Leopold-Franzens-University Innsbruck

Institute of Computer Science
Research Group Quality Engineering

MASTER THESIS

Impact of Layout Properties on the Understandability of Process Models

Katja Sonderegger

Supervisor: Assoc.-Prof. Dr. Barbara Weber, Dr. Andrea Burattin

Innsbruck, June 18, 2016

Contents

Abstract

Process modelling is widely used in business and science and is often used by companies for representing important business processes. The process is represented in a simplified, graphical way which can be used for documentation, communication and testing purposes.

The appearance of process models is a very important area of research, as it has an impact on the understandability of the process model. A lot of research is done in the area of process model understandability concerning influencing factors, the modelling language and visual features of a model. However, the layout of a process model and its influence on the understandability of a process model still needs further research. In particular, it is unclear whether certain layout properties do have a significant impact on the understandability of a process model.

This thesis takes a look at two layout properties: the change of direction and the visibility of blocks. It investigates whether these properties have an impact on the understandability of process models.
To get empirical data on the actual impact of these properties on the understandability of the process models, a controlled experiment was conducted. In this experiment 68 students from the University of Ulm participated.

While the change of direction does not have a significant effect on the response variables accuracy and duration in general, it has a significant effect on the response variable duration for global question types. The visibility of blocks feature has a significant impact on the duration in general and additionally has a significant impact on the accuracy for global question types. These results should be able to help improving automated layout algorithms and guidelines for the process model creation, as well as the training of future process modellers.

Acknowledgement

At this point I would like to thank all the people that supported me in completing this master thesis. Without their support, feedback and encouragement this thesis would not have been possible.

First of all I want to thank **Assoc.-Prof. Dr. Barbara Weber** for guiding me the whole time and for the feedback and motivation along the way. I highly appreciated working with you.

My thanks also go to **Prof. Pnina Soffer** and **Dr. Andrea Burattin** for all the discussions and constructive feedback during the development of the experiment.

Further I want to thank **Dr. Jakob Pinggera** who helped me a lot during the preparation and the finalization of the experiment and **Mag. Manuel Neurauter** who took the time to have a look at the questions and models of the experiment and guided me through the wide area of data analysis and statistics.

I also want to thank **Prof. Dr. Manfred Reichert** and **Johannes Schobel** who allowed me to use their students for the experiment.

Additionally I want to thank my boyfriend and my friends for proof reading this thesis and motivating me throughout my studies.

Last but not least I want to thank **my parents**, who are always there for me and supported me while studying computer science.

1. Introduction

1. Motivation

In the last decades business process modelling grew to be a very promising and essential area of research [7] [46]. Therefore, also the usage of process models in companies has increased [1]. Process models are used for representing important processes, as process modelling helps to evolve a shared view of the process. Process modelling represents the process in a graphical, simplified way and can be used for documentation, as well as testing and communication purposes [8].

Modelling notations provide elements to represent a business process in a graphical way, as activities, events and the overall process flow. Every notation has a focus on a certain area of usage. The used modelling notation in this thesis is called BPMN (Business Process Model Notation) and is one of the commonly used modelling notations for business processes [38].

There has been a lot of research on the quality of process models and the general understandability of process models. Workflow patterns [54] provide the modeller with patterns that help to model frequently occuring workflow requirements, for example parallelism. This helps to reduce the complexity of the process model creation as well as the reading of the process model, as cognitive load theory shows that mental effort for tasks is reduced by *chunking* the data [65].

Many modelling tools provide the user with automatic layout functions to align their activities and straighten edges, so that the model has a cleaner layout [25] [10]. The impact of layout on the understandability of a process model has already been researched [11] [50], but there are still gaps that have to be further researched concerning the impact of the layout on the understandability of a process model. Concerning the understandability of process models, there are only a few validated metrics available. Especially the influence of block visibility and the change of direction have not yet been considered.

Therefore, the goal of this thesis is to address this gap by looking at two layout properties (i.e., the visibility of blocks and the change of direction) and to investigate whether these properties have an impact on the understandability of process models. These layout properties were chosen, due to the investigations of earlier research [12] [2], that found that those properties may have a significant effect on the understandability of a process model. To measure the understandability impact of these layout properties, a controlled experiment is conducted.

This controlled experiment should give more empirical insight into the influence of the mentioned layout properties on the understandability of a process model. The results of this experiment have a potential impact for the improvement of automatic layout algorithms and set guidelines for process modelling concerning the layout of a process model.

2. Related Work

Closely related to this work is work on process model understandability in general and in particular secondary notation like layout.

A large amount of the related research focuses on the factors that influence understandability of process models and other model representations. While some of these papers focus on the modelling language and its semantic clarity [35] or the graphical elements of the language [37] [44], others focus on visual features of a model [47] [53]. Mendling et al. have identified certain model quality dimensions [32] and factors [33], which have an influence on the understandability of a process model. Features that influence understandability have been determined [29] [40] [46] [64], and metrics that try to measure complexity and other features of a model have been developed [15] [28] [41] [56] [57].

As psychology shows, appearance has an effect on understandability [26] [37] and therefore the influence of layout on the understandability of process models has to be further researched [11] [50]. Some work has already been done in this area of research. Figl and Strembeck [12] conducted an experiment to determine, whether the flow direction of a process model has an effect on the understandability of the process model. In this experiment, only consistent flow directions and no direction changes within the process model were taken into account. Bernstein et al. [2] have identified certain layout properties that could have an influence on the understandability of process models. The significance of some of these layout properties needs to be the focus of further research and is not yet clear. Weidlich et al. [59] have identified a difference in the accuracy of changes on process models depending on whether the changes are sequential or circumstantial. These results are the reason for the differentiation of the question types in this thesis. There are also a lot of algorithms available, that focus on the automatic layout of process models [16] [17] [19] [21] [42].

Metrics have been proposed [3] to measure the understandability of a process model concerning certain layout properties, but they have not been empirically evaluated yet. This work picks up this gap and to determine the influence of the layout on the understandability of a process model, this thesis conducts a controlled experiment with two chosen layout properties.

3. Research Objective

In particular, the thesis will look at two layout properties and their influence on the understandability of a process model.

Change of direction The first layout property examined by this thesis is the influence of the *change of direction* on the understandability of a process model. This property has been identified as an important property in [2], but the impact on the understandability has not been further evaluated. Therefore this thesis tries to get an empirical insight on the impact on understandability of this property.

Block visibility The second layout property investigated in this thesis is the *block visibility* of a process model and whether it has an influence on the understandability of a process model. Structuredness has been identified as an important property by Mendling and Reijers [31]. The importance of visual structuredness and the impact of a structured model without the layout showing the blocks clearly, has not yet been investigated. Therefore this thesis tries to get an empirical insight on the impact of this layout property on the understandability of a process model.

The following research questions can be derived from these objectives:

Research Question 1 *Does the change of the direction have a significant influence on the understandability of a process model?*

Research Question 2 *Does the block visibility have a significant influence on the understandability of a process model?*

4. Research Method

The applied research method is shown in Figure 1.1 [63] and is further explained in the following section.

First of all, the problem was formulated and the first literature research was conducted to get an overview of the already existing related work in this area of research. After this step the research questions were developed and the focus of the thesis was set to two layout properties in particular: the *change of direction* and the *block visibility.*

As this thesis goal is to get an empirical insight on the impact of the two mentioned layout properties on the understandability of process models, the next step was the

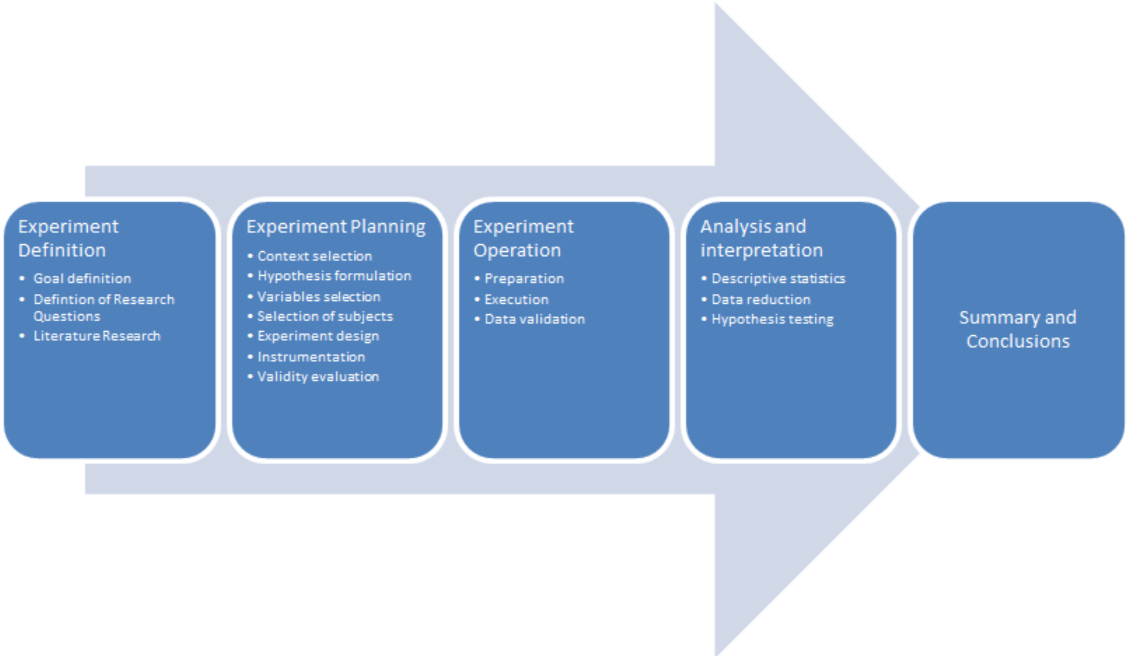

Figure 1.1.: Overview of the applied research method, adapted from [63].

planning of the experiment. In this step, the experimental design was developed. In the experimental design the models, the questions and the hypotheses were developed. The kind of experiment applied in this thesis is a so-called *controlled experiment* [20], as the main goal of this thesis is to get an empirical insight on the influence of the chosen layout properties on the understandability of a process model.

After the planning phase, the operation phase started with the preparation of the experiment, which was set up in the *Cheetah Experimental Platform (CEP)* [39]. Then the experiment was conducted and the gathered data was validated.

Following, the data was described with the help of descriptive statistics, before the data was further analysed to finally accept or reject the hypotheses.

5. Structure

This chapter gave an introduction on the topic of this thesis, an overview of the related work and the goal of this thesis as well as an explanation of the used research method.

Chapter 2 describes the main concepts that are used in this thesis and links to related work in this area of research, while Chapter 3 describes the setup and the execution of the experiment in detail.

The results of the experiment are discussed in Chapter 4.

Finally, Chapter 5 concludes the thesis with a discussion about future work.

2. Concepts

In this chapter an explanation of business processes is given as well as an overview of the related work of this thesis. Afterwards the main concepts and properties that are used in this thesis are described in detail.

1. Business Process

A *business process* is a process that has the purpose of reaching a certain goal or to fulfill a business function [45] [60]. Formally, a business process is described in the WMC specification [52] as:

> *"A set of one or more linked procedures or activities which collectively realise a business objective or policy goal, normally within the context of an organisational structure defining functional roles and relationships."*

A business process has a control flow, which defines a partial order relationship indicating in which temporal order the activities of a process will be executed [4].

2. Process Model

A *process model* is a visual representation of a process in a certain domain, e.g. a business process [36]. This representation can be modelled with the help of different notations. Some of these notations are UML Activity Diagram [9], BPMN (Business Process Modelling Notation) [38], Petri nets [55] and EPC (Event Driven Process Chain) [27]. This thesis will focus on the BPMN notation, which will be explained in the following.

BPMN The *Business Process Model and Notation (BPMN)*[1] is a commonly used notation for the modelling of business processes. BPMN is developed by *The Object Management Group (OMG).* The most current version of this standard is BPMN 2.0, which is available since 2011 and supports a lot of different use cases.

[1] `www.bpmn.org`, accessed on: 09.01.2016

The specification of BPMN 2.0 [38] states: *"The primary goal of BPMN is to provide a notation that is readily understandable by all business users, from the business analysts that create the initial drafts of the processes, to the technical developers responsible for implementing the technology that will perform those processes, and finally, to the business people who will manage and monitor those processes. Thus, BPMN creates a standardized bridge for the gap between the business process design and process implementation."*

A short overview of the main BPMN elements [14] can be seen in Figure 2.1, and will be explained next.

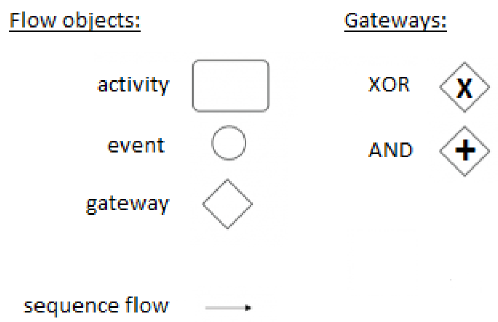

Figure 2.1.: The core elements of BPMN.

The *flow objects* of BPMN 2.0 are activities, gateways and events, which can be seen in Figure 2.1. While *activities* are things that have to be done, the *gateways* are the conditions when these activities have to be executed and additionally there are *events* that can happen. These flow objects are connected through arrows, which represent the overall *sequence flow*, which is also called *control flow*.

A process model consists of activities and gateways that are connected through sequence flows (arrows). The overall process shows the sequences from a start event to the end event of a process. An example of a process can be seen in Figure 2.2 - here a small process is represented, which explains the process of making a cup of coffee.

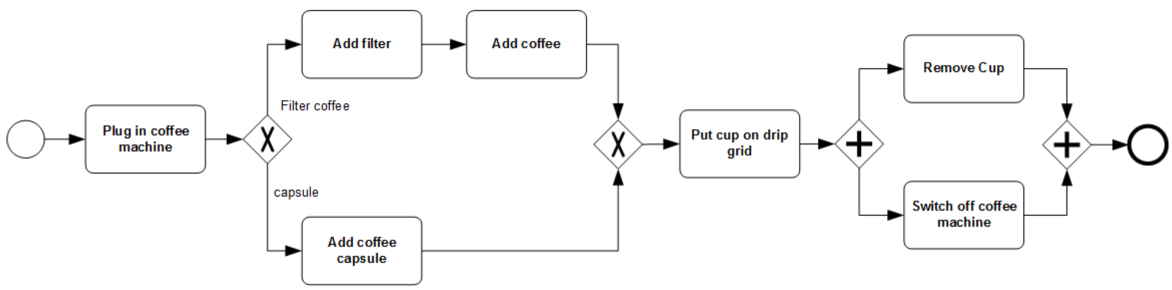

Figure 2.2.: A small process for coffee-making modelled in BPMN.

In Figure 2.2, a process is shown that starts with a start event, then there is an activity *plug in the coffee machine*. After that - depending on the type of coffee machine - the activities for the filter machine, or the activity for the capsule machine are chosen. This is modelled with the use of a XOR-gateway. The activities *remove cup* and *switch off coffee machine* can be executed in parallel and are therefore connected with an AND-gateway. After that an end event shows that the process is finished.

Besides the already mentioned core elements of BPMN, there are certain patterns that occur very often in a BPMN model. The most often used control flow patterns are shown in Figure 2.3[2].

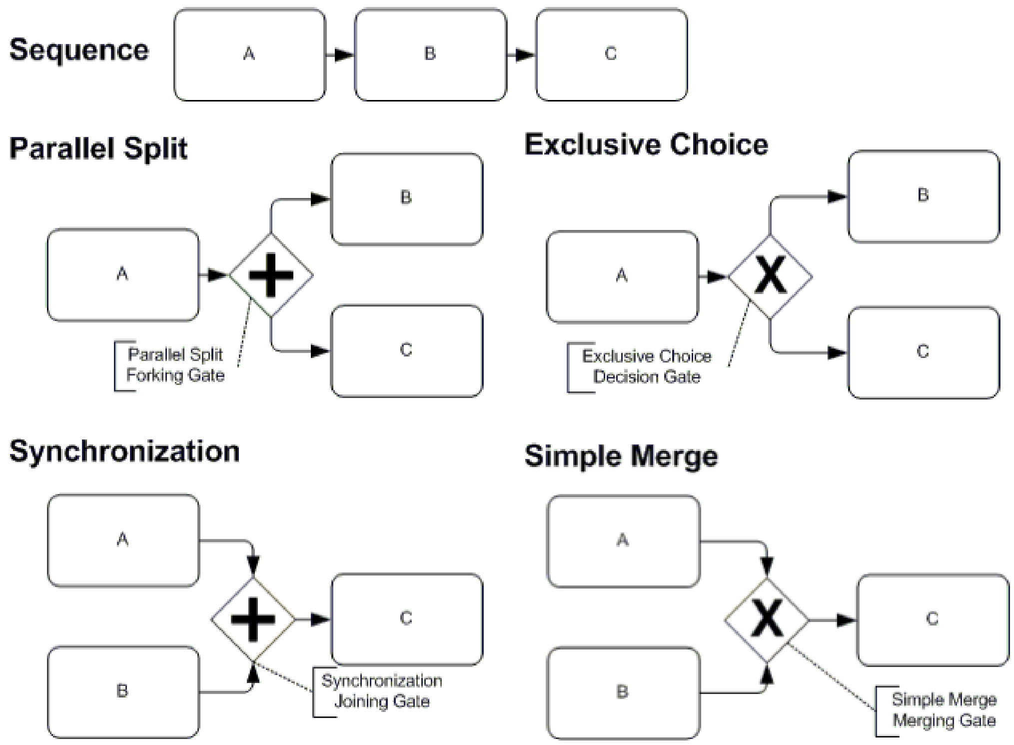

Figure 2.3.: The most often used control flow patterns. Source: http://www.workflow-patterns.com/vendors/documentation/BPMN-pat.pdf

Subsequently the control flow patterns of Figure 2.3 are explained:

- The *Sequence Pattern* is an ordered series of activities, where an activity can start as soon as the predecessor activity has completed [61].

[2]`http://www.workflowpatterns.com/vendors/documentation/BPMN-pat.pdf`, accessed on: 11.01.2015

- The *Parallel Split* is also called *AND-Split* and allows two or more activities to be performed concurrently [61].
- The *Synchronization* or *AND-Join* gathers different paths created by the Parallel Split and waits until every path has completed, before the process can continue [61].
- The *Exclusive Choice* or *OR-Split* splits the process into two or more different paths. Only one of those paths can be chosen, depending on the condition the path has [61].
- The *Simple Merge Pattern* or *OR-Join* gathers different paths and joins them into a single path [61].

3. Process Model Quality

Process Model Quality is an essential and critical aspect of process modelling and can be divided into three types of quality: syntactic, semantic and pragmatic quality [22] [23] [18].

- *Syntactic* quality addresses the correspondence of the graphical representation of the process and the notation language used to model the process. Most of the syntactic errors can be detected automatically by verification algorithms [18].
- *Semantic* quality addresses the completeness and validity of a process model and is up to now only detectable by humans.
- The third quality type - the *pragmatic* quality addresses the correspondence of the graphical representation of the process and the interpretation of it by a human. This quality type can be described as the understandability of a process model and is the main focus of this thesis, which is explained in the following section.

3.1. Process Model Understandability

The *pragmatic* quality or understandability of a process model is the main focus of this thesis. While the understandability of a process model is influenced by many factors, this thesis looks at the impact of the layout of a process model.

Layout

Vered Bernstein and Pnina Soffer [2] describe different measurable layout properties that seem to be important for the understandability of a process model. In [2] four main properties are identified:

- Edges (length, style, crossing, text on edges)
- the model's general layout (number of ending points, orthogonal segments)
- the model's direction (horizontal direction, vertical direction and the change in the model's direction)
- and the alignment of elements

For this thesis an experiment is conducted, which looks at the following layout properties in particular: the change of direction and the visibility of blocks, which is one aspect of the alignment property.

Change of Direction

The first layout property that is used in this experiment is the change of direction and the influences of the change of direction on the overall understandability of a process model. The model's direction has an impact on the understandability of process models [12] - therefore it would be interesting to find out, if a change in the model's direction also has an impact on the understandability of a process model. There are different operationalizations available and in this thesis the BP metric explained in the following has been chosen to calculate if the overall direction has changed.

Behavioural Profiles (BP) The *Behavioural Profiles (BP)* [3] metric analyses the position of a node and checks to which extent the layout reflects the temporal logical ordering of activities. For each relation this metric looks at the position of the source node and the position of the target node and checks if the source node is *graphically before* the target node.

A *behavioural relation* is a relation between a source node and a target node. In the following the three different behavioural relations, that are needed for the metric calculation, are explained on the example model in Figure 2.5.

A *strict relation* is a relation between two activities, where the order of the activities never changes (A always occurs before B in Figure 2.5) - this typically stands for a sequence.

An *exclusive relation* is a relation between two or more activities, where the order changes (C cannot appear before D, D cannot appear before C in Figure 2.5) - this stands for an OR-Split and OR-Join.

An *interleaving relation* is a relation between two or more activities, where the order might change (E might appear before F, F might appear before E in Figure 2.5) - this stands for an AND-Split and AND-Join.

```
Input: G = (V, E, L_V, L_E):
t_strict ← 0
correct_East ← 0
correct_South ← 0
BP ← BehavioralProfiles(G) /* Compute all behavioral relations */
foreach bp ∈ BP do
    if #_relation(bp) = → then /* Only strict order relations */
        /* Extract the coordinates of the central points of the
           source and target nodes */
        (s_x, s_y) ← L_V(#_source(bp))
        (t_x, t_y) ← L_V(#_target(bp))
        if s_x < t_x then /* Check for the East direction */
            correct_East ← correct_East + 1
        end
        if s_y < t_y then /* Check for the South direction */
            correct_South ← correct_South + 1
        end
        t_strict ← t_strict + 1
    end
end
return max{correct_East, correct_South} / t_strict /* Final consistency
    score as the dominant direction, divided by the total
    number of strict relations */
```

Figure 2.4.: Pseudo algorithm for the BP metric [3].

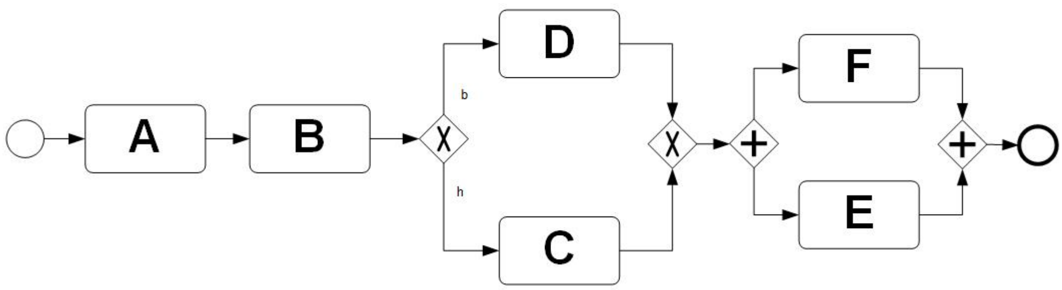

Figure 2.5.: Example model for the BP metric.

In detail the BP metric algorithm shown in Figure 2.4 works as follows: A process graph G = (V, E, L_V, L_E) is given, and a relationship b with the relation type (r), which determines, if it is a sequence ($\rightarrow$), OR (+) or AND (||) and (s) as the source node and (t) as the target node. Additionally there exists a procedure BehavioralProfiles(G) which extracts all behavioral relations out of the graph G. In short the algorithm initializes its variables and extracts the behavioral profile relations out of the graph. Then it looks at each behavioral profile relation and checks, whether it is a strict relation. If it is a strict relation, it computes the temporal location of both nodes and whether the first node is graphically before the second node. The final score is the division of the amount of graphically before relations and the total number of strict relations.

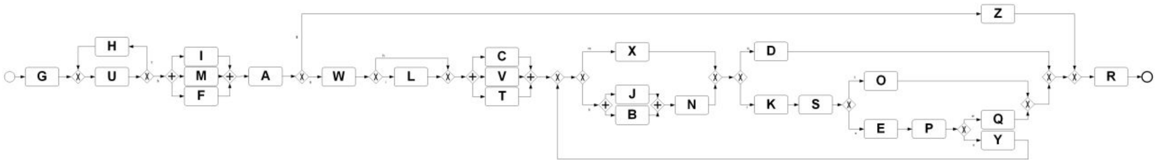

Figure 2.6.: Model 1 in the version for group A. This model has a low change of direction.

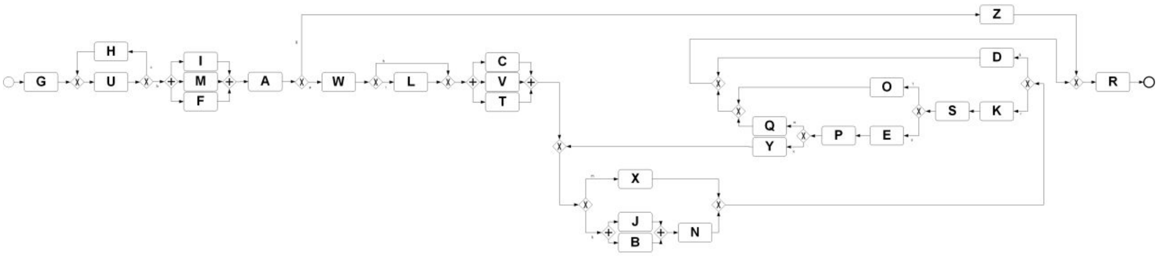

Figure 2.7.: Model 1 in the version for group B. This model has a higher change of direction, than Model 1 of group A.

Figure 2.6 and Figure 2.7 show the same model, where in Figure 2.7 the change of direction is higher than in Figure 2.6.

The value for the flow consistency of Figure 2.6 is with 0.932 very high. Note that the value of Figure 2.6 is not 1, due to the loop in the model which also counts as a change of direction. The BP metric for Figure 2.7 is much lower with a flow consistency of only 0.795 - which is due to the block containing activities D, K, S, O, E, P, Q, Y that has a reverse flow direction.

Visibility of Blocks

The other property that was used in this experiment is the block visibility. In this property the visual structuredness of the elements of the process model is of great importance. As Mendling and Reijers [32] found out, structuredness does have an effect on the understandability of a process model. Therefore the second property that would be interesting to investigate, is the impact of the visibility of blocks on the understandability of a process model.

In Figure 2.8 and Figure 2.9 the position of the activities is the same in both models. While in Figure 2.9 all blocks are visible and all gateways are aligned, in Figure 2.8 the gateways are not aligned and therefore the visibility of blocks is lower and the workflow patterns can not be easily recognized. As Zugal et al. [65] investigated, the recognition

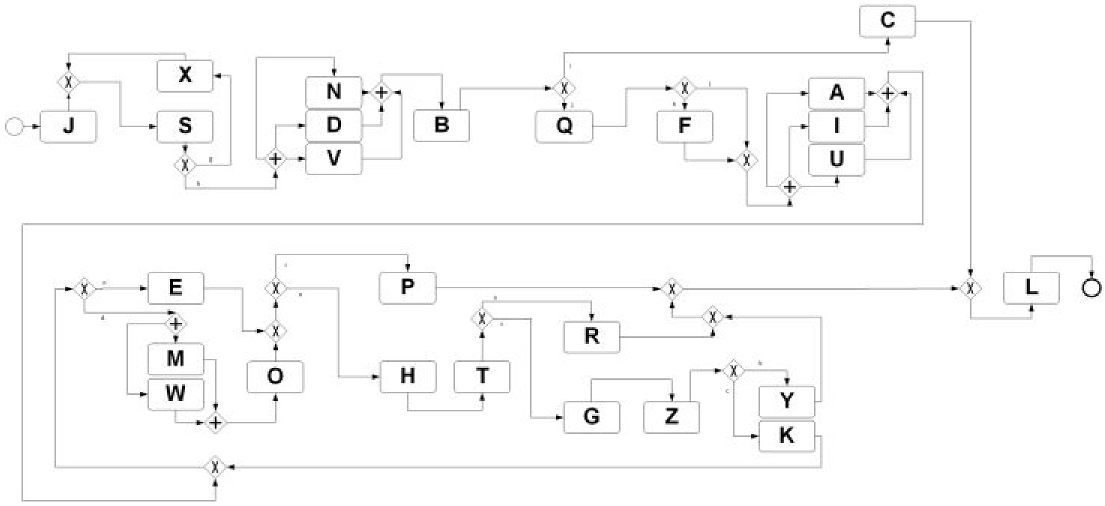

Figure 2.8.: Model 2 in the version for group A.

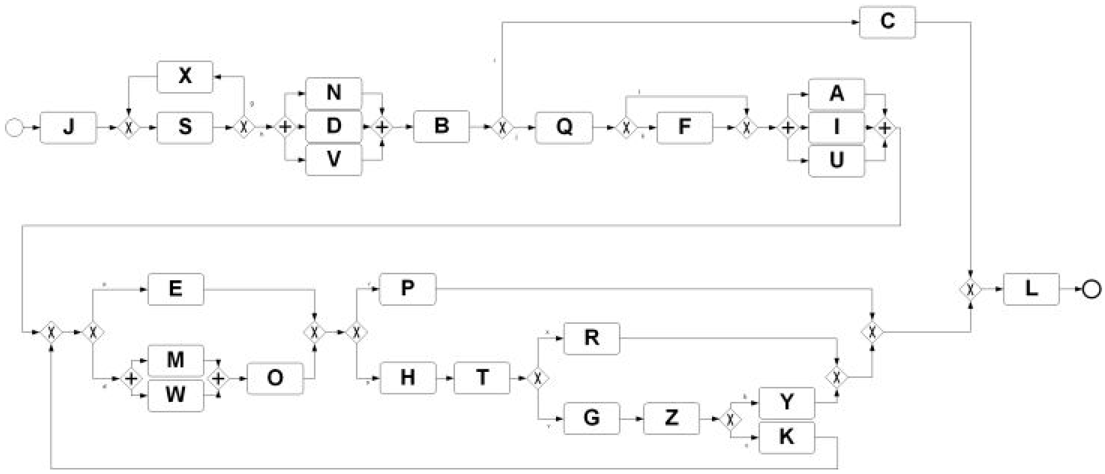

Figure 2.9.: Model 2 in the version for group B

of certain workflow patterns (chunks) in a process model reduces the mental effort for reading and therefore understanding the model [65].

3. Experiment Planning and Execution

In this chapter the setup of the experiment and the most important elements are explained, as well as the experiment design and the execution of the experiment.

1. Experimental Setup

The basic setup of an experiment consists of the consecutive elements [63]:

- Subjects
- Factor and Factor Levels
- Objects
- Parameters
- Response Variables
- Research Questions and Hypotheses

All of these elements are explained in the subsequent section and a visual representation of how these elements work together in an experiment, can be found in Figure 3.1.

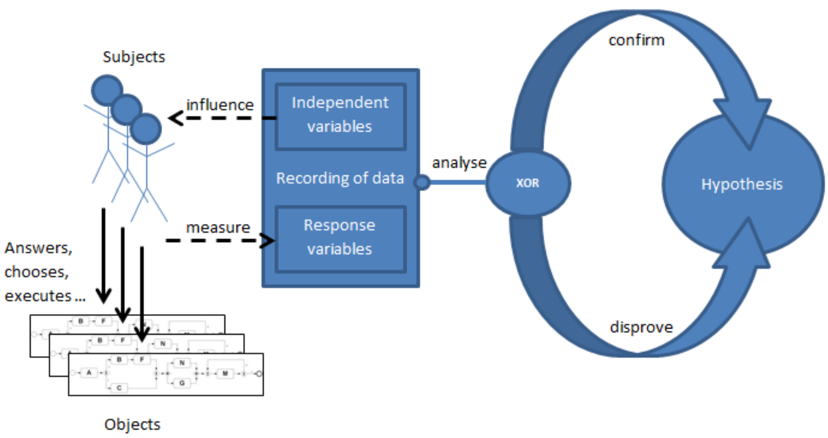

Figure 3.1.: Basic concept of an experiment, adapted from [49].

1.1. Subjects

A subject is a person who applies techniques or - in this experiment - answers questions on the experiment units [20]. The subjects that will take part in this experiment are moderately familiar modelers that are already familiar with basic concepts of BPMN modelling. The basic knowledge is needed, as this experiment aims at the understandability of the BPMN models concerning a specific layout. Overall, 68 subjects were involved in this experiment. The subjects were mainly students of the University of Ulm that either study computer science or economics and took part in the courses *Business Process Management* or *Business Process Intelligence*.

1.2. Factor and Factor Levels

A *factor* or *independent variable* is a characteristic that has an influence on the response variable. The possible values of a factor are called *factor levels* [20]. This experiment has two factors for each layout property that is investigated in this thesis: the **model layout type** and the **question type**. The differentiation into a local and a global *question type* is made as a result of the findings of [59], that changes on process models differ in accuracy depending on whether the changes are sequential or circumstantial, which will be further explained in the section about the *objects* of this experiment.

For the investigation of the layout property *change of direction* there are two factors with two alternatives each:

- The factor *model layout type* has the values **low change of direction** and **high change of direction**.
- The factor *question type* has the values **local question** and **global question**.

For the investigation of the layout property *block visibility* there are two factors with two alternatives each:

- The factor *model layout type* has the values **low block visibility** and **high block visibility**.
- The factor *question type* has the values **local question** and **global question**.

Table 3.1 summarizes the factors and factor levels used in this experiment.

1.3. Objects

The term *object* is used for the objects on which the experiment is run [20]. An object or *experimental unit* in this experiment is the combination of a model and the questions that are asked for this model. Models 1 and 3 were developed to gather information about whether the change of direction has an influence on the understandability of the

Factor	**Factor Level 1**	**Factor Level 2**
Model Layout Type (change of direction)	low	high
Question Type	local	global

Factor	**Factor Level 1**	**Factor Level 2**
Model Layout Type (block visibility)	low	high
Question Type	local	global

Table 3.1.: Factors and factor levels of the experiments on the impact of the two different layout properties.

process model, while models 2 and 4 were developed to collect data to identify the impact of the block visibility on the understandability of the process models.

Table 3.2 shows the different models concerning the layout property *change of direction*. Each model pair consists of one model with a low number of changes in the direction and one model with a high number of changes in the direction. For each model pair six questions are asked - three questions of the local question type and three questions of the global question type.

Object	**Model Layout Type - change of direction**
Model 1a	low number of changes
Model 1b	**high number of changes**
Model 3a	**high number of changes**
Model 3b	low number of changes

Table 3.2.: Models of the experiment concerning the layout property change of direction.

Table 3.3 shows the different models concerning the layout property *block visibility*. Each model pair consists of one model with a low block visibility and one model with a high block visibility. For each model pair six questions are asked - three questions of the local question type and three questions of the global question type.

Object	**Model Layout Type - block visibility**
Model 2a	**low block visibility**
Model 2b	high block visibility
Model 4a	high block visibility
Model 4b	**low block visibility**

Table 3.3.: Models of the experiment concerning the layout property block visibility.

In the following, it is explained how the models and questions were developed for this experiment.

Model Development

As a preparation for the experiment, process model pairs were created which only differ in the investigated layout properties (cf. paragraph parameters for details).

The recommendations of [31] on the modelling of processes were taken into account while the models were created. Each of the model pairs were checked for the following constraints [23]:

- *Correctness and Executability* Mendling and Reijers [32] define soundness and structuredness as important correctness criteria for imperative process models, e.g. BPMN models. Each model pair is sound (no deadlocks or conflicts) and syntactically correct.
- *Representativeness* To ensure the representativeness of the experimental objects and to avoid highly complex models only the core elements of BPMN were used. Each model pair contains the following control flow patterns: sequence, parallel split, synchronization, exclusive choice, simple merge pattern and loop.

Question Development

For each model six questions were developed, which are the same for each model pair. These questions have two different types - three of them are so-called *local questions* while the other three are *global questions*.

As cognitive load theory shows, the working memory of the human brain can not remember more than five [6] or up to seven (plus/minus 2) elements at a time [34]. Therefore we decided to take the number five as our threshold to distinguish between questions that are easily handled by the short term memory and the questions that seem to be harder, as the subject has to remember more than five elements at a time to answer the question. This distinction was developed, as the layout should not have an effect on questions concerning a small part of the process, that can be remembered easily.

The definitions of these two question types are as follows:

- *Local* questions can be answered by remembering less than five activities or gateways.
- *Global* questions can only be answered by remembering more than five activities or gateways.

The models as well as the questions for this experiment can be found in Appendix A.

1.4. Parameters

All the variables and parameters that will not influence the result are called *invariable*. This includes all the controlled parameters, that are fixed to check only the independent variables and their influence [20]. In this experiment the parameters are related to the main categories of factors that influence the understandability of a process model (cf. [32], [29]):

- Personal factors
- Task factors

Personal factors

Personal factors like the individual foreknowledge of each subject in the area of process modelling as well as foreknowledge concerning the domain of the process, can have a great influence on the outcome of the experiment.

Therefore these factors were controlled. The foreknowledge was controlled by selecting subjects for the experiment with uniform experience in process modelling. By selecting students of the same university which all attended the same course, the variation of the foreknowledge of the subjects was reduced. Still the foreknowledge of the subjects can vary, as each subject has an own history of previously visited courses. Therefore a foreknowledge survey was answered by each of the subjects at the beginning of the experiment, to be able to analyse possible impacts of this variation. As the validation of this survey, which can be seen in Section 1.1 of Chapter 4, showed, the foreground knowledge of the groups was homogeneous.

To avoid the impact of a subjects' domain knowledge on the result, the activity labels for the models in this experiment are all single letters, which were randomly assigned.

Task factors

Typical model characteristics that could influence the understandability of a process model - besides the layout properties to be investigated and thus were controlled - are the following factors:

Notation For the model creation BPMN was used for all four models, as it is a well-known standard [38].

Complexity of the model The size of the process models (number of activities, size of gateways) is similar across all model pairs. Additionally each model pair - as the model pairs visualize the same process with a different layout - always have the same number of gateways, edges and activities.

Control-flow patterns All four model pairs contain the core elements and control-flow patterns of BPMN: sequence, parallel split, synchronization, exclusive choice, simple merge pattern and loop.

Layout of the model In all models the influencing factors, like edge crossing and non-orthogonal edges [2] [32] were controlled, except for the ones under investigation. Additionally, it was ensured that all model pairs fit on a screen to avoid scrolling.

Labels To neutralize the impact of the labelling of activities on the understandability of a process model [30], the activity labels for the models in this experiment are all single letters, which were randomly assigned.

1.5. Response Variables

A response variable is the quantitative result of an experiment [20]. To test the hypotheses, the consecutive response variables need to be returned in this experiment:

- **Accuracy** - the number of correctly answered tasks
- **Duration** - the time needed to complete the tasks

1.6. Research Questions and Hypotheses

Research questions are the questions that are asked before the execution of the experiment. The hypotheses are formally stated and are either accepted or rejected after the data analysis [63].

Research Question RQ_0 *Does the change of direction have a significant influence on the understandability of a process model?*

Hypothesis H_1 *The change of direction influences the understandability of a process model concerning the response variable accuracy, for global questions. The more changes there are, the less accurate the answers on the process model get, for global questions.*

Hypothesis H_2 *The change of direction influences the understandability of a process model concerning the response variable accuracy, for local questions. The more changes there are, the less accurate the answers on the process model get, for local questions.*

Hypothesis H_3 *The change of direction influences the understandability of a process model concerning the response variable duration, for global questions. The more changes there are, the longer is the duration of the understanding process, for global questions.*

Hypothesis H_4 *The change of direction influences the understandability of a process model concerning the response variable duration, for local questions. The more changes there are, the longer is the duration of the understanding process, for local questions.*

Research Question RQ_1 *Does the block visibility have a significant influence on the understandability of a process model?*

Hypothesis H_5 *The visibility of blocks influences the understandability of a process model concerning the response variable accuracy, for global questions. The less blocks are visible, the less accurate the answers on the process model get, for global questions.*

Hypothesis H_6 *The visibility of blocks influences the understandability of a process model concerning the response variable accuracy, for local questions. The less blocks are visible, the less accurate the answers on the process model get, for local questions.*

Hypothesis H_7 *The visibility of blocks influences the understandability of a process model concerning the response variable duration, for global questions. The less blocks are visible, the longer is the duration of the understanding process, for global questions.*

Hypothesis H_8 *The visibility of blocks influences the understandability of a process model concerning the response variable duration, for local questions. The less blocks are visible, the longer is the duration of the understanding process, for local questions.*

2. Experimental Design

This thesis conducted two interleaving experiments with two factors each - the model layout and the question type. Subjects are divided into two groups each of which is working on both factor levels of both factors in both experiments. Each subject group receives four models with six questions. For each process model the same questions are asked - three global questions and three local questions.

The experiment is divided into four phases. In each phase, the subject is presented a model which differs depending on the subject group, and six questions on the model, which are the same for both model variants. After the questions for the model are answered, the next model appears and again the questions on the model have to be answered. The four model pairs are distributed to two groups (group A and group B). While group A gets model 1 with a low change of direction, group B gets model 1 with a high change of direction first. Then group A gets model 2 with a low block visibility while group B gets model 2 with a high block visibility. This structure repeats itself with models 3 and 4, but the factor levels are interchanged. Therefore all factor levels are covered, but in a different order, which can be nicely seen in Figure 3.2.

The experiment was conducted with independent groups and the students were randomly assigned to these groups.

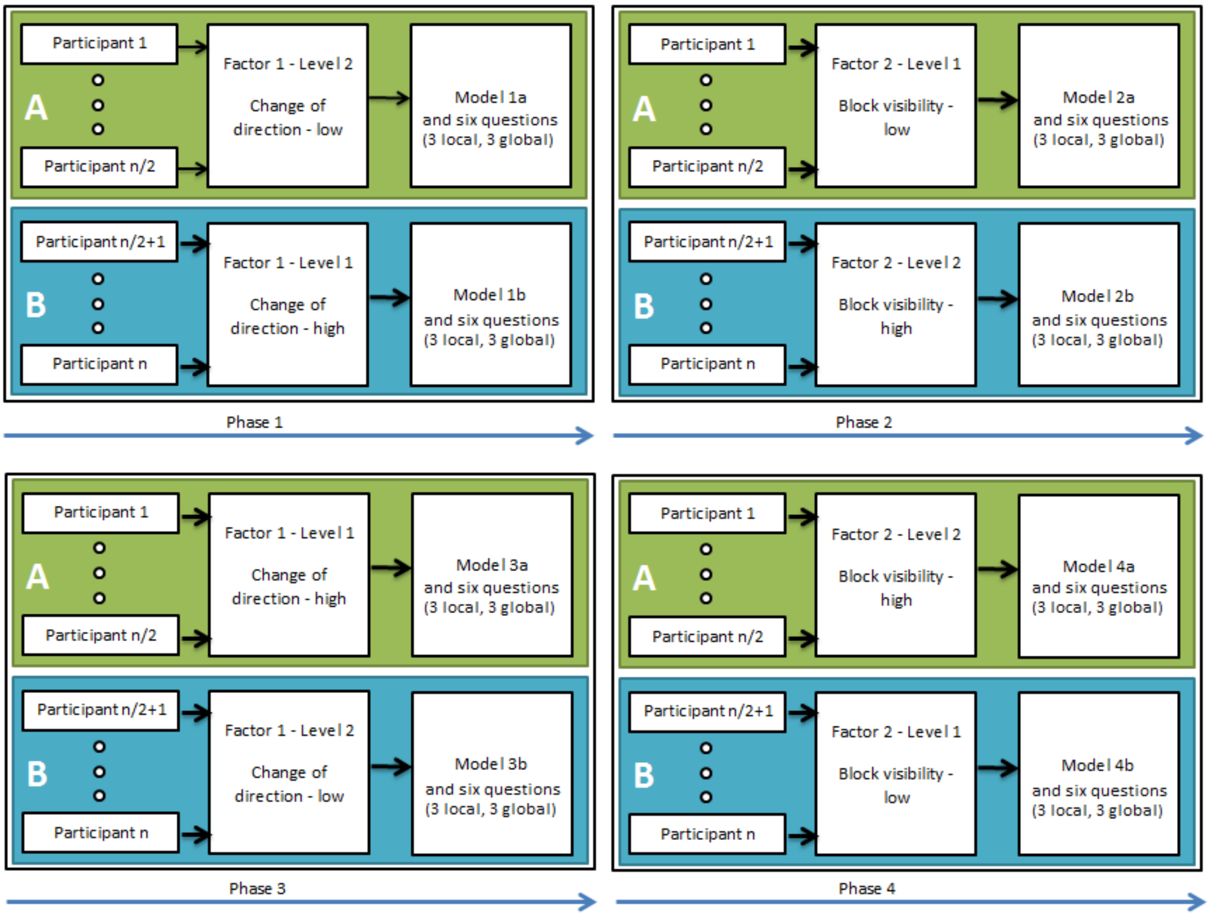

Figure 3.2.: Overall experiment with phases.

3. Experimental Execution

This section describes in detail, how the experiment was prepared and executed.

3.1. Preparation

For the experiment four model pairs were created. Each model pair is equivalent on a semantic level but differs in one layout property. Two model pairs differ in the layout property *change of direction*, while the other two pairs differ in the layout property *block visibility*. For each of these model pairs six questions were constructed - three of them are local questions, while the other three are global questions.

To ensure a balance in the difficulty of the questions and to avoid the influence of too easy or too complex questions, a pilot-test was conducted. The criteria applied for the development of the questions - besides the already mentioned - are the following:

- *Typical constructs* of questions for imperative process models were used to ensure that they cover relevant aspects of understandability.
- The *phrasing* of the question had to be adjusted to the question type, whether the question is local and can be answered by remembering less than five elements or a global question where more than five elements have to be remembered.

To ensure, that the foreknowledge in process modelling is gathered, a foreknowledge test on process modelling was used, which was developed by MMag. Dr. Kathrin Figl [13].

To assure that the models and questions are clear and not misleading and that the questions can be answered in the available amount of time, the pilot-test was conducted. In total, four subjects were asked to report on ambiguities regarding the setup of the experiment. The questions and setup was adapted due to the feedback given. A statistical reliability analysis was not possible, as the number of participants in the pre-test was too small.

3.2. Execution

The experiment was set up with the *Cheetah Experimental Platform (CEP)* [39]. CEP is a platform designed to conduct experiments in the area of business process models and is also able to let subjects model their own processes and analyse the overall modelling process. The gathered data can be exported as an Excel file or SPSS.

The experiment execution started on the 14th of January 2016. The last experiment data was received on the 27th of January 2016. Overall, 68 students from the University of Ulm, who took part in the courses *Business Process Management* or *Business Process Intelligence*, participated in this experiment.

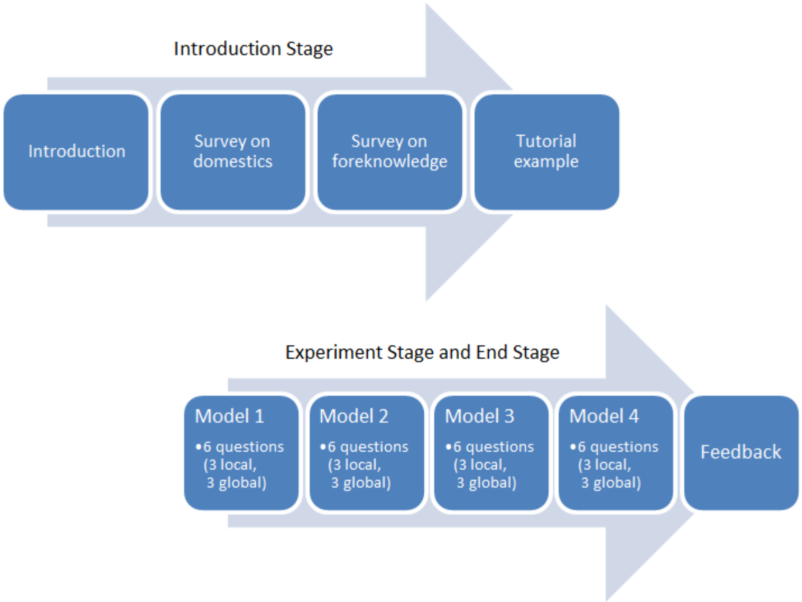

Figure 3.3.: Structure of the experiment.

The experiment structure is shown in Figure 3.3. Each student received a handout with an introduction, hints for the experiment and an explanation on how to download and execute the experiment. The groups were assigned randomly. After the groups were allocated, each subject downloaded CEP, entered the code of the assignment sheet and started the experiment with the answering of a questionnaire on their demographics, followed by a questionnaire on their foreknowledge in process modelling. Then an example model with example questions was given, to show how answering works in CEP. During the experiment stage, each subject had to answer six questions for each of the four models. Finally, the subjects could fill out a feedback form and upload the data.

4. Data Analysis and Interpretation

This chapter explains the analysis of the data, which was gathered during the experiment, and interprets the results of this analysis.

1. Data Analysis

After the execution of the experiment the gathered data has to be analysed to find out, whether the developed hypotheses can be accepted or have to be rejected. To analyse the data the software *SPSS* is used.

The data analysis can be divided into four tasks. First of all the subjects of the experiment and their foreknowledge are analysed, to ensure, that the two groups do not differ concerning their foreknowledge on BPMN. Then the gathered data is validated, to assure the consistency and plausibility of the data. Afterwards the data is further described in the so-called *Descriptive Analysis*. Finally the hypotheses are tested using inferential statistics and conclusions are drawn.

1.1. Subjects

This section gives an overview on the demographics of the subjects of this experiment. In addition, as personal modelling experience and knowledge can have a great impact on the understanding of process models, this section tested whether the two experimental groups did not differ in terms of foreknowledge.

The subjects of this experiment were mainly students of the University of Ulm that either study computer science or economics and took part in the courses *Business Process Management* or *Business Process Intelligence*. The age of the subjects ranges between 21 and 33 years with an average of 25 years. The profession of the subjects is besides four Academics of the University of Ulm, very consistent, as all other participants are students of the University.

As Figure 4.1 shows, the majority of the subjects have started process modelling a year ago or later. While the minimum amount of years is 0, the maximum amount of years is 10, with an average of 1.44 years.

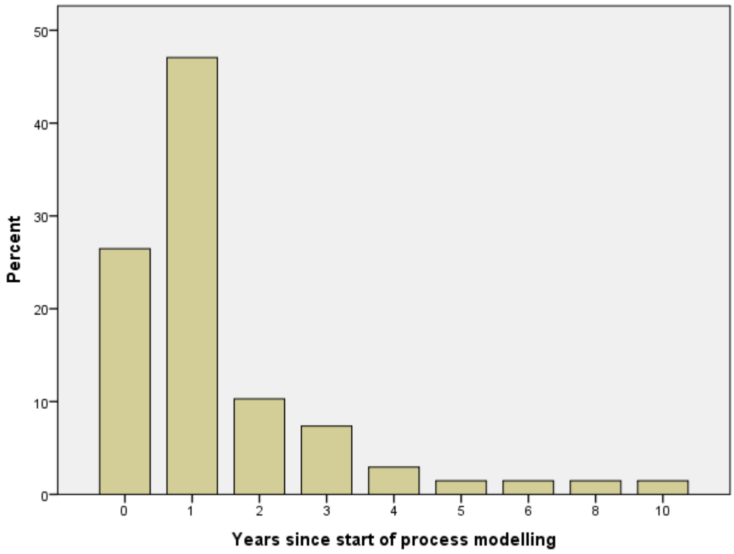

Figure 4.1.: The years since the start of process modelling of each subject.

To assess the foreknowledge of the subjects and to analyse, whether the experience in BPMN varied between the subjects of the two groups, the survey on the subjects foreknowledge in BPMN was conducted. For the analysis of the foreknowledge three items were taken into account:

- the familiarity of the subjects with the modelling notation BPMN
- the competence of the subjects in using BPMN
- the confidence of the subjects in understanding BPMN

The check for internal validity shows a Cronbach's Alpha of 0.936, therefore the scale seems suitable to assess foreknowledge.

Figure 4.2 shows the descriptive statistics of the data given for both groups. The data shows, that there is no significant difference between the two groups concerning their foreknowledge in BPMN. Additionally the data was tested for normal distribution with the help of the *Shapiro-Wilk-Test*. As the data is not normally distributed, a *Mann-Whitney-U-Test* on the mean of the data was executed, which showed no significant difference (p-value of 0.711 (>0.05)) between the two groups. Therefore it can be concluded, that the groups have a homogeneous experience in BPMN modelling and should thus not influence the results of this experiment.

	Variance	Minimum	Maximum	Group	Mean	N	Std. Deviation	Grouped Median	Std. Error of Mean
foreknowledge	1.61	1.33	7	A	3.43	33	1.28	3.24	0.22
				B	3.34	35	1.27	3.07	0.21

Figure 4.2.: Results of the survey on the foreknowledge of the subjects in BPMN by groups.

1.2. Data Validation

As soon as all the subjects completed the experiment, the gathered data was analysed. First of all the data was analysed with the focus on the consistency and plausibility of the data and whether the foreground knowledge was controlled.

The *consistency* of the data was ensured by checking all the available data and looking for missing entries. As all entries were complete, all 68 gathered data logs were taken into account for the data analysis. 33 of the data entries are from Group A, and 35 of the data entries are from Group B.

To ensure the *plausibility* of the gathered data, the data was analysed further with the so-called *box-whisker-plot diagram* [62]. This diagram visualizes all the data and shows possible anomalies and outliers. An outlier is a data point with an extreme value that doesn't fit into a certain range of the other distributed values. In detail the boxplot-diagram shows 50 % of the data in a box, while it eliminates the lower and upper 25 % of the data. The horizontal line inside the box is the median value of the distribution, while the upper and lower ends of the box are the hinges of the distribution. The vertical lines from the ends of the box connect the extreme data points to their respective hinges and are called whiskers [62].

The following Figures show the box-whisker-plot diagram for each model for the response variables duration and accuracy. For each model a total score of six could be reached in accuracy, if all six questions were answered correctly. For the response variable duration no limit was given.

Figure 4.3 shows boxplots for Model 1 for the response variable accuracy and the response variable duration. There are no outliers in Model 1.

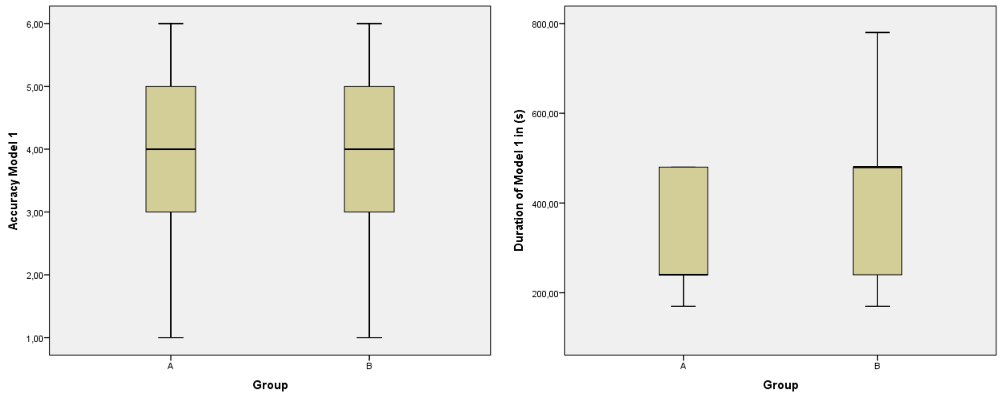

Figure 4.3.: Boxplot-Whisker-Diagram for the Accuracy and Duration of Model 1. In Model 1 the factor change of direction was investigated. While Model 1 for Group A has a low change of direction, Model 1 for Group B has a high change of direction.

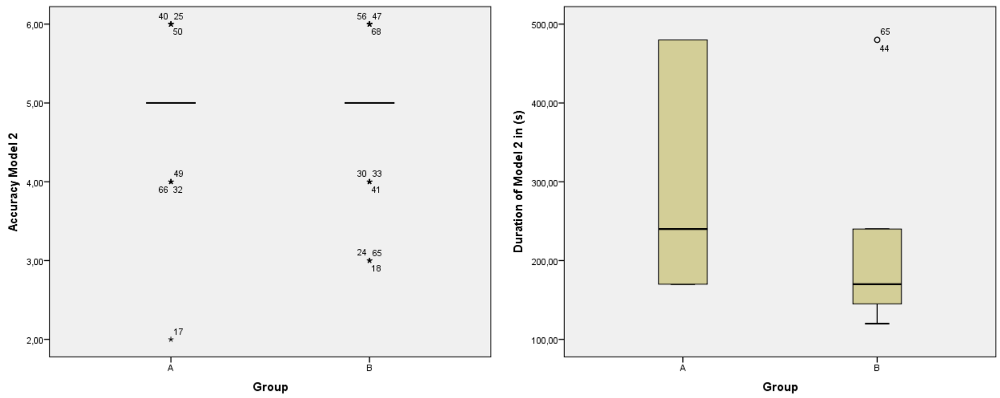

Figure 4.4.: Boxplot-Whisker-Diagram for the Accuracy and Duration of Model 2. In Model 2 the factor block visibility was investigated. While Model 2 for Group A has a low block visibility, Model 2 for Group B has a high block visibility.

Figure 4.4 shows the boxplots for Model 2. The results for the response variable accuracy show a lot of outliers in both groups and no visible box representing the 50 %. In this model the 50 %, typically visualized by a box, all have the same accuracy value of five right answers out of six, the box is represented by the median-line and is not visible as a box. There are no outliers concerning the response variable duration.

Figure 4.5 shows boxplots for Model 3. There is no outlier for the response variable accuracy. The results for the response variable duration have a few outliers in Group A and B.

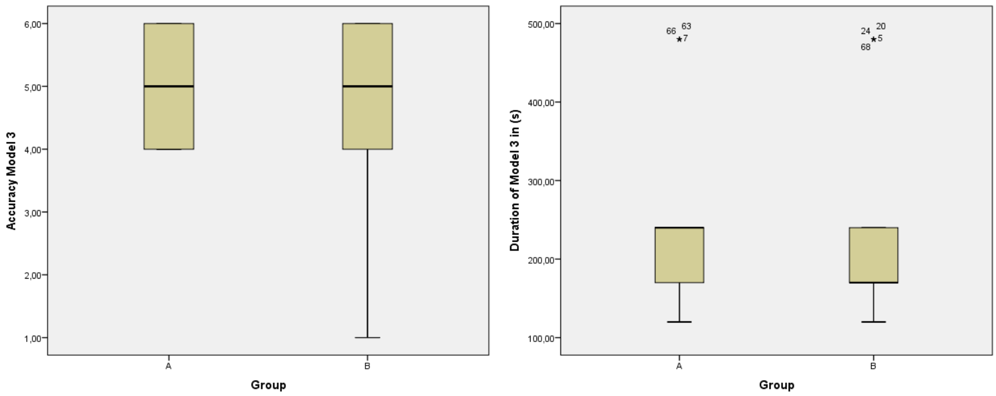

Figure 4.5.: Boxplot-Whisker-Diagram for the Accuracy and Duration of Model 3. In Model 3 the factor change of direction was investigated. While Model 3 for Group A has a high change of direction, Model 3 for Group B has a low change of direction.

Figure 4.6 shows boxplots for Model 4. There is no outlier for the response variable accuracy. The results for the response variable duration have a few outliers in Group A and B.

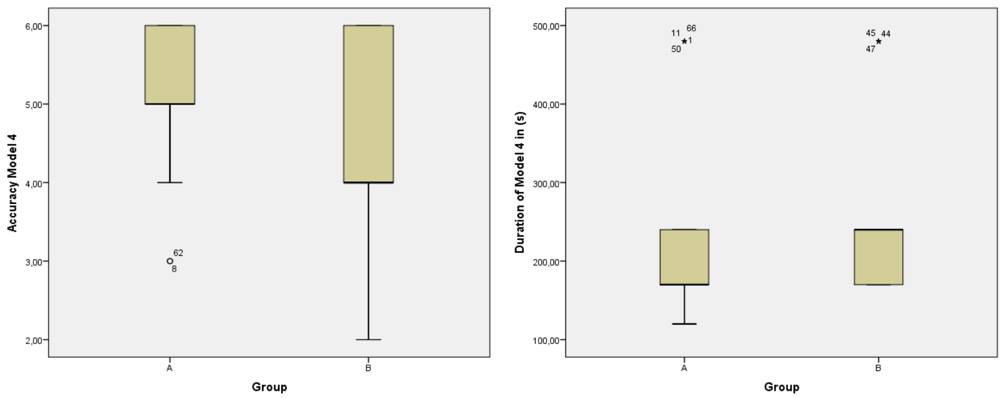

Figure 4.6.: Boxplot-Whisker-Diagram for the Accuracy and Duration of Model 4. In Model 4 the factor block visibility was investigated. While Model 4 for Group A has a high block visibility, Model 4 for Group B has a low block visibility.

The number of outliers was low under the given experimental conditions. In the case of response variable duration, no limit for the completion of the experiment was given and therefore the subjects could take as much time as they needed to complete the experiment. For the response variable accuracy the outliers (extreme values) for all model pairs differed. This suggests that the experimental setup was understood well by all subjects and therefore the outliers were kept in the data for the further analysis.

1.3. Descriptive Analysis

The descriptive analysis looks at all the gathered data of the experiment and shows the amount of data, minimum and maximum values and other statistical data of all the non-nominal response variables.

Response Variable Accuracy

The statistics for the response variable accuracy for the factor layout type can be seen in Figure 4.7. It shows the overall number of entries (N), the average number of reached points (mean), the standard error of mean, the standard deviation, the variance and the minimum and maximum values that occurred.

Model 1 and Model 3 have a changed layout concerning the change of direction, while Model 2 and 4 have a changed layout concerning the visibility of blocks. For the factor change of direction, Model 1 of Group A and Model 3 of Group B were taken into account for the low change of direction (i.e. factor level low), while Model 1 of Group B and Model 3 of Group A were taken into account for the high change of direction (i.e. factor level high). For the factor block visibility, Model 2 of Group A and Model 4 of Group B were taken into account for the low block visibility (i.e. factor level low), while Model 2 of Group B and Model 4 of Group A were taken into account for the high block visibility (i.e. factor level high). This can also be seen in Figure 4.7.

	Models	N		Mean	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
		Valid	Missing						
high change of direction	1 of Goup B, 3 of Group A	68	0	4.40	0.15	1.21	1.47	1	6
low change of direction	1 of Group A, 3 of Group B	68	0	4.38	0.15	1.25	1.55	1	6
low block visibility	2 of Group A, 4 of Group B	68	0	4.69	0.14	1.15	1.32	2	6
high block visibility	2 of Group B, 4 of Group A	68	0	5.04	0.11	0.89	0.79	3	6

Figure 4.7.: Descriptive Analysis of Accuracy for the different layout factors. The change of direction factor was investigated in models 1 and 3, while the block visibility factor was investigated in models 2 and 4.

Based on this descriptive statistic, the following can be derived:

- *The mean accuracy of the answers to the questions is higher for models with a high change of direction than for models with a low change of direction, regardless of the question type.*

- *The mean accuracy of the answers to the questions is higher for models with a high block visibility than for models with a low block visibility, regardless of the question type.*

In Figure 4.8, the different descriptive statistics for the factor question type can be seen.

	N		Mean	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
	Valid	Missing						
Accuracy local	68	0	10.49	0.18	1.52	2.31	6	12
Accuracy global	68	0	8.03	0.27	2.21	4.90	1	12

Figure 4.8.: Descriptive Analysis of Accuracy for the different question types.

Based on this descriptive statistic, the following can be derived:

The mean accuracy of the answers to the questions is higher for local questions than for global questions, regardless of the model layout type.

Model Type	Question Type	N		Mean	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
		Valid	Missing						
high change of direction	local	68	0	2.56	0.08	0.70	0.49	0	3
	global	68	0	1.84	0.10	0.82	0.68	0	3
low change of direction	local	68	0	2.59	0.08	0.67	0.46	1	3
	global	68	0	1.79	0.11	0.87	0.76	0	3

Figure 4.9.: Descriptive Analysis of Accuracy for the layout change of direction and the different question types.

Figure 4.9 shows the descriptive statistics for the layout factor change of direction and the question type factor. Based on this descriptive statistic, the following can be derived:

- *Regardless of the change of direction, local questions always yield higher accuracy when compared to global questions.*
- *For the factor change of direction we cannot say that a low change of direction results in a lower accuracy in general, only for global questions.*

Figure 4.10 shows the descriptive statistics for the layout factor block visibility and the question type factor. Based on this descriptive statistic, the following can be derived:

- *Regardless of the block visibility, local questions always yield higher accuracy when compared to global questions.*
- *High block visibility results in a higher accuracy within the same question type for both local and global questions.*

Model Type	Question Type	N		Mean	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
		Valid	Missing						
low block visibility	local	68	0	2.66	0.07	0.56	0.32	1	3
	global	68	0	2.03	0.10	0.85	0.72	0	3
high block visibility	local	68	0	2.68	0.07	0.58	0.34	1	3
	global	68	0	2.37	0.07	0.54	0.30	1	3

Figure 4.10.: Descriptive Analysis of Accuracy for the layout block visibility and the different question types.

Response Variable Duration

The statistics for the response variable duration for the factor layout type (i.e., change of direction and block visibility respectively) can be seen in Figure 4.11. It shows the overall number of entries (N), the average amount of minutes (mean), the standard error of mean, the standard deviation, the variance and the minimum and maximum values that occurred.

Model 1 and Model 3 have a changed layout concerning the change of direction, while Model 2 and 4 have a changed layout concerning the visibility of blocks. For the factor change of direction, Model 1 of Group A and Model 3 of Group B were taken into account for the low change of direction (i.e. factor level low), while Model 1 of Group B and Model 3 of Group A were taken into account for the high change of direction (i.e. factor level high). For the factor block visibility, Model 2 of Group A and Model 4 of Group B were taken into account for the low block visibility (i.e. factor level low), while Model 2 of Group B and Model 4 of Group A were taken into account for the high block visibility (i.e. factor level high). This can also be seen in Figure 4.11.

	Models	N		Mean (minutes)	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
		Valid	Missing						
high change of direction	1 of Goup B, 3 of Group A	68	0	3.78	0.18	1.44	2.09	1.83	12.45
low change of direction	1 of Group A, 3 of Group B	68	0	3.36	0.14	1.18	1.39	1.71	7.22
low block visibility	2 of Group A, 4 of Group B	68	0	3.46	0.12	0.97	0.94	2.02	7.05
high block visibility	2 of Group B, 4 of Group A	68	0	2.79	0.10	0.85	0.72	1.10	5.60

Figure 4.11.: Descriptive Analysis of Duration for the different layout factors. The change of direction factor was investigated in models 1 and 3, while the block visibility factor was investigated in models 2 and 4.

Based on this descriptive statistics, the following can be derived:

- *The mean duration of the answers to the questions is higher for models with a high change of direction than for models with a low change of direction, regardless of the question type.*

- *The mean duration of the answers to the questions is higher for models with a low block visibility than for models with a high block visibility, regardless of the question type.*

	N		Mean (minutes)	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
	Valid	Missing						
Duration local	68	0	5.54	0.18	1.47	2.15	3.08	10.20
Duration global	68	0	7.84	0.26	2.16	4.68	5.10	20.92

Figure 4.12.: Descriptive Analysis of Duration for the different questions.

Figure 4.12 shows the descriptive statistics for the different question types. Based on this descriptive statistic, the following can be derived:

The mean duration of the answers to the questions is higher for global questions than for local questions, regardless of the model layout type.

Model Type	Question Type	N		Mean (minutes)	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
		Valid	Missing						
high change of direction	local	68	0	1.50	0.09	0.75	0.56	0.53	3.88
	global	68	0	2.28	0.11	0.94	0.88	1.13	8.82
low change of direction	local	68	0	1.34	0.09	0.70	0.49	0.51	3.55
	global	68	0	2.01	0.09	0.75	0.56	0.84	5.19

Figure 4.13.: Descriptive Analysis of Duration for the layout change of direction and the different question types.

Figure 4.13 shows the descriptive statistics for the layout factor change of direction and the question type factor. Based on this descriptive statistic, the following can be derived:

- *Regardless of the change of direction, local questions always yield lower duration values when compared to global questions.*

- *For the factor change of direction we can say that a high change of direction results in a higher duration value within the same question type.*

Model Type	Question Type	N		Mean (minutes)	Std. Error of Mean	Std. Deviation	Variance	Minimum	Maximum
		Valid	Missing						
low block visibility	local	68	0	1.51	0.06	0.47	0.22	0.45	2.97
	global	68	0	1.95	0.09	0.70	0.50	0.95	4.31
high block visibility	local	68	0	1.18	0.05	0.42	0.17	0.52	2.40
	global	68	0	1.60	0.07	0.58	0.34	0.58	3.54

Figure 4.14.: Descriptive Analysis of Duration for the layout block visibility and the different question types.

Figure 4.14 shows the descriptive statistics for the layout factor block visibility and the question type factor. Based on this descriptive statistic, the following can be derived:

- *Regardless of the block visibility, local questions always yield lower duration values when compared to global questions.*
- *For the factor block visibility we can say that a low block visibility results in a higher duration value within the same question type.*

In the next step we will test whether the observed differences are statistically significant.

1.4. Testing the Hypotheses

In this section the hypotheses have to be statistically tested to reject or accept them. There are two methods to find out, whether a finding is significant: *parametric* and *non-parametric* tests. While parametric tests are typically used for interval data or ratio data, non-parametric tests are used for nominal data and ordinal/rank-order data [51].

With a non-parametric test the detection of a significant effect in the response variable is difficult. Therefore parametric tests are preferred for the analysis of the data of the experiment. To carry out a parametric test, certain constraints have to be fulfilled by the data [20]. For the execution of a parametric test, the data has to have a normal distribution and a homogeneity of variance. If the data doesn't fulfill both assumptions, a non-parametric test has to be conducted, which is less sensitive than the parametric test.

Therefore the normal distribution is tested first. To test, if the data is normally distributed, a so-called *Shapiro-Wilk* test [43] can be conducted. In Figure 4.15 the result of the Shapiro-Wilk test is shown. As the outcome of the normal distribution test results in a value lower than 0.05 for most of the data, the data is not distributed normally. The data can not be analysed with the help of a parametric test since one of the mandatory assumptions is not fulfilled.

As the data has no normal distribution, a non-parametric test has to be used to find out whether there are significant results. The chosen non-parametric test is the so-called *Wilcoxon-Test*. If the obtained result has a significance value lower than 0.05, the difference is significant. If the significance value is higher than 0.05, the difference is not significant.

Type		Measurement	Shapiro-Wilk			Normal Distribution
			Statistic	df	Sig.	
high change of direction		Accuracy	0.90	68	0.00	no
		Duration	0.73	68	0.00	no
low change of direction		Accuracy	0.89	68	0.00	no
		Duration	0.94	68	0.00	no
low block visibility		Accuracy	0.86	68	0.00	no
		Duration	0.90	68	0.00	no
high block visibility		Accuracy	0.83	68	0.00	no
		Duration	0.97	68	0.07	yes
local questions		Accuracy	0.86	68	0.00	no
		Duration	0.93	68	0.00	no
global questions		Accuracy	0.97	68	0.06	yes
		Duration	0.73	68	0.00	no
high change of direction	local	Accuracy	0.66	68	0.00	no
		Duration	0.92	68	0.00	no
	global	Accuracy	0.86	68	0.00	no
		Duration	0.57	68	0.00	no
low change of direction	local	Accuracy	0.63	68	0.00	no
		Duration	0.88	68	0.00	no
	global	Accuracy	0.87	68	0.00	no
		Duration	0.89	68	0.00	no
low block visibility	local	Accuracy	0.62	68	0.00	no
		Duration	0.93	68	0.00	no
	global	Accuracy	0.83	68	0.00	no
		Duration	0.91	68	0.00	no
high block visibility	local	Accuracy	0.59	68	0.00	no
		Duration	0.92	68	0.00	no
	global	Accuracy	0.71	68	0.00	no
		Duration	0.96	68	0.02	no

Figure 4.15.: Result of the *Shapiro-Wilk* test for the normal distribution of the data. The variable *df* shows the number of elements, while the variable *Sig.* shows the significance value of the result.

In the following, the results of the *Wilcoxon-Test* [5] are given for each of the response variables accuracy and duration.

Response Variable Accuracy

For the response variable *Accuracy*, the following results were reached.

For factor change of direction: With a p-value of 0.825 (>0.05) for the global question type (H_1) and a p-value of 0.898 (>0.05) for the local question type (H_2), hypothesis H_1 and H_2 can be rejected. There is no significant difference between the results for the global questions type nor for the local questions type.

For factor block visibility: With a p-value of 0.001 (<0.05) for the global question type, hypothesis H_5 can be accepted, i.e. there is a significant difference between the results in the global questions type. As this difference does not occur in the local questions type (p-value of 0.908 (>0.05)), hypothesis H_6 can be rejected.

Response Variable Duration

For the response variable *Duration*, the following results were reached.

For factor change of direction: With a p-value of 0.008 (<0.05) for the global question type, hypothesis H_3 can be accepted, i.e. there is a significant difference between the results in the global questions type. As this difference does not occur in the local questions type (p-value of 0.187 (>0.05)), hypothesis H_4 can be rejected.

For factor block visibility: With a p-value of 0.003 (<0.05) for the global question type (H_7), and a p-value of 0.000 (<0.05) for the local question type (H_8), hypothesis H_7 and H_8 can be accepted, as there is a significant difference between the results in the global questions type and in the local questions type. Therefore, the block visibility has a significant influence on the understandability of a process model for the response variable duration, regardless of the question type.

2. Interpretation of Results

In this section the results of the data analysis are interpreted.

Testing Hypothesis H_1, H_2, H_3, H_4

For the response variable accuracy, the hypothesis that more changes in direction yield in less accurate answers to the questions on the model, could be rejected for both question types. Therefore a high change of direction does not influence the accuracy of the answers on a model.

The hypothesis, that a high change of direction results in a longer duration to answer questions of the global question type, could be confirmed. While the hypothesis, that a high change of direction extends the duration for answering local questions, could not be confirmed by this test.

Testing Hypothesis H_5, H_6, H_7, H_8

The hypothesis, that the lower the block visibility, the less accurate the answers to the questions on a model, could be confirmed for the global question type. For the hypothesis, that the lower block visibility results in less accurate answers on questions on a model for the local question type, could not be confirmed. Therefore a low block visibility results in less accurate answers only for global questions.

The hypothesis, that a low block visibility results in a longer duration to answer questions on a model, could be confirmed for both the local and the global question type. Therefore the low block visibility has a significant influence on the duration of the answering of questions on a process model, regardless of the question type.

5. Conclusion

This thesis described the preparation and execution of an experiment on the impact of the model layout on the understandability of a process model. The experiment looked at the two layout properties change of direction and block visibility and validated, whether they have an impact on the understandability of a process model (accuracy and duration).

1. Main Findings

In this section, the main findings of this thesis and the results of the research questions are presented.

1.1. Change of Direction

The first research question tries to find out, whether the change of direction has a significant influence on the understandability of a process model. Regardless of the question type, a high change of direction did not decrease the accuracy of the answers on the questions of the model. A high change of direction increases the duration to answer questions on the model for the global question type, but not for the local question type.

1.2. Block Visibility

The second research question tries to find out, whether the block visibility has a significant influence on the understandability of a process model. The low block visibility only decreased the accuracy of the answers for questions of the global question type. Regardless of the question type, a low block visibility did increase the duration to answer questions on the model.

2. Research Limitations

The research of this thesis has a few limitations, which are discussed in the following.

The external validity is the degree to which the results of the experiment can be generalized beyond the experimental setup [51].

The majority of the subjects was rather new to the topic of process modelling, as the average amount of years in BPMN modelling is at 1.44 years, and therefore the results of this experiment cannot be generalized to experts in process modelling. Also the small number of participants in the pre-test is a limitation which results in a lower internal consistency, than the internal consistency with more participants.

Due to our considerations on the concentration span of the subjects, we decided to conduct this experiment with a few, well planned models and a few questions. Therefore we cannot generalize to models and questions in general. Still we tried to cover all essential patterns and considered different types of questions.

To counter these and other limitations, it is recommended to repeat the experiment with a larger number of models and questions and additionally it is recommended to repeat it with experts in process modelling, to be able to generalize the results of this experiment to experts.

3. Outlook

Future work in this area of research is needed to fill the lack of empirical insight on the understandability of process models.

It would be interesting to conduct an experiment concentrating on the impact of the question type on the understandability of a process model, as it seems to have a significant effect on the accuracy and duration of the answers given. Global questions, where the subject has to remember more than five activities or gateways seem to be more dependent on the layout than local questions that can be answered by looking at a small part of the model.

A. Models and questions of the experiment

The experiment can be found under the following links: `http://bpm.q-e.at/tmp/CEP_Layout_10122015_x86.exe` (32 bit) and `http://bpm.q-e.at/tmp/CEP_Layout_10122015_x64.exe` (64 bit).

1. Models

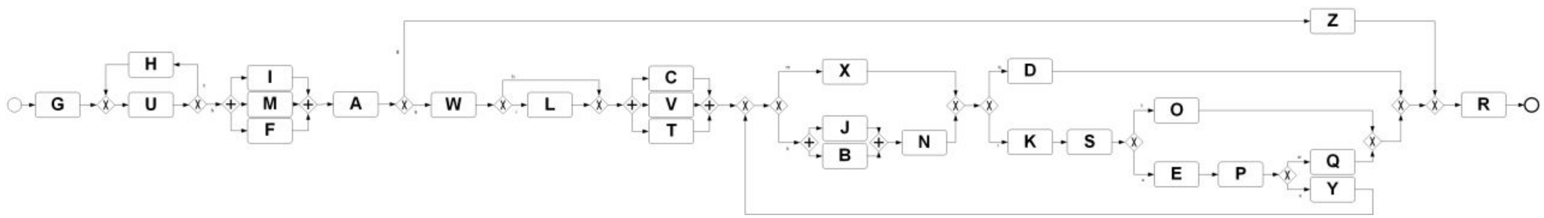

Figure A.1.: Model 1 of group A.

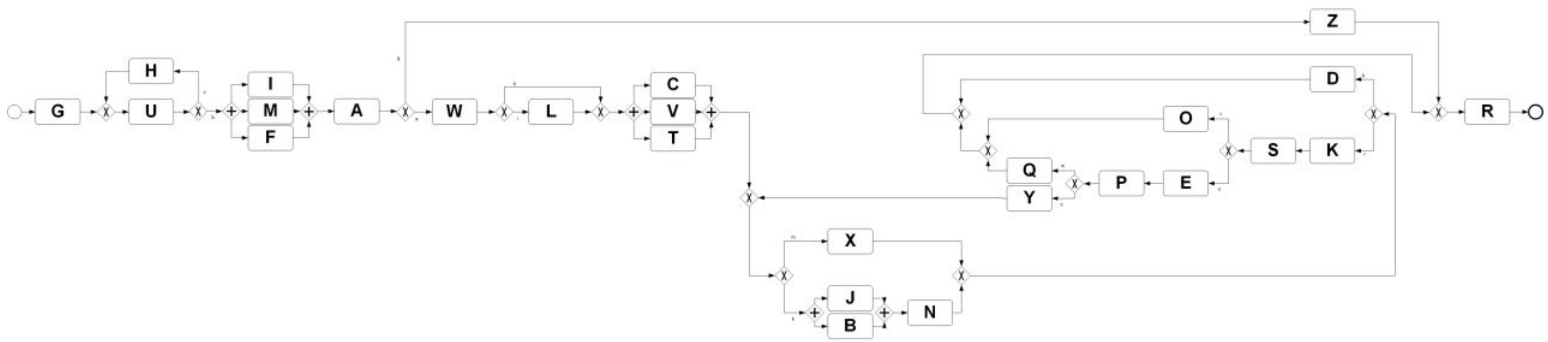

Figure A.2.: Model 1 of group B.

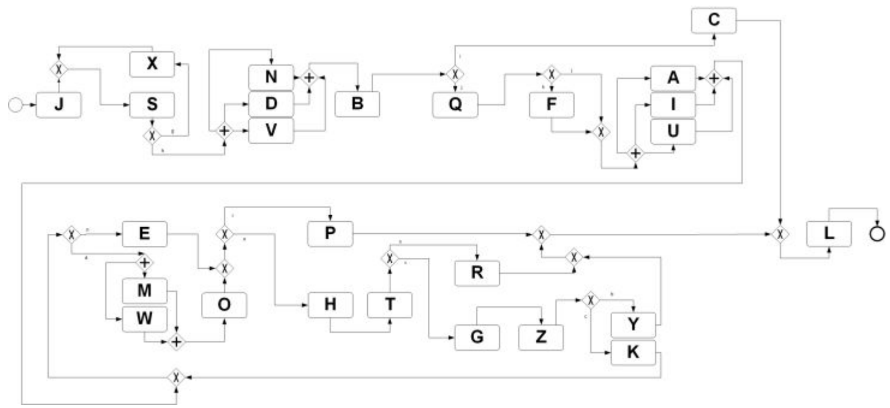

Figure A.3.: Model 2 of group A.

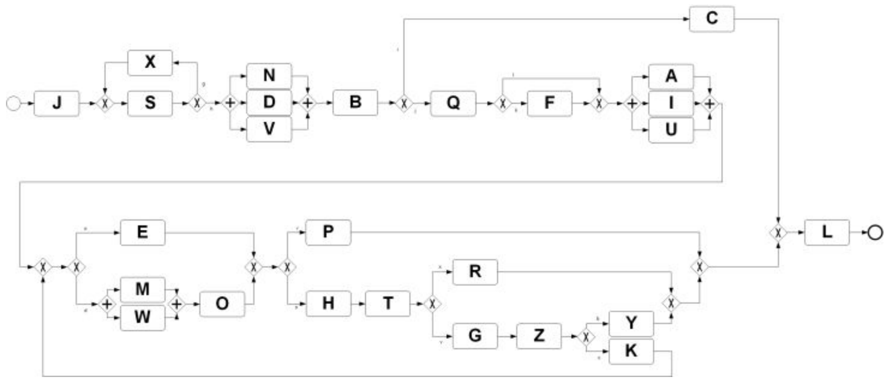

Figure A.4.: Model 2 of group B.

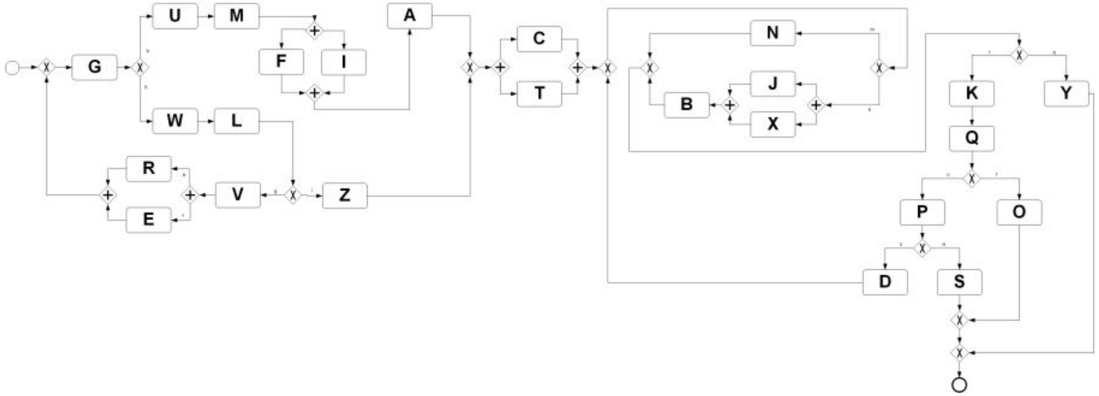

Figure A.5.: Model 3 of group A.

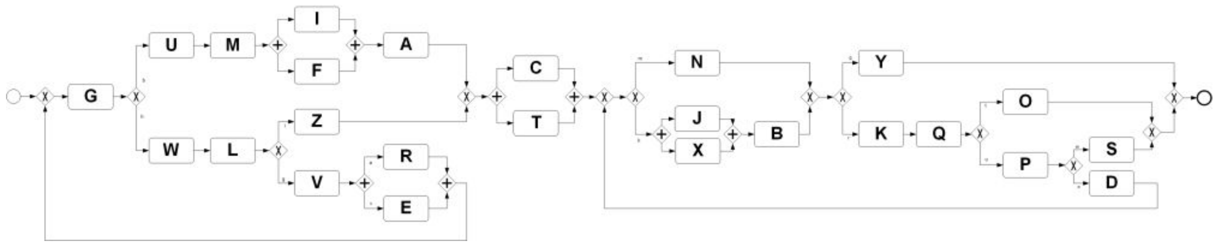

Figure A.6.: Model 3 of group B.

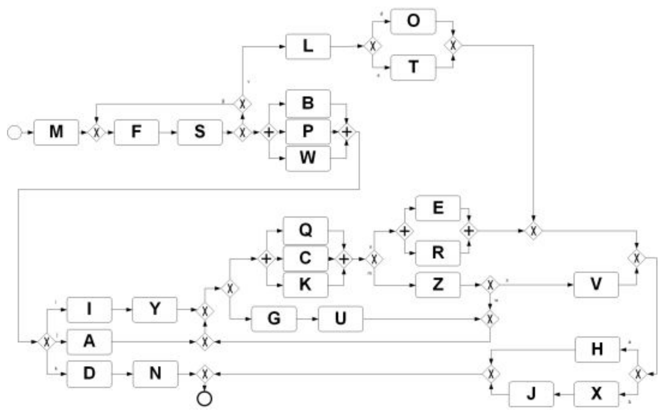

Figure A.7.: Model 4 of group A.

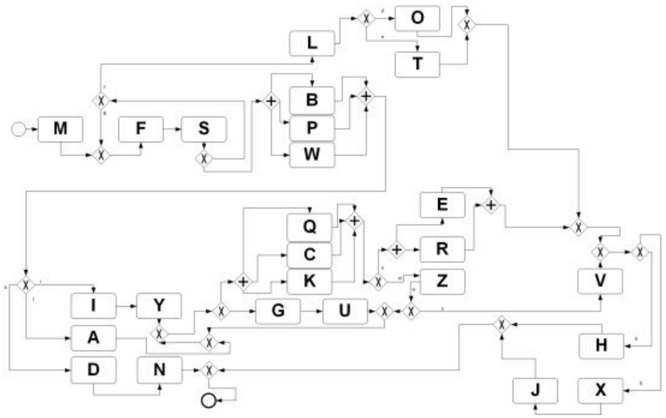

Figure A.8.: Model 4 of group B.

2. Questions

2.1. Questions for model 1

- Q is always directly followed by R.
- <Y, J, B> is a valid sub-trace.
- Either O or D can be executed, but not both.
- In every trace with J and Q, X can't occur.
- If S and Y are part of a trace, S can't occur more often than Y.
- If O occurs in a trace, N and X can't be both in it.

2.2. Questions for model 2

- N and B can be executed in parallel.
- <H, T, R> is a valid sub-trace.
- X can be executed several times for a single trace.
- If W appears in a trace, R can't be in it.
- If the sequence <T,R> appears in a trace, G can't be in it.
- If Q is executed, W is always executed as well.

2.3. Questions for model 3

- V can be executed more than once in a single trace.
- Y and K can be executed in parallel.
- <W, L, V> is a valid sub-trace.
- If B and Y are both part of a trace, B can't occur more often than Y.
- There is a trace, where C occurs more often than N.
- There is no trace, where D can happen at any point after the execution of U.

2.4. Questions for model 4

- <F, S, L, T> is a valid sub-trace.
- I and A can be executed in parallel.
- K can be performed several times for a single trace.
- If C is part of a trace, it always co-occurs with J.
- R and V can never occur in the same trace.
- If the sequence <F, S> has been executed, only L or F can be executed directly afterwards.

B. Metric results

Figure B.1 shows the results of the BP-Metric calculation for all four models and for each group.

	Alignment	Direction [%]				Model flow consistency
		West	East	North	South	
Model 1a	0.264	0.051	0.949	0.500	0.500	0.932
Model 1b	0.204	0.288	0.712	0.471	0.529	0.795
Model 2a	0.081	0.122	0.878	0.521	0.479	0.727
Model 2b	0.169	0.068	0.932	0.500	0.500	0.727
Model 3a	0.072	0.490	0.510	0.326	0.674	0.684
Model 3b	0.117	0.036	0.964	0.526	0.474	0.868
Model 4a	0.096	0.150	0.850	0.462	0.538	0.708
Model 4b	0.062	0.203	0.797	0.511	0.489	0.708

Figure B.1.: Metric results of all models and groups.

In Figure B.1 the basic data of each of the models is given: The *Alignment* value of the model indicates the amount of aligned elements in the model. The *Direction* values indicate the percentage of edges going into a certain direction (North, South, East or West). The *Model Flow consistency* value indicates the percentage of the model that has the same flow as the overall model - the more consistent the flow is, the less changes of direction there are in the model. This value is the resulting value of the BP-metric.

List of Figures

List of Tables

Bibliography

[1] Ruth Sara Aguilar-Saven. Business Process Modelling: Review and Framework. *International Journal of production economics*, 90(2):129–149, 2004.

[2] Vered Bernstein and Pnina Soffer. Identifying and Quantifying Visual Layout Features of Business Process Models. In *Enterprise, Business-Process and Information Systems Modeling*, pages 200–213. Springer, 2015.

[3] Andrea Burattin, Vered Bernstein, Manuel Neurauter, Pnina Soffer, and Barbara Weber. Detection and Quantification of Flow Consistency in Business Process Models, 2016.

[4] Jorge Cardoso. *Handbook of Research On Business Process Modeling*. IGI Global, 2009.

[5] Gregory W Corder and Dale I Foreman. *Nonparametric Statistics: A Step-by-Step Approach*. John Wiley & Sons, 2014.

[6] Nelson Cowan. The Magical Mystery Four: How Is Working Memory Capacity Limited, And Why? *Current directions in psychological science*, 19(1):51–57, 2010.

[7] Islay Davies, Peter Green, Michael Rosemann, Marta Indulska, and Stan Gallo. How do Practitioners Use Conceptual Modeling In Practice? *Data & Knowledge Engineering*, 58(3):358–380, 2006.

[8] Juliane Dehnert and Wil MP Van Der Aalst. Bridging the Gap Between Business Models and Workflow Specifications. *International Journal of Cooperative Information Systems*, 13(03):289–332, 2004.

[9] Marlon Dumas and Arthur HM Ter Hofstede. UML Activity Diagrams as a Workflow Specification Language. In *UML 2001 - The Unified Modeling Language. Modeling Languages, Concepts, and Tools*, pages 76–90. Springer, 2001.

[10] Philip Effinger and Gero Decker. Layout Techniques Coupled With Web2.0-based Business Process Modeling. In *Graph Drawing*, pages 417–418. Springer, 2009.

[11] Philip Effinger, Nicole Jogsch, and Sandra Seiz. On a Study of Layout Aesthetics for Business Process Models Using BPMN. In *Business Process Modeling Notation*, pages 31–45. Springer, 2010.

[12] Kathrin Figl and Mark Strembeck. On the Importance of Flow Direction in Business Process Models. In *9th International Conference on Software Engineering and Applications*, 2014.

[13] Kathrin Figl and B. Weber. Individual Creativity in Designing Business Processes. In *Proc. HC-PAIS '12*, pages 294–306, 2012.

[14] Jakob Freund and Bernd Rücker. *Praxishandbuch BPMN 2.0.* Carl Hanser Verlag GmbH Co KG, 2014.

[15] Volker Gruhn and Ralf Laue. Complexity Metrics for Business Process Models. In *9th international conference on business information systems (BIS 2006)*, volume 85, pages 1–12. Citeseer, 2006.

[16] Thomas Gschwind, Jakob Pinggera, Stefan Zugal, Hajo A Reijers, and Barbara Weber. Edges, Structures, and Constraints: the Layout of Business Process Models. Technical report, Technical Report RZ 3825, IBM Research, Zurich, 2011.

[17] Thomas Gschwind, Jakob Pinggera, Stefan Zugal, Hajo A Reijers, and Barbara Weber. A Linear Time Layout Algorithm for Business Process Models. *Journal of Visual Languages & Computing*, 25(2):117–132, 2014.

[18] Cornelia Haisjackl, Jakob Pinggera, Pnina Soffer, Stefan Zugal, Shao Yi Lim, and Barbara Weber. Identifying Quality Issues in BPMN Models: an Exploratory Study. In *Enterprise, Business-Process and Information Systems Modeling*, pages 217–230. Springer, 2015.

[19] Michael Himsolt. Comparing and Evaluating Layout Algorithms within GraphEd. *Journal of Visual Languages and Computing*, 6(3):255–273, 1995.

[20] Natalia Juristo and Ana M Moreno. *Basics of Software Engineering Experimentation.* Springer Science & Business Media, 2013.

[21] Ingo Kitzmann, Christoph König, Daniel Lübke, and Leif Singer. A Simple Algorithm for Automatic Layout of BPMN Processes. In *Commerce and Enterprise Computing, 2009. CEC'09. IEEE Conference on*, pages 391–398. IEEE, 2009.

[22] John Krogstie. *Model-based Development and Evolution of Information Systems: A Quality Approach.* Springer Science & Business Media, 2012.

[23] John Krogstie, Guttorm Sindre, and Håvard Jørgensen. Process Models Representing Knowledge for Action: a Revised Quality Framework. *European Journal of Information Systems*, 15(1):91–102, 2006.

[24] Beate List and Birgit Korherr. An Evaluation of Conceptual Business Process Modelling Languages. In *Proceedings of the 2006 ACM symposium on Applied computing*, pages 1532–1539. ACM, 2006.

[25] A. Malesevic, D. Brdjanin, and S. Maric. Tool for Automatic Layout of Business Process Model Represented by UML Activity Diagram. In *EUROCON, 2013 IEEE*, pages 537–542, July 2013.

[26] Richard E Mayer. *Multimedia Learning*. Cambridge university press, 2009.

[27] Jan Mendling. *Event-driven Process Chains (epc)*. Springer, 2008.

[28] Jan Mendling. Metrics for Process Models: Empirical Foundations of Verification. *Error Prediction, and Guidelines for Correctness*, 6, 2008.

[29] Jan Mendling, Hajo A Reijers, and Jorge Cardoso. What Makes Process Models Understandable? In *Business Process Management*, pages 48–63. Springer, 2007.

[30] Jan Mendling, Hajo A Reijers, and Jan Recker. Activity Labeling in Process Modeling: Empirical Insights and Recommendations. *Information Systems*, 35(4):467–482, 2010.

[31] Jan Mendling, Hajo A Reijers, and Wil MP van der Aalst. Seven Process Modeling Guidelines (7PMG). *Information and Software Technology*, 52(2):127–136, 2010.

[32] Jan Mendling and Mark Strembeck. Influence Factors of Understanding Business Process Models. In *Business information systems*, pages 142–153. Springer, 2008.

[33] Jan Mendling, Mark Strembeck, and Jan Recker. Factors of Process Model Comprehension - Findings From a Series of Experiments. *Decision Support Systems*, 53(1):195–206, 2012.

[34] George A Miller. The Magical Number Seven, Plus Or Minus Two: Some Limits On Our Capacity For Processing Information. *Psychological review*, 63(2):81, 1956.

[35] Simon K Milton, Jayantha Rajapakse, and Ron Weber. Ontological Clarity, Cognitive Engagement, and Conceptual Model Quality Evaluation: An Experimental Investigation. *Journal of the Association for Information Systems*, 13(9):657, 2012.

[36] Saurabh Mittal and José L Risco Martín. *Netcentric System of Systems Engineering With DEVS Unified Process*. CRC Press, 2013.

[37] Daniel L Moody. The physics of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. *Software Engineering, IEEE Transactions on*, 35(6):756–779, 2009.

[38] OMG. Business Process Model and Notation (BPMN) - Version 2.0. `http://www.omg.org/spec/BPMN/2.0/PDF`, January 2011.

[39] J. Pinggera, S. Zugal, and B. Weber. Investigating the Process of Process Modeling with Cheetah Experimental Platform. In *Proc. ER-POIS '10*, pages 13–18, 2010.

[40] Helen C Purchase. Which aesthetic has the greatest effect on human understanding? In *Graph Drawing*, pages 248–261. Springer, 1997.

[41] Helen C Purchase. Metrics for Graph Drawing Aesthetics. *Journal of Visual Languages & Computing*, 13(5):501–516, 2002.

[42] Helen C Purchase, Jo-Anne Allder, and David Carrington. User Preference of Graph Layout Aesthetics: A UML Study. In *Graph Drawing*, pages 5–18. Springer, 2000.

[43] Nornadiah Mohd Razali and Yap Bee Wah. Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics*, 2(1):21–33, 2011.

[44] Manfred Reichert. Visualizing Large Business Process Models: Challenges, Techniques, Applications. In *Business Process Management Workshops*, pages 725–736. Springer, 2012.

[45] Manfred Reichert and Barbara Weber. *Enabling Flexibility In Process-Aware Information Systems: Challenges, Methods, Technologies*. Springer Science & Business Media, 2012.

[46] Hajo Reijers, Jan Mendling, et al. A Study Into the Factors That Influence the Understandability of Business Process Models. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 41(3):449–462, 2011.

[47] SB Rinderle, Ralph Bobrik, Manfred Reichert, and Thomas Bauer. Business Process Visualization-Use Cases, Challenges, Solutions. 2006.

[48] Marcello La Rosa, Arthur HM Ter Hofstede, Petia Wohed, Hajo A Reijers, Jan Mendling, and Wil MP Van der Aalst. Managing Process Model Complexity via Concrete Syntax Modifications. *Industrial Informatics, IEEE Transactions on*, 7(2):255–265, 2011.

[49] M. Schier. Adoption of Decision Deferring Techniques in Plan-driven Software Projects. Master's thesis, University of Innsbruck, Institute of Computer Science, 2008.

[50] Matthias Schrepfer, Johannes Wolf, Jan Mendling, and Hajo A Reijers. The Impact of Secondary Notation on Process Model Understanding. In *The Practice of Enterprise Modeling*, pages 161–175. Springer, 2009.

[51] David J Sheskin. *Handbook of Parametric and Nonparametric Statistical Procedures*. crc Press, 2003.

[52] WMC Specification. Workflow Management Coalition Terminology & Glossary (Document No. WFMC-TC-1011). *Workflow Management Coalition Specification*, 1999.

[53] Alexander Streit, Binh Pham, and Ross Brown. Visualization Support for Managing Large Business Process Specifications. In *Business Process Management*, pages 205–219. Springer, 2005.

[54] Wil MP van Der Aalst, Arthur HM Ter Hofstede, Bartek Kiepuszewski, and Alistair P Barros. Workflow Patterns. *Distributed and parallel databases*, 14(1):5–51, 2003.

[55] Wil MP Van Der Aalst, Arthur HM Ter Hofstede, and Mathias Weske. Business Process Management: A Survey. In *Business process management*, pages 1–12. Springer, 2003.

[56] Irene Vanderfeesten, Jorge Cardoso, Jan Mendling, Hajo A Reijers, and WMP Van der Aalst. Quality Metrics for Business Process Models. *BPM and Workflow handbook*, 144:179–190, 2007.

[57] Irene Vanderfeesten, Hajo A Reijers, Jan Mendling, Wil MP van der Aalst, and Jorge Cardoso. On a Quest for Good Process Models: the Cross-Connectivity Metric. In *Advanced Information Systems Engineering*, pages 480–494. Springer, 2008.

[58] Barbara Weber, Manuel Neurauter, Jakob Pinggera, Stefan Zugal, Marco Furtner, Markus Martini, and Pierre Sachse. Measuring Cognitive Load During Process Model Creation. In *Information Systems and Neuroscience*, pages 129–136. Springer, 2015.

[59] Matthias Weidlich, Stefan Zugal, Jakob Pinggera, Dirk Fahland, Barbara Weber, Hajo A Reijers, and Jan Mendling. The Impact of Sequential And Circumstantial Changes On Process Models. In *ER-POIS workshop at CAISE, Springer*, pages 43–54, 2010.

[60] Mathias Weske. *Business Process Management: Concepts, Languages, Architectures.* Springer Science & Business Media, 2012.

[61] Stephen A White. Process Modeling Notations and Workflow Patterns. *Workflow handbook*, 2004:265–294, 2004.

[62] David F Williamson, Robert A Parker, and Juliette S Kendrick. The Box Plot: a Simple Visual Method to Interpret Data. *Annals of internal medicine*, 110(11):916–921, 1989.

[63] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. *Experimentation in Software Engineering.* Springer Science & Business Media, 2012.

[64] S. Zugal. *Applying Cognitive Psychology for Improving the Creation, Understanding and Maintenance of Business Process Models.* PhD thesis, University of Innsbruck, Department of Computer Science, 2013.

[65] Stefan Zugal, Jakob Pinggera, and Barbara Weber. Assessing Process Models with Cognitive Psychology. In *EMISA*, volume 190, pages 177–182, 2011.