not to demonstrate that a certain visualization type is better than another one (such result is only a collateral result). The core contribution of the paper, instead, is the methodological procedure suggesting how to come from analysis purposes to a selection of suitable visualization types. In the light of this, the limitations concerning sample size are less critical.

Another limitation concerns the limited number of questions presented to the participants. While such questions are certainly relevant and representative for Process Modeling Behavior Analysis, they only constitute a subset of relevant questions. Moreover, as with every eye tracking study, areas of interest need to be defined with care. Although we followed the guidelines proposed by [24] and defined areas of interest beforehand based on hypotheses, there is still an element of subjectivity involved.

6.2 Impact

The framework presented in this paper has important implications for practice. Specifically, we devised a formal and systematic procedure to come from analysis purposes to a selection of visualization types. The framework itself as well as the procedure to assess alternative visualization types is domain independent.

It is important to point out that the described framework is not limited to the analysis of data coming from the process of process modeling. Instead, several other projects and studies can benefit from it. In particular, in case of investigations concerning Process Modeling Behavior Analysis, it is possible to reuse the framework as it is, just by changing the research questions. This is due to the fact that data coming from this domain are intrinsically characterized by events and phases and, therefore, the dimensions as well as the categorization presented in Sects. 3.2.2 and 3.2.3 can be directly applied. However, this framework can be used also in other domains, where data is less similar to the PMBA case. In such scenarios, however, it is necessary to reconfigure it by following the five steps presented in Sect. 3.2 (e.g., identify the domain-relevant questions, corresponding dimensions addressed).

An example of a domain that can benefit from the framework (going beyond Process Modeling Behavior Analysis) is interfaces that are used by operators to provide timely and accurate answers according to different stimuli. An example of such system is the software to support control tower operators in an airport. In this case, the questions to answer are very specific and easy to qualify for domain experts (i.e., step 1 in Sect. 3.2; e.g., Is it possible to grant landing permission to a flight that requires it?). At the same time, the identification of the dimensions can be achieved easily (i.e., step 2 in Sect. 3.2; e.g., the priority of the request, if the request refers to air space or runways). With this information, it is possible to locate each question and each visualization available as specific configurations in the dimensions (i.e., step 3 and 4 in Sect. 3.2) and therefore proceed with the analysis (i.e., step 5 in Sect. 3.2). This example, only roughly sketched, gives a glimpse of the capabilities of the ViP framework in domains completely different from the one investigated in the paper.

7 Conclusion and future work

In this paper, we proposed a novel methodological approach for matching visualization types and analytical purposes. Using the case of Process Modeling Behavior Analysis as running example, we developed the ViP framework that gets instantiated for our domain with three fundamental dimensions: timing, actions, and instance handling. We then applied the framework to selected visualization types for Process Modeling Behavior Analysis and conducted experiments using eye tracking to compare them whenever more than one alternative per analytical purpose could be identified. The analysis of the collected eye tracking data included a statistical analysis using fixation-derived measures and the analysis of reading patterns using a novel process mining-based technique that is able to overcome known limitations of existing scanpath analysis tools (e.g., transition matrices). The process mining-based technique was complemented with saccade-derived measures where appropriate.

It is important to notice that the whole framework can be used in any other domain. Of course some applications might require additional configurations, but the flexibility of the procedure is described with one example, from a completely different domain, in Sect. 6.2. Future work will generalize our approach to further domains, e.g., the EAM Pattern Catalog [28] (cf. Sect. 2.1) could be used as large application example in the Enterprise Architecture Management domain.

Another avenue of future work will explore the potential of process mining for the analysis of eye tracking data in further detail. In this paper, we focused on a frequency-based analysis. Future work could explore the suitability of performance-based analyses as well.

Acknowledgements This work is partially funded by the Austrian Science Fund Project “The Modeling Mind: Behavior Patterns in Process Modeling” (P26609).

Appendix: Detailed descriptive statistics for timing dimension

See Tables 10, 11, 12, and 13.

Table 10 Descriptive statistics for event characteristics and relative timing

	N	Rhythm-Eye for events				PPM chart			
		Min	Max	Median	SD	Min	Max	Median	SD
Answer correctness	15	1.000	1.000	1.000	0.000	0.000	1.000	1.000	0.458
Answering time	15	11.010	87.630	34.940	20.680	19.450	105.190	49.400	27.057
Tot. fixation duration graph	15	2.250	27.910	8.330	8.178	4.080	55.360	22.360	16.967
Fixation count graph	15	10.000	108.000	34.000	29.613	20.000	266.000	82.000	73.647

Table 11 Descriptive statistics for phase characteristics and relative timing

	N	Rhythm-Eye for phases				Modeling phase diagram			
		Min	Max	Median	SD	Min	Max	Median	SD
Answer correctness	15	1.000	1.000	1.000	0.000	0.000	1.000	0.000	0.516
Answering time	15	9.830	81.640	18.170	17.893	3.060	100.110	21.270	26.534
Tot. fixation duration graph	15	1.910	26.490	6.490	6.204	2.430	60.410	15.190	14.643
Fixation count graph	15	8.000	117.000	24.000	28.102	9.000	259.000	66.000	58.807

Table 12 Descriptive statistics for event characteristics and absolute timing

	N	Rhythm-Eye for events				PPM chart			
		Min	Max	Median	SD	Min	Max	Median	SD
Answer correctness	15	1.000	1.000	1.000	0.000	0.000	1.000	1.000	0.352
Total duration	15	6.580	32.510	15.230	6.753	7.880	81.230	18.990	18.368
Tot. fixation duration graph	15	1.160	10.700	4.530	2.895	3.850	43.330	8.690	10.600
Fixation count graph	15	5.000	50.000	21.000	12.927	22.000	208.000	40.000	50.336

Table 13 Descriptive statistics for phase characteristics and absolute timing

	N	Rhythm-Eye for phases				Modeling phase diagram			
		Min	Max	Median	SD	Min	Max	Median	SD
Answer correctness	15	1.000	1.000	1.000	0.000	0.000	1.000	1.000	0.258
Total duration	15	7.500	27.970	13.610	6.143	5.450	33.400	8.700	7.343
Tot. fixation duration graph	15	2.020	12.630	4.710	3.227	2.950	21.240	5.960	5.099
Fixation count graph	15	11.000	52.000	23.000	12.755	13.000	88.000	25.000	23.405

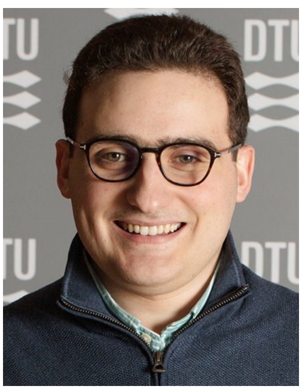
References

1. van der Aalst, W.M.: Process Mining, 2nd edn. Springer, Berlin (2016). https://doi.org/10.1007/978-3-662-49851-4
2. Basili, V.R., Selby, R.W., Hutchens, D.H.: Experimentation in software engineering. IEEE Trans. Softw. Eng. **12**(7), 733–743 (1986). http://dl.acm.org/citation.cfm?id=9775.9777
3. Bennett, K.B., Flach, J.M.: Display and Interface Design—Subtle Science. Exact Art. CRC Press, Boca Raton (2011)
4. Berson, A., Smith, S.J.: Data Warehousing, Data Mining, & OLAP. McGraw-Hill, New York (2004)
5. Bertin, J.: Semiology of Graphics: Diagrams, Networks. Maps. University of Wisconsin Press, Madison (1983)
6. Borner, K., Polley, D.E.: Visual Insights: A Practical Guide to Making Sense of Data. MIT Press, Cambridge (2014)
7. Buckl, S., Ernst, A., Matthes, F., Schweda, C.: Enterprise architecture management patterns for enterprise architecture visioning. In: Proceedings of the 14th Annual European Conference on Pattern Languages of Programming (EuroPLoP 2009). Irsee, Germany (2009)
8. Burattin, A., Kaiser, M., Neurauter, M., Weber, B.: Eye tracking meets the process of process modeling: a visual analytic approach. In: TAProViz (2016)
9. Cairo, A.: The Functional Art. Voices that Matter. Pearson Education (2010)
10. Chen, C.: Information Visualization. Springer, London (2010)
11. Chen, C., Härdle, W.K., Unwin, A. (eds.): Handbook of Data Visualization. Springer, Berlin (2008)
12. Claes, J., Vanderfeesten, I., Pinggera, J., Reijers, H.A., Weber, B., Poels, G.: A visual analysis of the process of process modeling. IseB **13**(1), 147–190 (2015). https://doi.org/10.1007/s10257-014-0245-4
13. Claes, J., Vanderfeesten, I.T.P., Gailly, F., Grefen, P., Poels, G.: The structured process modeling method (SPMM) what is the best way

for me to construct a process model? Decis. Support Syst. **100**, 57–76 (2017). https://doi.org/10.1016/j.dss.2017.02.004
14. Codd E.F., C.S.B.S.C.T.: Providing olap (on-line analytical processing) to user analysts: an it mandate (1993)
15. Cooper, A.: About Face: The Essentials of Interaction Design, 4th edn. Wiley, Hoboken (2014)
16. Few, S.: Information Dashboard Design: The Effective Visual Communication of Data. O'Reilly, Sebastopol (2006)
17. Frank, U., Heise, D., Kattenstroth, H.: Use of a domain specific modeling language for realizing versatile dashboards. In: J.P. Tolvanen, M. Rossi, J. Gray, J. Sprinkle (eds.) Proceedings of the 9th OOPSLA Workshop on Domain-Specific Modeling (DSM), Helsinki Business School, Helsinki, 2009 (2009)
18. Gärdenfors, P.: Conceptual Spaces. MIT Press, Cambridge (2000)
19. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. **24**(6), 631–645 (1999). https://doi.org/10.1016/S0169-8141(98)00068-7. http://www.sciencedirect.com/science/article/pii/S0169814198000687
20. Gulden, J.: Visually comparing process dynamics with rhythm-eye views. In: Dumas, M., Fantinato, M. (eds.) Business Process Management Workshops: BPM 2016 International Workshops, Rio de Janeiro, Brazil, September 19, 2016. Revised Papers, Springer, Berlin (2016)
21. Gulden, J., van der Linden, D., Aysolmaz, B.: Requirements for research on visualizations in information systems engineering. In: ENASE Conference (2016)
22. Gulden, J., Reijers, H.A.: Toward advanced visualization techniques for conceptual modeling. In: CAiSE Forum. CEUR (2015)
23. Hichert, R., Faisst, J., et al.: International business communication standards (ibcs®) v. 1.1 (2017). http://www.ibcs-a.org/
24. Holmqvist, K.: Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press, Oxford (2011)
25. IEEE Task Force on Process Mining: Process Mining Manifesto. In: F. Daniel, K. Barkaoui, S. Dustdar (eds.) Business Process Management Workshops, pp. 169–194. Springer-Verlag (2011)
26. Johnson, M., Lakoff, G.: Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. Basic Books, New York (1999)
27. Juristo, N., Moreno, A.M.: Basics of Software Engineering Experimentation (2001)
28. Khosroshahi, P.A., Hauder, M., Schneider, A.W., Matthes, P.D.F.: Enterprise architecture management pattern catalog version 2.0. Tech. rep., Technical University Munich, Munich (2015)
29. Kirk, A.: Data Visualization: a successful design process. Packt Publishing (2012)
30. Kleiner, N.: Can business process changes be cheaper implemented with workflow management systems? Innovations Through Information Technology **1 and 2**, 529–532 (2004)
31. Kurzhals, K., Fisher, B.D., Burch, M., Weiskopf, D.: Eye tracking evaluation of visual analytics. Information Visualization **15**(4), 340–358 (2016). https://doi.org/10.1177/1473871615609787
32. Pinggera, J.: The Process of Process Modeling. Ph.D. thesis, University of Innsbruck, Department of Computer Science (2014)
33. Pinggera, J., Soffer, P., Fahland, D., Weidlich, M., Zugal, S., Weber, B., Reijers, H., Mendling, J.: Styles in business process modeling: an exploration and a model. Softw. Syst. Modeling **14**(3), 1055–1080 (2015). https://doi.org/10.1007/s10270-013-0349-1
34. Pinggera, J., Soffer, P., Fahland, D., Weidlich, M., Zugal, S., Weber, B., Reijers, H.A., Mendling, J.: Styles in business process modeling: an exploration and a model. Softw. Syst. Modeling **14**(3), 1055–1080 (2015). https://doi.org/10.1007/s10270-013-0349-1
35. Pinggera, J., Zugal, S., Weidlich, M., Fahland, D., Weber, B., Mendling, J., Reijers, H.: Tracing the process of process modeling with modeling phase diagrams. In: Proc. ER-BPM '11, pp. 370–382 (2012)
36. Poole, A., Ball, L.J.: Eye tracking in human-computer interaction and usability research: current status and future. In: Prospects", Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc (2005)
37. Poole, A., Ball, L.J., Phillips, P.: In Search of Salience: A Response-time and Eye-movement Analysis of Bookmark Recognition, pp. 363–378. Springer London, London (2005). https://doi.org/10.1007/1-84628-062-1_23
38. Sharp, H.: Interaction Design. Wiley, Hoboken (2011)
39. Shmueli, G., Bruce, P., Yahav, I., Patel, N., Lichtendahl, K.: Data Mining for Business Analytics: Concepts, Techniques, and Applications in R. Wiley (2017). https://books.google.de/books?id=ETwuDwAAQBAJ
40. Sjoeberg, D.I.K., Hannay, J.E., Hansen, O., Kampenes, V.B., Karahasanovic, A., Liborg, N., Rekdal, A.C.: A survey of controlled experiments in software engineering. IEEE Trans. Softw. Eng. **31**(9), 733–753 (2005). https://doi.org/10.1109/TSE.2005.97
41. Spence, R.: Information Visualization. Prentice Hall, Upper Saddle River (2007)
42. Surma, J.: Business Intelligence: Making Decisions Through Data Analytics. Business Expert Press, New York (2011)
43. Tufte, E.R.: The Visual Display of Quantitative Information. Graphics Press, Cheshire (1983)
44. Weber, B., Gulden, J., Burattin, A.: Designing visual decision making support with the help of eye-tracking. In: S. Nurcan, J. Gulden (eds.) BPMDS 2017 Radar Proceedings. CEUR (2017)
45. Zelkowitz, M.V., Wallace, D.R.: Experimental models for validating technology. Computer **31**(5), 23–31 (1998). https://doi.org/10.1109/2.675630

Jens Gulden fills the position of an Assistant Professor at the chair for Information Systems and Enterprise Modeling at University of Duisburg-Essen. He previously was affiliated, among others, with Free University Berlin and Technical University Berlin and was a visiting scientist at Middlesex University London and University of Toronto. Jens's research interests focus on exploring non-linguistic modes of expressing meaning as part of the interactions in socio-technical information systems. In the course of these activities, he examines and develops methods for designing information visualizations and visual modeling languages. His teaching activities cover the design and use of domain-specific languages and visual notations, as well as the methodical use of model-driven specification and development techniques for software system design and implementation. Jens Gulden also carries a university degree in philosophy, which he obtained in the areas of embodied cognition and philosophy of pictures.

Andrea Burattin is Associate Professor at the Technical University of Denmark, Denmark. Previously, he worked as Assistant Professor at the same university and as postdoctoral researcher at the University Innsbruck (Austria) and at the University of Padua (Italy). In 2013, he obtained his Ph.D. degree from a joint Ph.D. School between the University of Bologna and Padua (Italy). The IEEE Task Force on Process Mining awarded to his Ph.D. thesis the Best Process Mining Dissertation Award for theses defended in 2012–2013. A reworked version of his thesis has then been published as Springer Monograph in the LNBIP series. He served as organizer of BPI workshop since 2015 and special sessions on process mining at IEEE CIDM since 2013. He is also in the program committee of several conferences such as BPM, CAiSE, and IJCAI. His research interests include process mining techniques and, in particular, online processing of event data.



Amine A. Andaloussi is a Ph.D. candidate at the Technical University of Denmark. He received his M.Sc. degree from the same university in 2018. His main research is about the understandability of business process models. With an academic background in computer science and strong research interest in cognitive psychology, Amine investigates the way end users engage with business process models in order to improve the understandability of process modeling notations and enhance the existing modeling tool support.

Barbara Weber is Professor for Software Systems Programming and Development at the University of St. Gallen, Switzerland. She is Chair for Software Systems Programming and Development and Director of the Institute of Computer Science. In addition, she holds a part-time full professor position at the Department of Applied Mathematics and Computer Science with the Technical University of Denmark. Barbara's research interests include process model understandability, process of process modeling, process flexibility, and user support in flexible process-aware systems as well as neuro-adaptive information systems. Barbara has published more than 150 refereed papers, for example, in Nature Scientific Reports, Information and Software Technology, Information Systems, Data and Knowledge Engineering, Software and System Modeling, and Journal of Management Information systems and is co-author of the book "Enabling Flexibility in Process-aware Information Systems: Challenges, Methods, Technologies" by Springer.