# MASTER'S THESIS

## A study into the impact of the use of modularity in business process modeling on understandability

Milder, A.

**Award date:**
2021

# A study into the impact of the use of modularity in business process modeling on understandability

| | |
|---|---|
| Opleiding: | Open Universiteit, faculteit Betawetenschappen<br>Masteropleiding Business Process Management & IT |
| Degree programme: | Open University of the Netherlands, Faculty Science<br><br>Master of Science Business Process Management & IT |
| Course: | IM0602 BPMIT Graduation Assignment Preparation<br>IM9806 Business Process Management and IT Graduation Assignment |
| Student: | Arnaud Milder |
| Identification number: | |
| Date: | 03-07-2021 |
| Thesis supervisor | Dr.ir. I. Vanderfeesten |
| Second reader | Dr. B. Roelens |
| Version number: | 1.0 |
| Status: | Final version |

# Abstract

Process models are used for communication about processes to various stakeholders. An important group of stakeholders are the employees involved in these processes, as the models present an overview of their activities. Since not all stakeholders are experts in the field of process modeling, the models should not only accurately represent reality, but also be easy to understand.

Especially if models are large, they become more difficult to understand. That is the reason why modularization is applied in large and/or complex processes. Several modularization strategies have been developed to increase the understandability of process models. One of those techniques is vertical modularization, which requires a user to navigate through the main process and sub-processes in order to read the process model.

This study has focused on the vertical modularization technique, using two different strategies, the so-called 'overview+detail' and 'focus+context' strategies. The purpose of the study was to determine if there would be a significant difference, with respect to understandability, between those two interface strategies. This significant difference was not established.

Follow-up studies could include other modularization techniques to see whether a significant difference in understandability of models can yet be established.



# Key terms

Business Process Modeling, Modularity, Sub-processes, Understandability, Vertical modularization, Interface strategy

# Summary

With process modeling, business processes are mapped. Process models are not only used to analyze the processes, they are also used to communicate about the processes to different stakeholders. An important group of stakeholders are the employees involved in these processes, since the models present an overview of their activities which are part of a business process. This imposes requirements on the understandability of the models for laypersons in the field of process modeling. One way to make models more understandable is to use modularization. Using modularization, a process model is composed of smaller subsystems. Especially in the case of large and complex models, modularization can help to increase their understandability.

Several techniques for applying modularization in models have been found in the existing scientific literature. This research contributes to the scientific literature by examining whether the use of one of these techniques, vertical modularization, provides a difference in understandability for the use of business process models by non-experts in the field of process modeling.

When vertical modularization is used in a model, a user has to navigate through the main process and sub-processes in order to read the process model. To be able to navigate, several interface strategies can be used to vertically modularize a process model. This research has examined and compared two interface strategies, ‘overview+detail’ and ‘focus+context’. The purpose of the research was to determine whether there is a difference in understandability between models in which the 'overview+detail' or 'focus+context' strategy is used.

To detect a possible difference between the use of the two interface strategies, an online experiment was designed in which participants were asked to answer questions about a fictitious process model. Participants were given a brief explanation of the design of the experiment and the model used. They were then asked some questions to assess their knowledge and experience with process modeling, after which the actual questions about the models followed.

Participants were randomly divided into two groups. Both groups were presented with the same model and the same sets of questions. Two sets of questions were used, one with models using the ‘overview+detail’ strategy and the other set using the ‘focus+context’ strategy. The question sets were presented to both groups in reverse order to mitigate the learning effect. Two dependent variables were defined: ‘task effectiveness’ and ‘task efficiency’. ‘Task effectiveness’ was operationalized by the number of correct answers. ‘Task efficiency’ was calculated by dividing the number of correct answers by the time it took a participant to answer these questions correctly.

Based on the literature research, the hypothesis of this study was formulated as:

**The use of ‘overview+detail’ interface strategy will result in better understanding of the process model when used by a layperson than the use of ‘focus+context’ interface strategy.**

The results of the experiment resulted in rejection of this hypothesis.

In addition to the independent variable 'interface strategy’, three confounding variables were examined: ‘theoretical modeling competency’, ‘practical modeling competency’ and ‘education’. Those variables were operationalized by asking the participants to self-asses their knowledge of process modeling and use of process models in daily practice. They were also asked to provide their highest level of education. Taking these confounding variables into account, no significant difference was found in the results for both interface strategies.

Finally, the difference between global and local questions was examined. When answering local questions, a participant only needs knowledge of a sub-process. In contrast, answering a global question requires not only knowledge of the sub-process but also of the big picture. When this distribution of questions was taken into account, it was found that participants had better ‘task

efficiency' when answering the local questions combined with the 'overview+detail' strategy than with the 'focus+context' strategy. More research is needed to statistically support this finding.

The purpose of this study was to determine whether, in terms of understandability, there would be a significant difference between the vertical modularization strategy using the 'overview+detail' strategy and the 'focus+context' strategy. This significant difference was not established.

The study concludes with some recommendations for practice and possible follow-up studies.

## Contents

## List of Figures

# List of Tables

# 1. Introduction

## 1.1. Background

Business process models have become indispensable in organizations. Not only are they used to analyze the processes, they are also used for communication about the processes to various stakeholders. An important group of stakeholders are the employees involved in these processes, since the models provide an overview of their activities which are part of a business process. They show when and by whom a process step should be performed. Therefore, it is necessary that the models accurately reflect reality, but are also easy to understand for all parties involved. Model comprehension is a primary quality factor of a process model (Figl, 2017). After all, if a model is not understood, the model is of no use.

## 1.2. Exploration of the topic

In this research, understandability is defined as the degree to which information contained in a process model can be easily understood by a reader of that model (Reijers & Mendling, 2011).

The understandability of a model depends on many factors. Dikici et al. (2018) divides these factors into process model factors and personal factors, as shown in Figure 1.

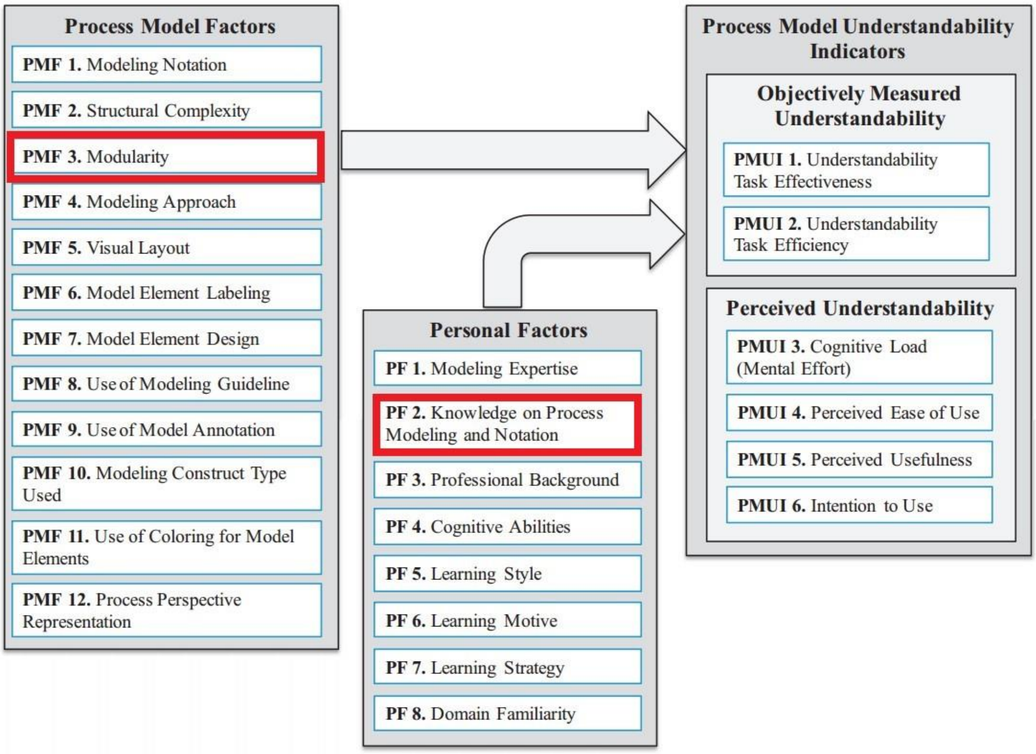


*Figure 1 - An integrated framework of process model understandability (Dikici et al., 2018)*

Not only does the way a process model is visualized play a role in the understandability of the model, the characteristics of the person reading the model are also relevant. The combination of all factors from the framework, as shown in Figure 1, determine the degree of understanding. This means, among other things, that when a model is very clear for a certain target group, it does not necessarily have to be comprehensible for other target groups as well.

In the case of large and/or complex process models, applying modularization (PMF 3) is considered an important factor for the understandability (Turetken et al., 2019). Modularity is defined by Reijers and Mendling (2008) as the design principle of having a complex system composed of smaller subsystems that can be managed independently yet function together as a whole. At the highest level one can show the general structure of the model, while at a lower level more detailed information can be shown.

Regarding personal factors, it is important to bear in mind that process models within organizations are often used to communicate to persons who often lack in-depth theoretical knowledge on process modeling notation (PF2). However, as mentioned before, it is paramount that they understand the model. Therefore, it is relevant to examine the influence of the use of modularization together with a person's knowledge of process modeling on the understandability of a model.

## 1.3. Problem statement

A process model can easily become very large and complex. For example, there can be many exceptions to the main process. Modularization of these large and complex models can be used to increase their understandability (Turetken et al., 2019). There are several ways in which modularization of a model can be designed.

The way in which modularization contributes to the understandability may also depend on the target group. The result of modularization on the understandability could be different for people who are experts in business process modeling versus people who are laypersons in that field.

Therefore, the problem statement is:

**Little research has been done on which modularization strategy is better for laypersons' understanding of process models.**

In this study, laypersons are considered to be people who are no experts in the field of process models and/or process modeling in general, but also do not have specific domain knowledge of the process being modelled.

## 1.4. Research objective and questions

The aim of this study is to find out which modularization strategy results in a better understanding of a business process model by laypersons.

The main research question is:

**Which modularization strategy is best in terms of understandability of a process model for laypersons?**

In order to answer the main question, the following sub-questions have been defined:

1. **Which process model modularization strategies are described in the literature?**
2. **What is the impact of those process model modularization strategies on the understandability of the process model?**
3. **Which of those strategies are selected as most interesting for further research in practice?**
4. **Which of the selected modularization strategies actually results in the highest understandability of the process by laypersons?**

## 1.5. Motivation and relevance

How and when processes should be modularized is often a topic of discussion. The goal of this research is to contribute to the body of knowledge on this subject. The study aims to provide an overview of all modularization strategies described in the literature and their impact on the understandability of process models.

After the literature review, a well-considered selection of the most promising modularization strategies will be further investigated. This research will show which notation leads to better understanding of the process model. Because process models are increasingly used to clarify work processes to people who are not adept at creating and using models, it is very important that those process models can be understood by laypeople without too much effort.

The outcome of this research will increase knowledge about the relationship between modularization and understandability and initiate guidelines on how modularization should be put into practice in the future. With these guidelines, one can create better models that can be easily understood by laypersons.

## 1.6. Main lines of approach

In order to answer the main research question, an empirical research will be conducted. The following sections are structured as follows. Section 2 presents the theoretical framework. It provides a list of modularization strategies which are described in the literature. This section also indicates which strategies are most promising regarding the increase of understandability. Section 3 gives a justification of the empirical research that has been carried out. Section 4 describes the implementation and realization of the experiment and presents the outcomes of the experiment. The results are discussed in section 5, followed by conclusions and suggestions for further research.

# 2. Theoretical framework

To answer the first three sub-questions as stated in section 1.4, a literature review has been carried out (Saunders et al., 2019). This section provides the theoretical framework of the research conducted.

## 2.1. Literature review approach

To find and select the literature for the review, the defined stages of the PRISMA 2009 Flow Diagram, were used. Appendix A describes in more detail the way in which the search was planned and conducted i.e. the search criteria, search queries and selection criteria. Appendix A also shows the results of the search in each phase. After the search and selection of articles was completed, six articles remained, see Appendix A Table 26.

## 2.2. Results and conclusions

### 2.2.1. Modularization strategies

In this section the first sub-question will be answered:

**Which process model modularization strategies are described in the literature?**

Modularizing a model means splitting it up into smaller models, in other words sub-processes. These smaller parts should be intrinsically complete and able to be managed independently (Winter et al., 2020). In the literature, terms like modularity, decomposability and hierarchy are often used interchangeably. There is, however, a difference between them (H. A. Reijers et al., 2011, p. 882). For example, all three models depicted in Figure 2 are modular models. Model a is clearly a hierarchical model because all processes come from one process at the top. Model b, on the other hand, is not because there exists a cyclic dependency between a subset of the modules. Model c, in turn, is hierarchical but hardly decomposable given the amount of dependencies between the modules.

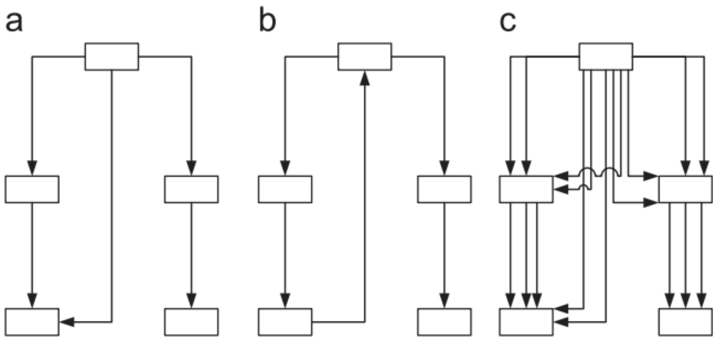


*Figure 2 - Examples of modular designs (H. A. Reijers et al., 2011)*

Therefore, when talking about modularization, it is paramount to clearly indicate what is meant by it. Winter et al. (2020) define three types of modularization: horizontal, vertical, orthogonal. Horizontal modularization, see Figure 3, divides the model into several smaller models with the same level of abstraction.

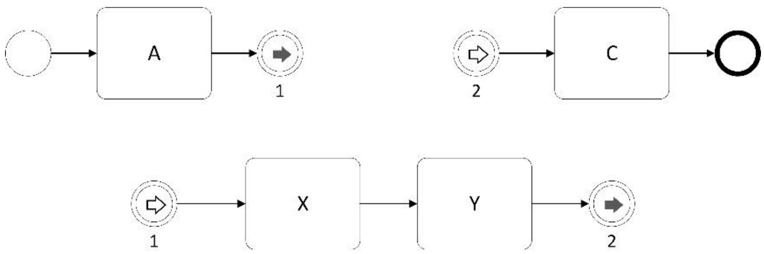


*Figure 3 - Horizontal modularization (Winter et al., 2020)*

Vertical modularization, also called hierarchical structuring, consists of a process model that is broken down into smaller sub-processes at a deeper level. While the main process has a high level of abstraction, the underlying sub-processes are more detailed, see Figure 4.

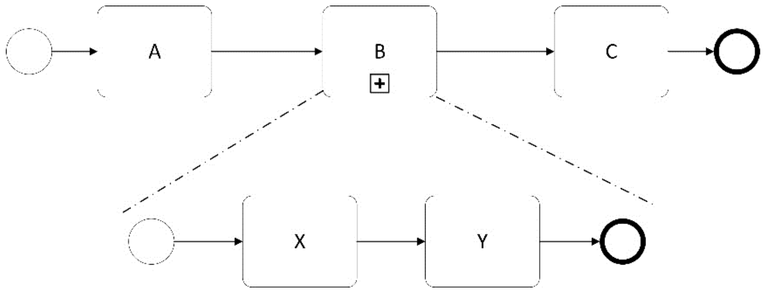


*Figure 4 - Vertical modularization (Winter et al., 2020)*

The orthogonal modularization is based on exceptions and further cross-cutting concerns, see Figure 5 .

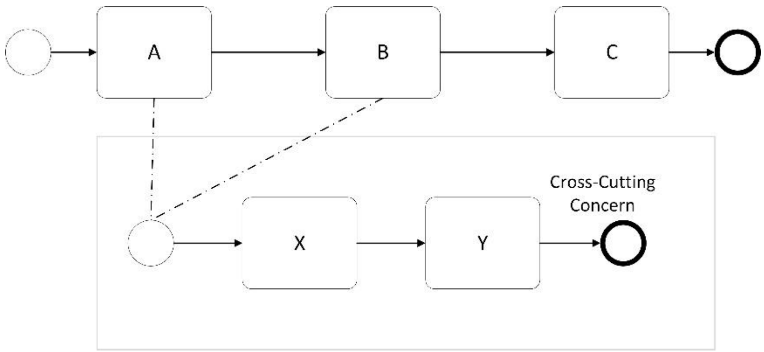


*Figure 5 - Orthogonal modularization (Winter et al., 2020)*

In addition to those three , Turetken et al. (2019) provide one extra modularization strategy, called the "Flattened view with groups", see Figure 6. These groups cluster a set of logically related model elements.

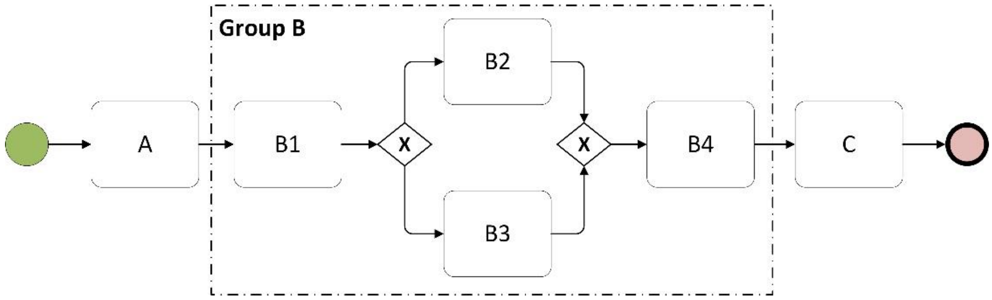


*Figure 6 - Flattened with groups modularization (Turetken et al., 2019)*

In the case of the vertical modularization strategy, see Figure 4, a user has to navigate through the main process and the sub-processes in order to read the process model (Figl et al., 2013). Cockburn et al. (2009) define four possible interface strategies with which a user can navigate through a vertical model. Examples of the strategies are described in Appendix C. Two of those strategies, called 'overview+detail' and 'focus+context', are commonly used in process modeling both on paper and on screen. Figure 7 shows examples of both strategies.

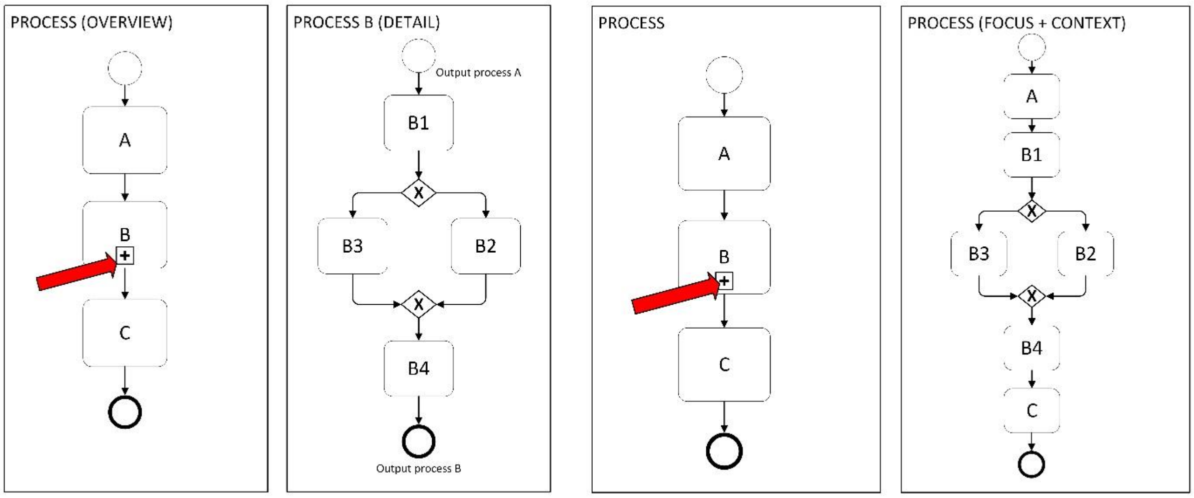


*Figure 7 - Overview+detail strategy (left) and Focus+context strategy (right)*

**So, in the literature four main modularization strategies have been found and two associated interface strategies.**

### 2.2.2. Impact of modularity on the understandability

This section takes a look at sub-question 2:

**What is the impact of the process model modularization strategies, found in section 2.2.1, on the understandability of the process model?**

Should modularization be applied, and if so, when? The literature does not yet provide an unambiguous answer to that question (H. A. Reijers et al., 2011, p. 883). The reason for this, is the fact that modularization creates multiple pieces of information from multiple (sub-)models or components that have to be processed and integrated simultaneously by the user. This causes an increase in cognitive load and can result in, what the literature calls, the split-attention effect (Figl et al., 2013).

Some research has already been done on the impact of modularization on the understandability of process models. Reijers and Mendling (2008), for example, state that modularization through the use of sub-processes appears to have a positive impact. However, this occurs only in large models and particularly in cases where insight into a local part of the model is required. In those cases, modularity shields a user from unnecessary information.

Another study by Turetken et al. (2019) compares the results of models (fully-flattened and modularized with 'groups') with models using vertical modularization in which sub-processes are shown in separate models ('overview+detail'). They conclude that flattened models are easier to understand than models with sub-processes shown separately. However, they also indicate that if vertical modularization is necessary and the comprehension of the models is critical, the use of groups to indicate the process elements that can be combined into a sub-process, see Figure 6, should be preferred to separately shown sub-processes.

In an earlier study by Figl et al. (2013), experts were asked by means of a questionnaire to evaluate the two interface strategies for vertical modularization: ‘overview+detail’ and ‘focus+context’. The study made the observation that the experts preferred the ‘overview+detail’ strategy, as described above, over the ‘focus+context’ strategy.

There are many factors that play a role in the outcome of these studies into understandability. One potential shortcoming of these studies is the fact that participants were expected to have knowledge of process modeling. Other limitations indicated include the size of the process model that was used in the study. In addition, what questions are asked is of great importance. There is a big difference between the situation if only knowledge of a sub-process (local) is needed when answering a question and when insight into the big picture is needed to answer a question. Not to be underestimated is the influence that the chosen notation of the model (e.g. BPMN or UML) can have. The use of medium on which the process models are presented, paper or monitor, has a considerable effect on the understandability as Turetken et al. (2019, p. 139) concluded in their study. Practical issues such as which monitor (size and resolution) a participant viewed the model on also play a role.

**After reviewing the literature and the studies already conducted, it can be concluded that not enough research has been done to be able to fully and unambiguously answer sub-question 2.**

For this reason, the following section will examine what research is interesting in order to provide more clarity on this subject and to be able to give a clearer answer to the sub-question.

### 2.2.3. Further research on modularization strategies

**Which of those strategies are selected as most interesting for further research in practice?**

As shown in section 2.2.2, more research needs to be done to indicate which modularization strategies have a positive impact on the understandability of a process model. Because not all modularization strategies and associated interface strategies can be investigated in this study, a choice must be made as to which strategy is the most interesting to investigate.

Nowadays, process models are mostly viewed digitally on computer screens, often printing the models is not allowed due to version control. This research will therefore focus on showing a model on a computer screen. Compared to the other strategies, the vertical modularization strategy is well suited to fit a large model on a computer screen and keep it readable without having to scroll.

When presented with large and complex process models, a user may experience “map shock” (Moody, 2006) , in other words “the feeling of being overwhelmed by the scale and complexity of the display” (Blankenship & Dansereau, 2000). This may result in someone not using the model at all. Vertical modularization reduces the size and complexity of models at first sight and may therefore reduce the risk of “map shock”.

Most modeling techniques support modularization by decomposing processes into sub-processes (H. A. Reijers et al., 2011). When using vertical modularization, attention must be paid to how a user navigates to and from the sub-processes. This means that the interface strategy must also be examined.

**For this reason, the vertical modularization strategy with the interface strategies ‘overview+detail’ and ‘focus+context’, as described in section 2.2.2, are selected as most interesting for follow-up research.**

## 2.3. Objective of the follow-up research

Building on the research already conducted, this research will further contribute to answering the main research question and sub-questions. The aim of the follow-up research is to check whether the conclusions drawn in previous studies regarding vertical modularization and the interface strategies, also apply when participants have neither expert knowledge of process modeling nor domain knowledge of the modelled process. An example of this is the observation made by Figl et al. (2013) that experts prefer the ‘overview+detail’ strategy over the ‘focus+context’ strategy in a reading task.

In this way, the study contributes to the body of knowledge on how vertical modularization affects the understandability of process models.

# 3. Methodology

This chapter describes the design and tactics of the research strategy. The research strategy should facilitate the investigation into the impact of the interface strategy on the understandability of a process model that uses vertical modularization. More specifically, when such a model is used by a layperson.

## 3.1. Conceptual design: select the research method(s)

The main purpose of the research is to find the answer to sub-question 4:

**Which of the selected modularization strategies actually results in the highest understandability of the process by laypersons?**

As seen in section 2.2.3, the research will be limited to vertical modularization and the interface strategies: ‘overview+detail’ and ‘focus+context’.

This study has an explanatory purpose because it seeks to establish a causal relationship between the interface strategy of vertical modularization and the understandability (Saunders et al., 2019, p. 188) in practice. To demonstrate this relationship, data are collected and analyzed in an experiment. A hypothesis is formulated and the compatibility of the data, collected in the experiment, with this hypothesis is tested statistically. This is a main characteristic of a quantitative research design (Saunders et al., 2019, p. 178).

In the experiment, participants will be asked to answer questions about a process model of a fictitious process. To be able to answer the questions, participants must actually use the model they are presented with. The model will be presented with different interface strategies (independent variable). This difference in strategies may affect the understandability (dependent variable) of the process model (Saunders et al., 2019, p. 190).

From the above, it can be derived that the study will use a mixed method research design (Saunders et al., 2019, p. 181), because participants will be asked to complete a short questionnaire (qualitative) and to take part in an experiment (quantitative).

The participant will be divided into two groups. Both groups will be presented with a process model of the same process. The first group of participants is firstly presented with a set of questions using the ‘overview+detail’ interface strategy and then a comparable set of questions using the ‘focus+context’ interface strategy. The second group is presented with the same sets of questions, but in reverse order, see Table 1. By reversing the sets of questions, the learning effect can be removed from the results of the experiment.

*Table 1 - Order of Presentation of questions per participant group*

| | **First question set** | **Second question set** |
|---|---|---|
| Participant Group 1 | Model using ‘overview+detail’ | Model using ‘focus+context’ |
| Participant Group 2 | Model using ‘focus+context’ | Model using ‘overview+detail’ |

This way, it is possible to aggregate the results of both groups and compare the results within this new group to establish a possible difference between the two interface strategies with regard to the understandability of the models, see Appendix D for more details.

The dependent variable ‘understandability’ will be measured objectively using ‘task effectiveness’ and ‘task efficiency’ (Figl et al., 2013). Understandability can also be measured while using subjective variables ‘perceived usability’ and ‘perceived ease’ (Dikici et al., 2018). Those subjective variables are not included in this study. Personal factors, 'theoretical modeling competency', 'practical modeling competency' and 'education' are added as confounding variables. To operationalize these factors, the participants are asked to self-assess their knowledge of business process models, their experience with business process models and are asked about their highest education. By using a fictitious process, knowledge of the participants about the process does not play a role.

The resulting research model is displayed in Figure 8 (Turetken et al., 2019):

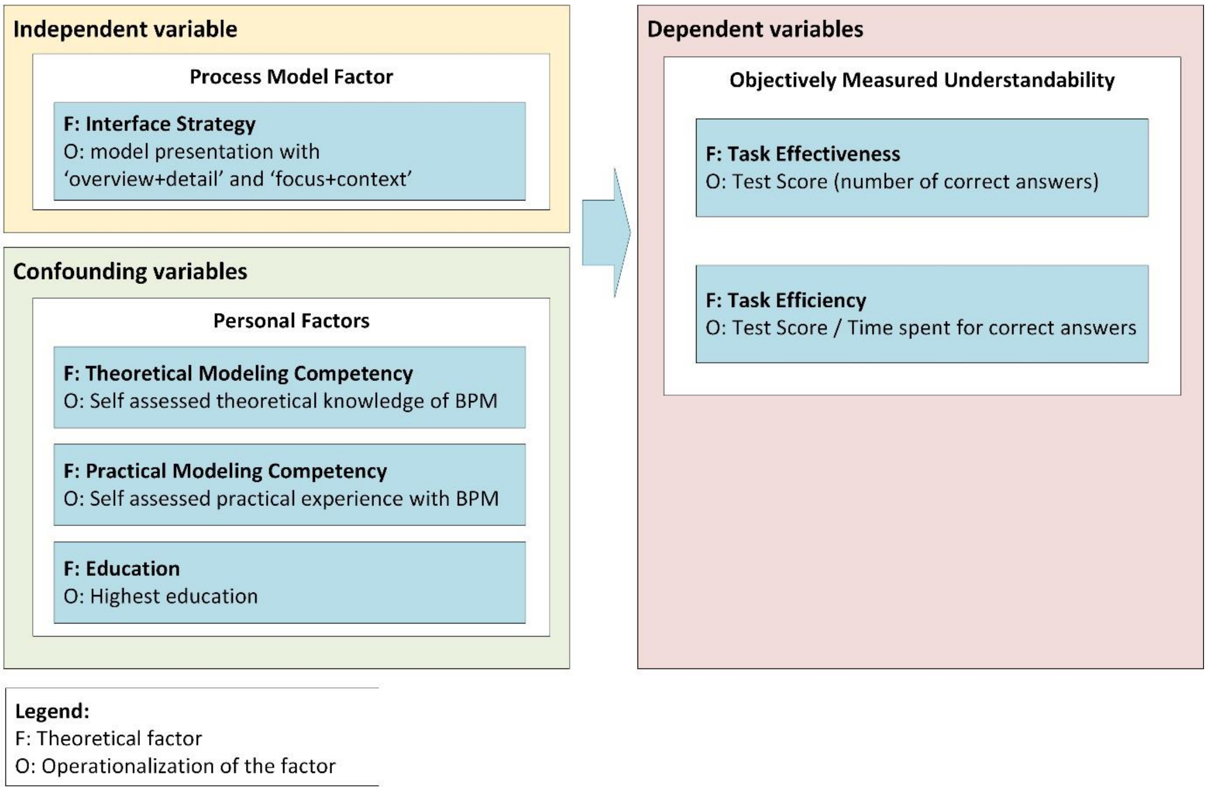


*Figure 8 - Research Model based on Turetken et al. (2019)*

Task effectiveness is an indication of whether a model leads to correct answers and will be measured by counting how many questions a participant answers correctly. 'task efficiency' is an indication whether a model is quickly understood and will be calculated by dividing the number of correct answers by the time spent on those correct answers.

Based on the research model and the conclusions from Figl et al. (2013), the following hypothesis is formulated:

**Hypothesis: the use of 'overview+detail' interface strategy will result in better understanding of the process model when used by a layperson than the use of 'focus+context' interface strategy.**

## 3.2. Technical design: elaboration of the method

As described in section 3.1, a short questionnaire will be mixed with an experiment to compare two interface strategies 'overview+detail' and 'focus+context'.

### 3.2.1. Questionnaire

Firstly, participants will be asked if they have certain knowledge of and/or experience with process models. They will also be asked to provide their educational background. The questions and possible answers, used to determine the confounding factors, were taken from Reijers and Mendling (2011), see Appendix D for details. It was decided not to take a standard Business Process Modeling

Competency test to establish the participants' level of theoretical knowledge as was done by Mendling et al. (2012) and Turetken et al. (2019). This would potentially discourage people who, in their own opinion, have insufficient knowledge of process modeling from participating. In this experiment, it was desirable for everyone, especially those with little knowledge, to participate.

### 3.2.2. Experiment

Participants will be randomly assigned to two groups. Each group will be presented with a process using interface strategy: 'overview+detail' and 'focus+context' in different order, see section 3.1. This means the experiment has a within-subjects approach, because every participant is exposed to an intervention (Saunders et al., 2019, p. 193). For both participant groups, the intervention consists of presenting the models of the first set of questions with a different interface strategy than the models of the second set of questions. Using two groups may negate a possible learning effect. Also, there is no need to perform a baseline measurement because the two interface strategies are compared with each other. The open source tool LimeSurvey is used to carry out the experiment.

The questions in the experiment have to be drawn up carefully. Some questions require an overview of the whole process (global). Others require only information about a specific part of the process represented in a sub-process (local). Other factors are also taken into account in drafting the questions. The exact way in which the questions are prepared is included in Appendix D. To answer these questions, a multiple choice questionnaire is used. One of the choices will always be 'I don't know'.

The process model is based on Model L from Reijers and Mendling (2011). The choice of the number of activities and nodes is explained in Appendix D, as well as the model itself.

The variable 'task effectiveness' will be measured by counting how many questions a participant answered correctly. 'Task efficiency' will be calculated by dividing a participants' number of correct answers by the time it took him or her to answer the questions correctly. Results will used to test the hypothesis.

### 3.2.3. Participants

Since this experiment focuses on laypersons, the indicated experience and knowledge is used to consider whether the results of certain participants should be removed because they cannot be counted as laypersons.

Based on the central limit theorem, an absolute minimum of 30 participants must participate for each interface strategy. In the experiment, each participant is presented with both strategies, see section 3.1. This way, a total of 30 participants have to participate in the experiment. However, because the participants are divided into two groups, the goal is to have a minimum of 20 participants per group, which means a total of 40 participants.

Participation in the experiment will be voluntarily and anonymously. Furthermore, the questionnaire and experiment will be conducted online, so participants need to have access to a device with an internet connection.

## 3.3. Data analysis

As mentioned in section 3.1, 'task efficiency' is calculated by dividing the number of correct answers by the time spent on those correct answers. In order to perform this calculation it is necessary to use a tool (LimeSurvey) which measures the time a participant spent answering the question.

Hypothesis testing will be used to tell whether the difference between the two vertically modularized models that use a different interface strategy is statistically significant (Saunders et al., 2019, p. 603). The data collected must be independent and normally distributed. Statistics will be used to calculate the p-value. If the data should turn out to be not normally distributed, then nonparametric tests will be used (Field, 2017, p. 283).

Data will be checked for outliers, e.g. did a participant answer all questions with 'I don't know' and/or were all answers provided in an unreliable short or long period of time.

The data will also be analyzed with regard to the answers provided in the questionnaire. This means that the results for 'task effectiveness' and 'task efficiency' will also be compared using the three confounding variables. These data can be used to see whether the education, experience and knowledge of a participant affect the results between the two models. However, the estimation of the amount of experience and theoretical knowledge is done by the participants themselves and is therefore subjective. Therefore, any conclusions drawn from it should be considered with the necessary care.

## 3.4. Reflection on research design

Reliability and validity are often used to judge the quality of a research (Saunders et al., 2019, p. 213). This section will reflect upon these aspects. Furthermore, the ethics of the research design will be established.

### 3.4.1. Reliability

Reliability of the study, as defined by Saunders et al. (2019), is achieved with the design of the study as provided in the previous sections. The description should allow other researchers to be able to repeat the research and come to the same conclusion.

To avoid errors in the experiment, a pre-test test will be conducted. A minimum of two persons will be asked to review the questionnaire and the models. Any errors in texts and models revealed by this test will be corrected.

To avoid participant bias, a quiet place where one can participate in the experiment undisturbed is recommended prior to the start of the experiment.

However, attention should be paid to the risk related to the experience of participants. The estimation of this experience is left to the participants and is therefore subjective. Another potential risk is the fact that participants are all recruited from the same social environment. This could mean that no proper reflection of the population has been established.

### 3.4.2. Validity

The research has been designed to make it possible to demonstrate that the difference in findings can be attributed to the two different interface strategies and not to other confounding variables. This is called internal validity (Saunders et al., 2019, p. 215). As described in section 3.3, the data will be checked for errors caused by participants who did not answer the questions seriously. By keeping the experiment from being too complicated, participants are prevented from dropping out during the experiment. It will not be possible for participants to save the intermediate results. This prevents participants from increasing their knowledge while conducting the experiment.

External validity is achieved by recruiting participants with different working backgrounds. Finally, construct validity has been taken into account as the questions and model are based on the earlier study by Reijers and Mendling (2011).

### 3.4.3. Ethical aspects

Participants in the study participate voluntarily, this should prevent subjecting them to any risk of embarrassment or any other disadvantage. They will also be asked to give their consent after being clearly explained how the experiment will take place. A participant can decide to stop with the experiment at any time. The results will be processed anonymously.

# 4. Results

This chapter describes the execution of the research strategy outlined in chapter 3. The general descriptive statistics will be presented, as well as the cleaning of the data, including looking for possible outliers. The hypothesis, as described in section 3.1, will be tested. In addition, the effect of confounding variables will be examined.

## 4.1. Data gathering

The pre-test, carried out by four fellow students, provided meaningful feedback. After processing their comments in LimeSurvey, the survey was published online from March 10 to March 26. A total of 56 people were invited by email to participate in the study. On March 17, 24 invitees, who had not yet completed the survey, received a reminder to participate. In the end, a total of 45 people completed the survey.

After closing the survey, the results were exported from LimeSurvey to an Excel spreadsheet. In Excel, the data were edited to be able to import them into SPPS. For example, the values ‘Yes’, ‘No’ or ‘I don’t know’ were converted to ‘Correct’, ‘Incorrect’ or ‘I don’t know’. In further analysis ‘I don’t know’ is considered an incorrect answer.

### 4.1.1. Data cleansing

Based on the answers given by the participants in the questionnaire, everyone was considered a layperson, see Appendix F for details. Therefore, this did not give rise to the removal of results from certain participants.

The data were also checked for outliers on the time it took the participants to complete the survey as a whole. One participant (ID = 44) took a very long time to finish the survey, over 48 minutes. Since this participant took a long time on all the questions, this was not considered an outlier. In contrast, another participant (ID = 1) took a very short time, just under six minutes. However, because he or she had answered all questions correctly, this participant evidently answered the questions seriously and was not labelled an outlier. In the end, no participants were removed based on the time it took them to complete the entire survey.

Subsequently, a check was carried out to see whether certain participants spent a lot of time on specific questions. This could indicate that they were distracted while answering the question and distort the results of the study. Figure 9 and Figure 10 show the boxplots for both the time spent on questions using the ‘overview+detail’ strategy (Figure 9) and on questions using the ‘focus+context’ strategy (Figure 10).

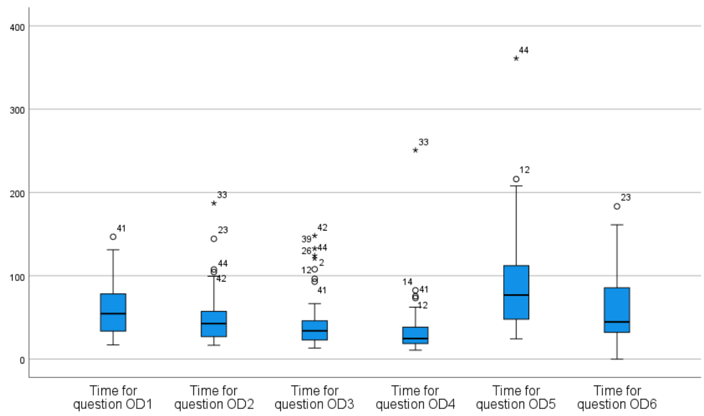


*Figure 9 - Boxplot time used per question ('overview+detail')*

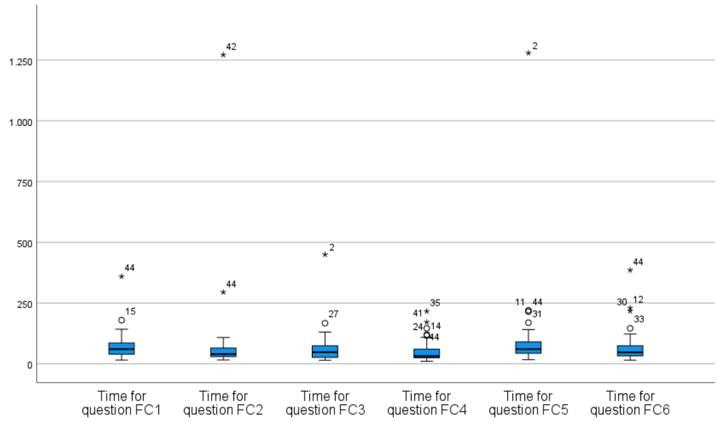


*Figure 10 - Boxplot time used per question ('focus+context')*

Z-scores were calculated for the time a participant spent on answering a question. Z-scores smaller than -3,29 and bigger than 3,29 were considered potential outliers (Field, 2017, p. 39). These are shown in Table 2. Then, for each potential outlier, participants' response pattern was examined. As discussed before, participant with ID = 44 had several potential outliers, but this fitted his/her pattern of answering the questions.

*Table 2 - Outlier analysis for time used per question*

| Question | Participant ID | Time for question (seconds) | Z-score | Outlier Yes/No | Answer correct/incorrect |
|---|---|---|---|---|---|
| **OD2** | **33** | **187,10** | **3,94469** | **Yes** | **Correct** |
| **OD4** | **33** | **250,81** | **5,83759** | **Yes** | **Incorrect** |
| OD5 | 44 | 361,01 | 4,02977 | No, fits the pattern | Correct |
| FC1 | 44 | 359,57 | 5,11756 | No, fits the pattern | Incorrect |
| **FC2** | **42** | **1.271,60** | **6,37872** | **Yes** | **Correct** |
| **FC3** | **2** | **449,83** | **5,65433** | **Yes** | **Incorrect** |
| **FC4** | **35** | **216,30** | **3,76598** | **Yes** | **Correct** |
| **FC5** | **2** | **1.278,98** | **6,36981** | **Yes** | **Correct** |
| FC6 | 44 | 385,73 | 4,66604 | No, fits the pattern | Correct |

Ultimately, six measurement were considered outliers, see bold figures in Table 2. These six outliers were related to four participants (ID = 2, 33, 35 and 42). The complete results of all four participants were removed from the dataset, to prevent distortion of the overall results. Leaving the results of 41 participants.

As described in section 3.1, the dependent variable 'task effectiveness' was operationalized by the number of correct answers. The second dependent variable ' was calculated by dividing the number of correct answers by the time spent on those correct answers. This value was subsequently converted into the number of correct answers per hour, as was done by Turetken et al. (2019).

### 4.1.2. Descriptive statistics

The participants were randomly divided into two groups by LimeSurvey. After data cleansing, twenty participants were assigned to group 1, see Table 3. They were first given the questions using the 'overview+detail' interface strategy and then the questions using the 'focus+context' strategy. Twenty-one participants were assigned to group 2. They were presented with the questions in reverse order.

*Table 3 - Participants per participant group*

| | Frequency | Percentage |
|---|---|---|
| Participant group 1 | 20 | 48,8 |
| Participant group 2 | 21 | 51,2 |
| Total | 41 | 100,0 |

Table 4 shows the number of correct, wrong and 'I don't know' answers per question for all participants.

*Table 4 – Overview of answers per question*

| | 'overview+detail' | | | | | | 'focus+context' | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Question | 1 | 2 | 3 | 4 | 5 | 6 | 1 | 2 | 3 | 4 | 5 | 6 |
| Correct | 39 | 38 | 38 | 32 | 25 | 35 | 34 | 38 | 39 | 36 | 27 | 36 |
| Incorrect | 2 | 3 | 3 | 8 | 16 | 5 | 7 | 3 | 2 | 4 | 14 | 5 |
| I don't know | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| Total | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 |

From Table 4, it can be seen that participants had the most difficulty with question 5 of both sets.

A comparison of the results of both participant groups revealed that a slight learning effect occurred when looking at the 'task effectiveness'. The Independent-Samples median test showed that participant group 2 scored significantly better (p = 0,042) than participant group 1 on questions with the 'overview+detail' interface strategy. Participant group 2 got these questions as the second set whereas participant group got these questions as the first set. For 'task effectiveness' no significant difference between both participant groups was found.

It can therefore be cautiously concluded that a learning effect occurred. For a detailed analysis, please refer to Appendix E. As stated in section 3.1, the learning effect was negated by presenting the sets of questions in reverse order. Therefore, further analysis of the data could be carried out by aggregating the results of both groups.

The average time it took participants to complete the survey was 15 minutes and 47 seconds. The fastest participant finished the survey in 5 minutes and 42 seconds. The longest time it took a participant to finish the survey was 48 minutes and 16 seconds. If a participant finished quickly, it is not necessarily an indication on how well he or she performed. Figure 11 shows how long it took participants to complete the survey, with time binned by 5 minutes.

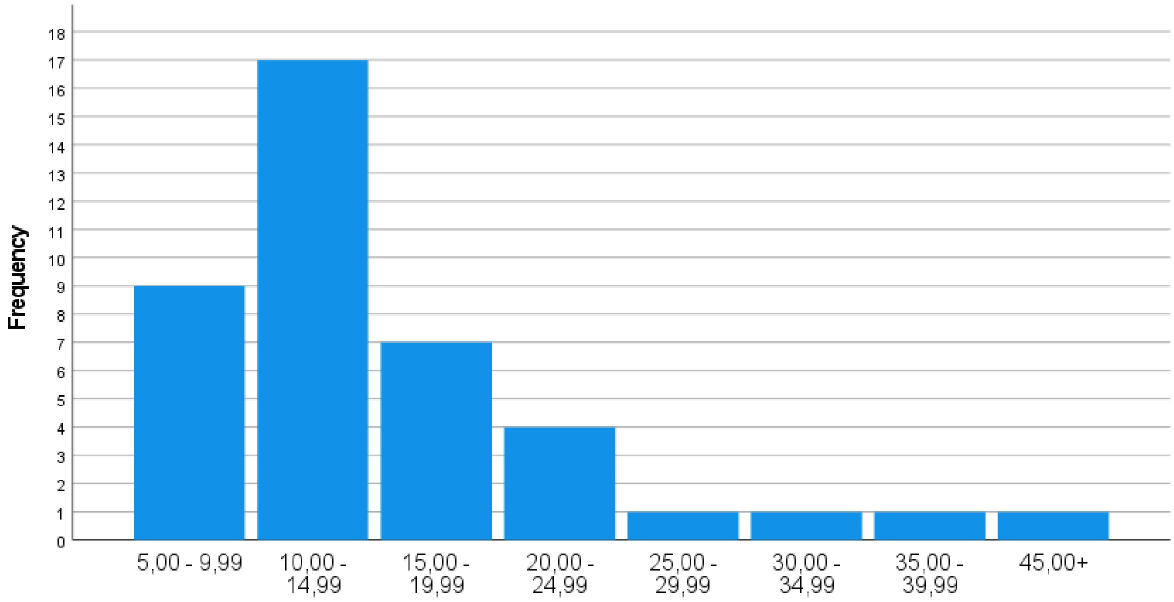


*Figure 11 - Histogram total time spent (binned per 5 minutes)*

The highest number of correct answers was twelve, the lowest six. Ten participants answered all twelve questions correctly. Two participants had only six questions right. The average of correct answers was 10,17.

An overview of the descriptives of the number of correct answers is shown in Table 5, Figure 12 shows the corresponding boxplot.

*Table 5 - Statistics of total number of correct answers*

| | Statistic | Std. Error |
|---|---|---|
| Mean | 10,17 | 0,277 |
| Median | 11,00 | |
| Variance | 3,145 | |
| Std. Deviation | 1,773 | |
| Minimum | 6 | |
| Maximum | 12 | |
| Skewness | -0,978 | 0,369 |
| Kurtosis | -0,012 | 0,724 |

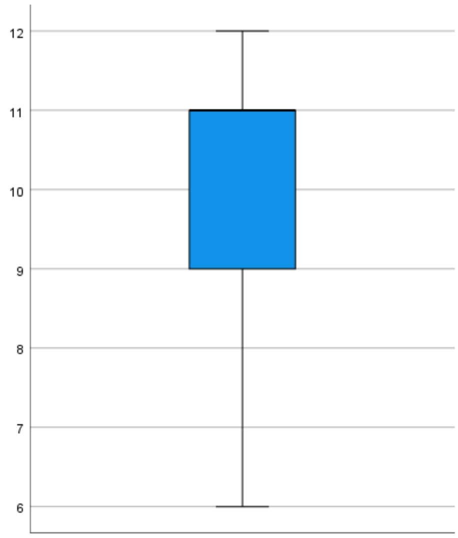


*Figure 12 - Boxplot number of correct answers*

Figure 13 shows how many participants got a certain number of questions right.

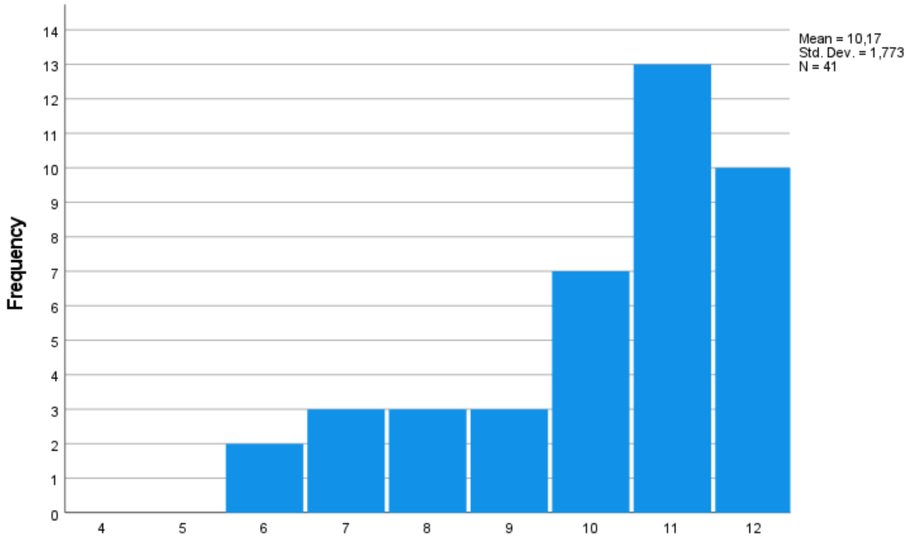


*Figure 13 - Histogram total number of correct answers*

Zooming in on the 'task effectiveness' per interface strategy, the highest number of correct answers for both strategies was six. The lowest number of correct answers for the 'overview+detail' strategy was three, for the 'focus+context it was one. To compare both strategies, Figure 14 shows a complete overview of the number of questions participants got right per interface strategy.

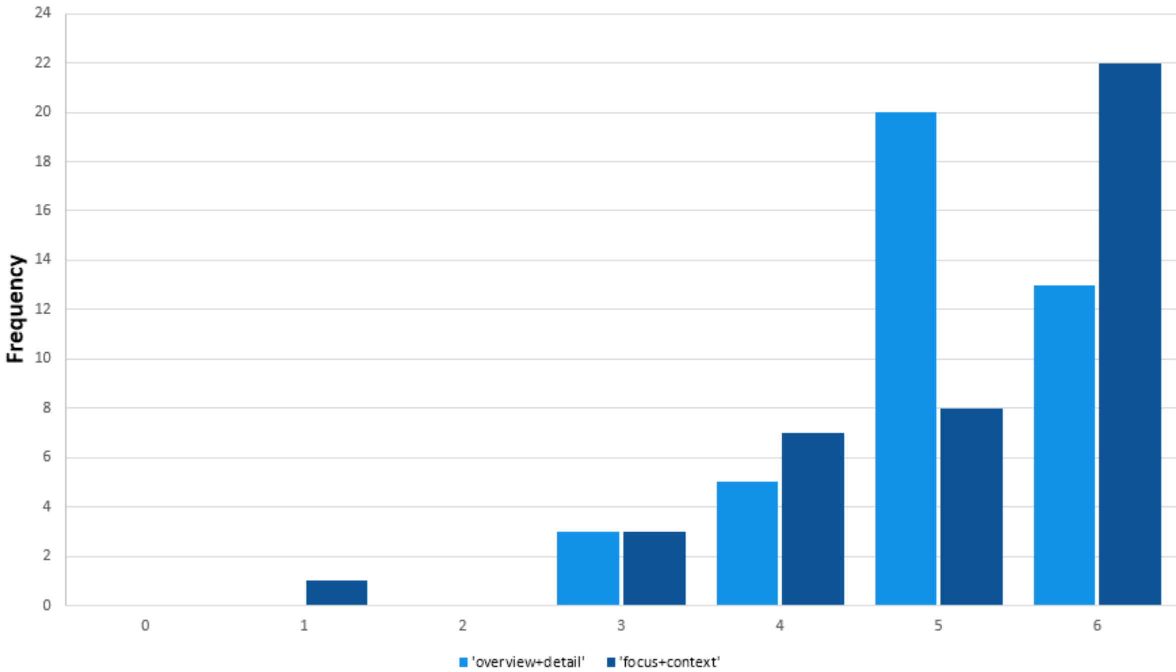


*Figure 14 - Histogram number of correct answers per interface strategy*

More than half (53,7%) of the participants answered all questions with 'focus+context' strategy correctly compared to 31,7% for the questions using the 'overview+detail' strategy, see Table 6 for the complete statistics.

*Table 6 - Frequency correct answers per interface strategy*

| Number of correct answers | Frequency correct answers | | | |
|---|---|---|---|---|
| | 'overview+detail' | | 'focus+context' | |
| | Frequency | Percentage | Frequency | Percentage |
| 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 1 | 2,4 |
| 2 | 0 | 0 | 0 | 0 |
| 3 | 3 | 7,3 | 3 | 7,3 |
| 4 | 5 | 12,2 | 7 | 17,1 |
| 5 | 20 | 48,8 | 8 | 19,5 |
| 6 | 13 | 31,7 | 22 | 53,7 |
| Total | 41 | 100,0 | 41 | 100,0 |

The statistics of the number of correct answers for the 'overview+detail' strategy is provided in Table 7, Figure 15 shows the boxplot.

*Table 7 - Statistics of total number of correct answers 'overview+detail'*

| | **Statistic** | **Std. Error** |
|---|---|---|
| Mean | 5,05 | 0135 |
| Median | 5,00 | |
| Std. Deviation | 0,865 | |
| Variance | 0,748 | |
| Minimum | 3 | |
| Maximum | 6 | |
| Skewness | -0,829 | 0,369 |
| Kurtosis | 0,376 | 0,724 |

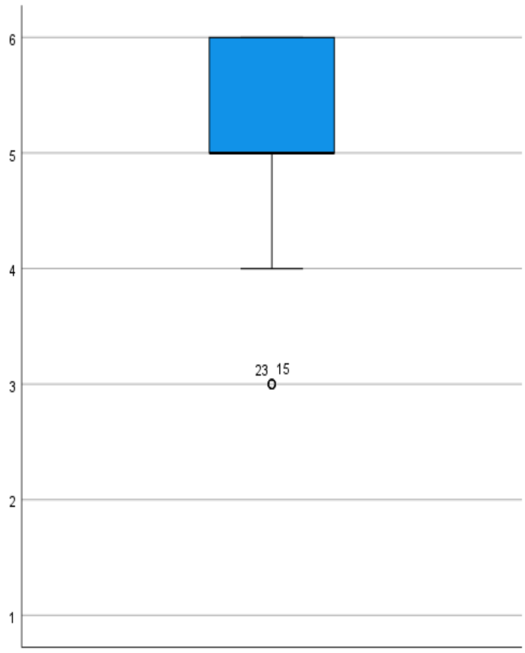


*Figure 15 - Boxplot number of correct answers 'overview+detail'*

Results for the ‘focus+context’ strategy are shown in Table 8, Figure 16 shows the corresponding boxplot.

*Table 8 - Statistics of total number of correct answers ‘focus+context’*

| | Statistic | Std. Error |
|---|---|---|
| Mean | 5,12 | 0,185 |
| Median | 6,00 | |
| Std. Deviation | 1,187 | |
| Variance | 1,410 | |
| Minimum | 1 | |
| Maximum | 6 | |
| Skewness | -1,471 | 0,369 |
| Kurtosis | 2,223 | 0,724 |

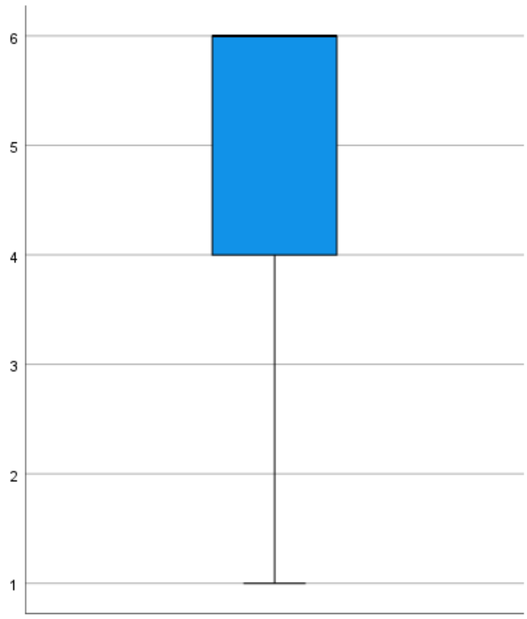


*Figure 16 - Boxplot number of correct answers ‘focus+context’*

In looking at the results, it is noticeable that the number of correct answers is quite high, as is the number of people who got all the questions right. Does this mean that the questions were too easy or is there another possible cause? Chapter 5 will reflect on this.

Looking at the ‘task efficiency’ per interface strategy, Figure 17 shows the number of correct answers per hour (binned) given by participants for both interface strategies.

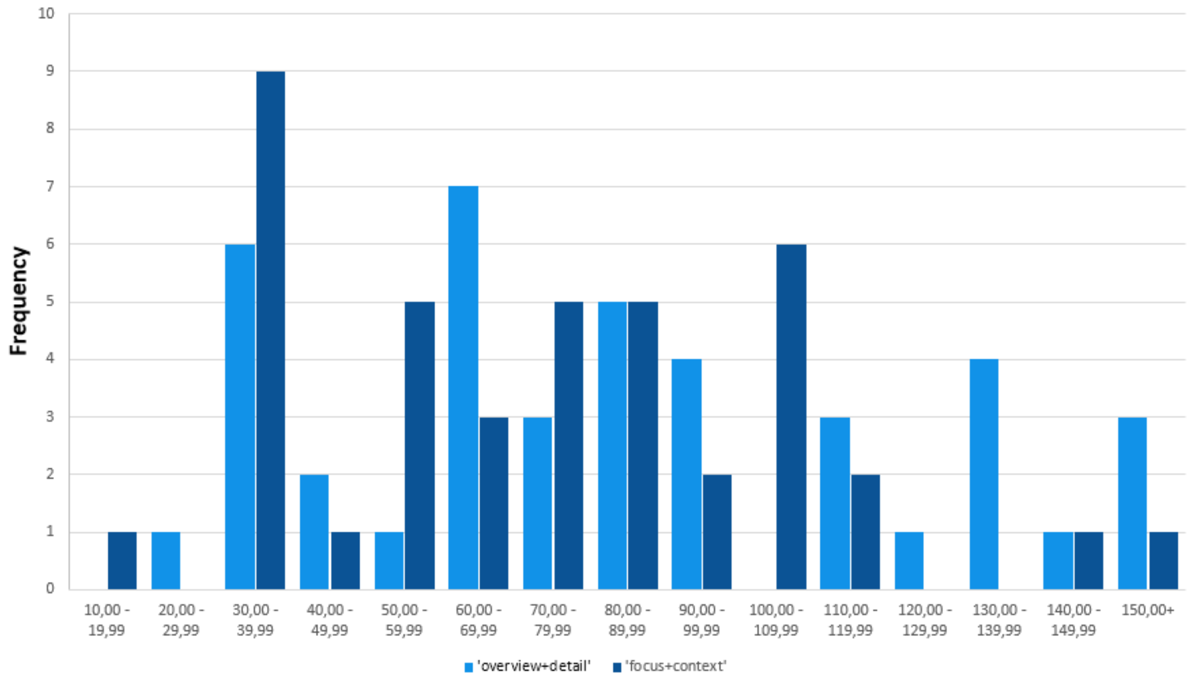


*Figure 17 - Histogram number of correct answers per hour (binned) per interface strategy*

The maximum number of correct answers per hour for questions using the 'focus+context' was higher (229,57) than for questions with the 'overview+detail' strategy (208,68). The minimum number of correct answers per hour was lower for questions with the 'focus+context' strategy than the 'overview+detail' strategy (14,20 versus 26,09). The average number of correct answers per hour was also lower for questions with the 'focus+context' (75,18) than for the 'overview+detail' strategy (87,48). Table 9 shows the mean, minimum and maximum of correct answers per hour for both interface strategies.

*Table 9 - Number of correct answers per hour per interface strategy*

| | **Correct answers per hour** | | | |
|---|---|---|---|---|
| | 'overview+detail' | | 'focus+context' | |
| | Statistic | Std. Error | Statistic | Std. Error |
| Mean | 87,48 | 6.836 | 75,18 | 5,981 |
| Median | 81,86 | | 74,32 | |
| Std. Deviation | 43,769 | | 38,299 | |
| Variance | 1915,744 | | 1466,803 | |
| Minimum | 26,09 | | 14,20 | |
| Maximum | 208,68 | | 229,57 | |
| Skewness | 0,860 | 0,369 | 1,635 | 0,369 |
| Kurtosis | 0,528 | 0,724 | 5,421 | 0,724 |

Figure 18 and Figure 19 show the boxplots of the 'task efficiency' for the 'overview+detail' and 'focus+context' strategy.

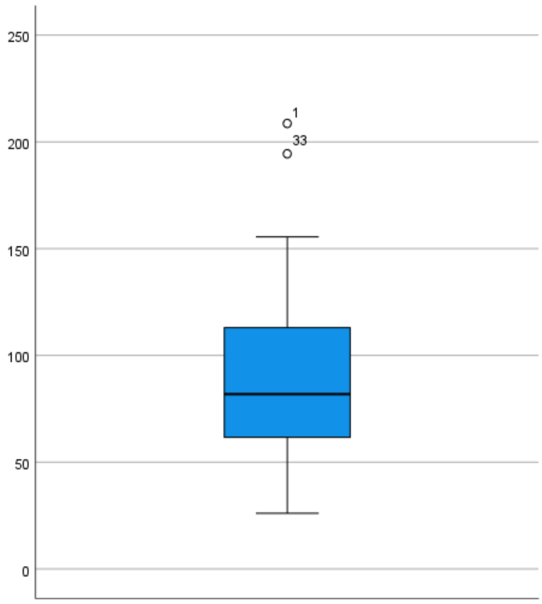


*Figure 18 - Boxplot task efficiency 'overview+detail'*

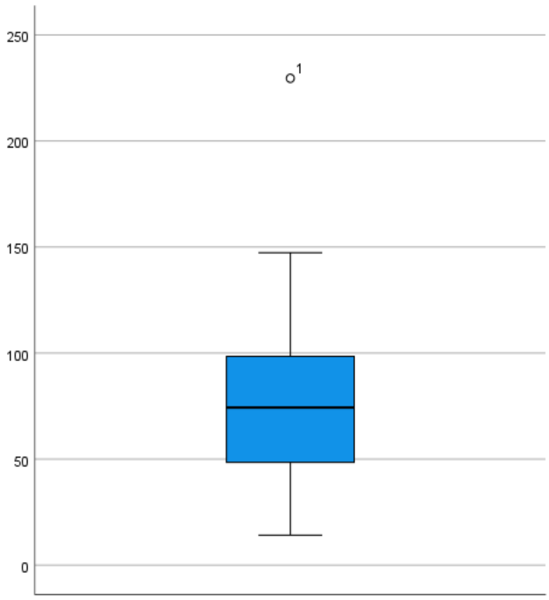


*Figure 19 - Boxplot task efficiency 'focus+context'*

Three confounding variables 'theoretical modeling competency', 'practical modeling competency' and 'education' were defined in section 3.1. Participants were asked to self-asses their theoretical knowledge about process modeling and practical use of process models. An overview of the answers to the questions about these personal factors is provided in Appendix F. None of the participants indicated that they had a lot of theoretical knowledge. This study focuses on the use of models by laypersons, so the group of participants seems appropriate in that respect.

## 4.2. Hypothesis testing

In order to determine if the independent variable ‘interface strategy’ has a significant impact on the dependent variables ‘task effectiveness’ and ‘task efficiency’, the data must first be checked for normality. Based on this result a test will be selected to check for significant differences.

### 4.2.1. Test for normality

By looking at the histograms per interface strategy, Figure 20 to Figure 23, one can already get an idea if the distribution of the data is normal.

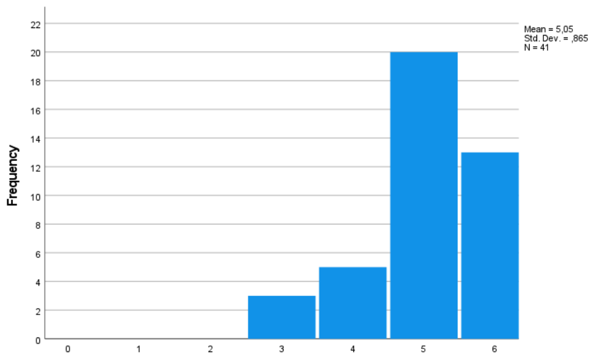


*Figure 20 - Histogram task effectiveness ‘overview+detail’*

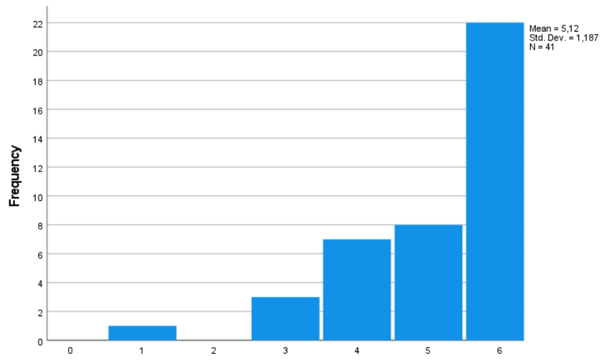


*Figure 21 - Histogram task effectiveness ‘focus+context'*

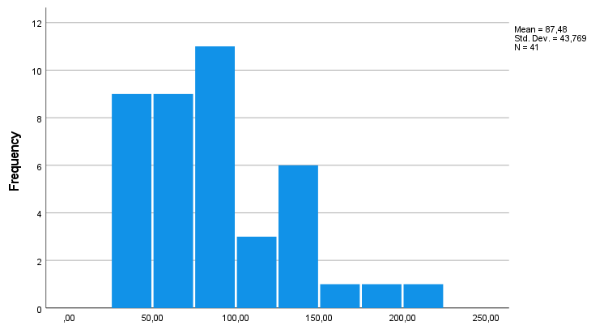


*Figure 22 - Histogram task efficiency ‘overview+detail’*

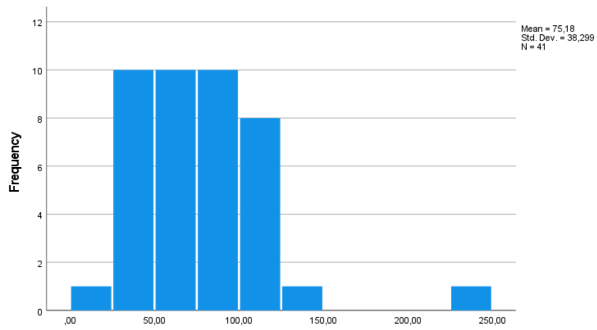


*Figure 23 - Histogram task efficiency ‘focus+context’*

Although the distribution of both dependent variables may not appear normal at first glance, this will be further checked with the Kolmogorov-Smirnov and Shapiro-Wilk tests (Field, 2017, p. 250). Both tests examine whether a distribution of scores is significantly different from a normal distribution. If a test is significant ($p < 0,05$), it indicates a deviation from normality. If the test is non-significant ($p > 0,05$), it is concluded that the data are normally distributed.

The results for variable ‘task effectiveness’ are shown in Table 10 and for variable ‘task efficiency’ in Table 11. All tests for ‘task effectiveness’ show a $p < 0,05$. It is therefore concluded that the data is not normally distributed. Note that in the tables the value ‘p’ is indicated by ‘Sig.’. For task efficiency, the results of the Kolmogorov-Smirnov test ($p > 0,05$) differ from the Shapiro-Wilk test ($p < 0,05$). As stated by Razali and Wah (2011) the Shapiro-Wilk test is the most powerful. Therefore, in this case, it was decided that the distribution of the data was also not normal.

*Table 10 - Tests of normality (task effectiveness)*

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| Number of correct answers ‘overview+detail’ | 0,282 | 41 | 0,000 | 0,818 | 41 | 0,000 |
| Number of correct answers ‘focus+context’ | 0,307 | 41 | 0,000 | 0,753 | 41 | 0,000 |

a. Lilliefors Significance Correction.

*Table 11 - Tests of normality (task efficiency)*

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| Number of correct answers per hour ‘overview+detail’ | 0,132 | 41 | 0,069 | 0,937 | 41 | 0,026 |
| Number of correct answers per hour ‘focus+context’ | 0,109 | 41 | 0,200* | 0,879 | 41 | 0,000 |

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction.

### 4.2.2. Selected test method

Due to the fact that the distribution of the data sets is non-normal, a non-parametric test has to be used. In this study two dependent variables have been defined, ‘task effectiveness’ and ‘task efficiency’, which both have continuous outcomes. The variables are studied independently of each other. Only one independent variable is defined, the ‘interface strategy’, which has two categorical values defined as ‘overview+detail’ and ‘focus+context’ strategy.

The results of all participants in this study are aggregated into one group. This means that the results for both the ‘overview+detail’ and the ‘focus+context’ questions come from the same group. This implies that the data are related. The non-parametric Wilcoxon signed-rank test is selected (Field, 2017, p. 297) because this test can be used when comparing sets of results that are related. In this test the median of a set of numbers are compared against a hypothetical median.

SPSS is used to perform the Wilcoxon signed-rank test.

### 4.2.3. Test task effectiveness

First, the test is used to determine whether there is a significant difference in the results for the dependent variable ‘task effectiveness’. The Wilcoxon signed-rank test tests the null hypothesis that states that the medians of both strategies are equal. Table 12 shows the result of the test.

*Table 12 - Hypothesis test summary for task effectiveness*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The median of differences between number of correct answers 'overview+detail' and number of correct answers 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,453 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

Because the resulting value from the test, p = 0,453, is higher than the critical value of 0,050, the null hypothesis of the Wilcoxon signed-rank test is retained. The outcome of the test shows that there is no difference in the medians between the number of correct answers for the 'overview+detail' and 'focus+context' interface strategies. It is therefore concluded that there is no significant difference in 'task effectiveness' between both interface strategies. See Appendix G for more details

### 4.2.4. Test task efficiency

The same Wilcoxon signed-rank test is used for the second dependent variable 'task efficiency'. The results are shown in Table 13.

*Table 13 - Hypothesis test summary for task efficiency*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The median of differences between number of correct answers per hour 'overview+detail' and number of correct answers per hour 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,236 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

Since the significance value of p = 0,236 is higher than the critical value of 0,050, the null hypothesis of the test is also held for 'task efficiency'. Therefore, it is concluded that there is no significant difference in 'task efficiency' between the two interface strategies. For more details, see Appendix G.

## 4.3. Confounding variables

The confounding variables, 'theoretical modeling competency', 'practical modeling competency' and 'education' were operationalized by asking the participant to self-assess their theoretical knowledge of business process modeling and their practical experience with process models. They were also asked to indicate their highest level of education.

The data were analyzed to see whether these confounding variables have a significant impact on the understanding of the process models. The Independent-Samples median test in SPSS was used for this analysis. Participants were divided into separate groups by using their self-assessed levels of 'theoretical modeling competency', 'practical modeling competency' and 'education'. The

Independent-Samples median test checks whether the medians of the dependent variables are the same across the categories of the confounding variables. This test was done for both dependent variables, 'task effectiveness' and 'task efficiency'. A summary of all the test values is given in next sections, see more details in Appendix H.

### 4.3.1. Theoretical modeling competency

The summary of the Independent-Samples median test for 'theoretical modeling competency' is shown in Table 14. The test is performed for 'task effectiveness' as well as 'task efficiency' for both interface strategies: 'overview+detail' and 'focus+context'.

*Table 14 - Hypothesis test summary theoretical knowledge*

**Hypothesis Test Summary**

| Dependent variable | Null Hypothesis | Sig.[a,b] | Decision |
|---|---|---|---|
| Task effectiveness | The medians of number of correct answers 'overview+detail' are the same across categories of theoretical knowledge. | 0,203 | Retain the null hypothesis. |
| | The medians of number of correct answers 'focus+context' are the same across categories of theoretical knowledge. | [c] | Unable to compute |
| Task efficiency | The medians of number of correct answers per hour 'overview+detail' are the same across categories of theoretical knowledge. | 0,099 | Retain the null hypothesis. |
| | The medians of number of correct answers per hour 'focus+context' are the same across categories of theoretical knowledge. | 0,618 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.
c. All test field values are less than or equal to the median.

Because all values for p are greater than 0,050 (values 'Sig.' in the table), it is concluded that the confounding variable 'theoretical modeling competency' has no significant impact on the understanding of the business process models, either in terms of 'task effectiveness' or 'task efficiency'.

### 4.3.2. Practical modeling competency

shows the outcomes of the Independent-Samples median tests for 'practical modeling competency'. The tests are performed for 'task effectiveness' as well as 'task efficiency', again for both interface strategies.

*Table 15 - Hypothesis test summary practical experience*

**Hypothesis Test Summary**

| Dependent variable | Null Hypothesis | Sig.[a,b] | Decision |
|---|---|---|---|
| Task effectiveness | The medians of number of correct answers 'overview+detail' are the same across categories of practical experience. | 0,289 | Retain the null hypothesis. |
| | The medians of number of correct answers 'focus+context' are the same across categories of practical experience. | [c] | Unable to compute |
| Task efficiency | The medians of number of correct answers per hour 'overview+detail' are the same across categories of practical experience. | 0,437 | Retain the null hypothesis. |
| | The medians of number of correct answers per hour 'focus+context' are the same across categories of practical experience. | 0,335 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.
c. All test field values are less than or equal to the median.

Also in these tests, all values for p are greater than 0,050 (values 'Sig.' in the table). Thus, it is concluded that the confounding variable 'practical modeling competency' does not have a significant impact on the understanding of the business process models, either in terms of 'task effectiveness' or 'task efficiency'.

### 4.3.3. Education

To see whether the highest level of educations plays a significant role in the results, the Independent-Samples median test is also performed for confounding variable 'education'. The summary for 'task effectiveness' as well as 'task efficiency' for both interface strategies.is shown in Table 16.

*Table 16 - Hypothesis test summary education*

**Hypothesis Test Summary**

| Dependent variable | Null Hypothesis | Sig.[a,b] | Decision |
|---|---|---|---|
| Task effectiveness | The medians of number of correct answers 'overview+detail' are the same across categories of highest education. | 0,786 | Retain the null hypothesis. |
| | The medians of number of correct answers 'focus+context' are the same across categories of highest education. | [c] | Unable to compute |

| | | | |
|---|---|---|---|
| Task efficiency | The medians of number of correct answers per hour ‘overview+detail’ are the same across categories of highest education. | 0,452 | Retain the null hypothesis. |
| | The medians of number of correct answers per hour ‘focus+context’ are the same across categories of highest education. | 0,452 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.
c. All test field values are less than or equal to the median.

For the tests that could be carried out, all values for p are greater than 0,050 (values ‘Sig.’ in the table). Therefore, it is concluded from this analysis that the confounding variable ‘education’ has no significant impact on the understanding of the business process models, either in terms of ‘task effectiveness’ or ‘task efficiency’.

## 4.4. Local and global questions

As mentioned in section 3.2.2, the questions can be divided into two groups, local and global questions. To be able to answer so-called local questions, one only needs to have knowledge about the part of the process represented by the sub-process. To answer the global questions, a participant needs to have an overview of the bigger picture and thus possess more information than is represented by the sub-process. Table 68 in Appendix I, shows the which questions are considered ‘local’ and which are ‘global’ questions.

In order to check whether the difference between these two types of questions leads to significant differences in ‘task effectiveness’ and ‘task efficiency’ when using the ‘overview+detail’ or ‘focus+context’ interface strategy, the results are grouped into two groups. One group with the results of the local questions and the other group containing the results of the global questions. After splitting the data, they are analyzed with the Wilcoxon signed-rank test. The Wilcoxon signed-rank test is used because the data from all participants are aggregated and therefore related. Details of the analysis are listed in Appendix I.

### 4.4.1. Task effectiveness

The null hypothesis of the Wilcoxon signed-rank test states that the medians of both strategies are equal. For both the local ($p = 0{,}157$) as well as the global questions ($p = 0{,}653$) the null hypothesis of the Wilcoxon signed-rank test is retained ($\alpha = 0{,}05$). It is concluded that when looking at ‘task effectiveness’ no significant difference between both interface strategies can be determined.

### 4.4.2. Task efficiency

Using the ‘overview+detail’ strategy together with local questions seems to result in a significant better result in ‘task efficiency’. The null hypothesis that the results are equal for local questions is rejected ($p = 0{,}004$) when comparing the ‘task efficiency’ of the ‘overview+detail’ and ‘focus+context’ interface strategies.

This may indicate that participants understand the process model faster when using the 'overview+detail' interface strategy in combination with local questions.

For the global questions, the null hypothesis of the Wilcoxon signed-rank test is retained ($p = 0{,}946$). So, in the case of global questions no significant difference between both strategies can be identified.

# 5. Conclusions, discussion and recommendations

This chapter will reflect on the research that has been carried out. The results of the literature review and the research design will be assessed. The significance of the results found in this study will be discussed, as well as how they relate to previous studies. Also, the impact of the results on the use of sub-processes in business process models in practice will be indicated. Finally, some recommendations for possible further research will be provided.

## 5.1. Conclusions

Based on the outcome of the survey, the hypothesis of this study: **the use of 'overview+detail' interface strategy will result in better understanding of the process model when used by a layperson than the use of 'focus+context' interface strategy** must be rejected.

The hypothesis was inspired by the observation made by Figl et al. (2013) that experts prefer the 'overview+detail' strategy to the 'focus+context' strategy in a reading task. Following this observation and further literature research, it was decided to investigate whether the use of these two different interface strategies in process models leads to differences in understanding among laypersons.

In the design of the study, the interface strategy, being 'overview+detail' and 'focus+context', was defined as the independent variable. Understandability was defined by two dependent variables 'task effectiveness' and 'task efficiency'.

This research did not demonstrate a significant difference in the results between the two interface strategies and the dependent variables in the case of laypersons using the models.

In addition to the independent variable, interface strategy, three personal factors were defined as confounding variables. Participants were asked to self-assess their 'theoretical modeling competency', their 'practical modeling competency' and were asked to provide their highest level of 'education'. This study revealed no significant difference between the three confounding variables and the two dependent variables.

This research also examined the difference in results between so-called 'local' and 'global' questions. A significant difference in the 'task efficiency' was observed between the 'overview+detail' and 'focus+context' strategies for participants answering local questions. The 'task efficiency' for the 'overview+detail' interface strategy was higher than for the 'focus+context' strategy when answering local questions. This could mean that using the 'overview+detail' interface strategy is to be preferred when only knowledge of a sub-process is needed.

## 5.2. Discussion – reflection

### 5.2.1. Reflection on the literature research

The literature review of this study sought to answer the following question:

**Which process model modularization strategies are described in the literature?**

The literature search resulted in a large number of documents from which it was difficult to select the relevant ones. This was partly due to the fact that terms such as: modularization, decomposition and sub-processes are used for many diverse things in the literature. Furthermore, the naming convention is ambiguous, e.g. vertical modularization and hierarchical structuring are often used interchangeably. From the experience acquired during the research, it would be recommendable to define and use unambiguous terms.

As a results of the literature review, four modularization strategies were identified: horizontal, vertical, orthogonal, and 'flattened with groups'. In the case of the vertical modularization strategy, two associated interface strategies: 'overview+detail' and 'focus+context' were of interest.

Next step in the literature review was finding the answer to the question:

**What is the impact of the process model modularization strategies on the understandability of the process model?**

It was concluded that the literature and the studies already conducted did not provide an unambiguously answer to this question. No study had examined all modularization strategies in one study. All studies found in the literature were limited to a subset of strategies.

As this study would otherwise be too broad, it too had to be limited to a select number of strategies. Therefore, the following question had to be answered:

**Which of those strategies are selected as most interesting for further research in practice?**

The vertical modularization strategy with the interface strategies 'overview+detail' and 'focus+context' was selected as most interesting for follow-up research.

Based on the results of the literature review, the hypothesis was limited to those two interface strategies linked to the vertical modularization strategy and was formulated as:

**The use of 'overview+detail' interface strategy will result in better understanding of the process model when used by a layperson than the use of 'focus+context' interface strategy.**

### 5.2.2. Reflection on the research strategy

A research strategy was defined to be able to answer the question:

**Which of the selected modularization strategies actually results in the highest understandability of the process by laypersons?**

An experiment was designed to try to establish a causal relationship between the interface strategy of vertical modularization and the understandability of a process model. LimeSurvey was used to set up an experiment in which participants were asked to answer questions about a process model of a fictitious process using the 'overview+detail' and the 'focus+context' interface strategy.

The use of LimeSurvey has limitations relating to a participant's interaction with the process displayed on the screen. For example, the software did not provide the ability to actually click on the process model to open or close sub-processes. Therefore, the choice was made to present the entire process model and adjacent to that the model with the sub-process opened in a way depending on the interface strategy. This could have affected the outcome of the experiment because the participants were already presented with the right sub-process opened in the model. Because this was the case for both interface strategies, it probably did not affect the comparison of both two strategies.

Fifty-six persons were asked to participate in the study. These people all belong to a certain social group such as colleagues, friends and family. Forty-five people did participate. It is possible that only people who are interested in process models did finish the experiment. This could mean that no proper reflection of the population, namely laypersons, was established. An interesting questions is: why did some invitees not participate? Is it because of lack of time or are they afraid to participate because they think they don't know enough about process models? Even so, by recruiting participants with different working backgrounds the external validity was sufficiently guaranteed.

The external validity was also substantiated by the fact that only six participants indicated to have rather strong theoretical knowledge of process modeling and none indicated to have strong theoretical knowledge. Regarding the ‘practical modeling competency’, the majority of the participants (32) indicated that they never or sometimes use process modeling in practice. Although this is a self-assessed estimate, it is reasonable to assume that the participants are not experts in process modeling.

Another point of attention is what device someone used when taking the survey. No instructions were given as to which device should be used, nor were participants asked which device they used. The results of a participant using a smartphone when answering the questions could be different than the results of a participant using a personal computer with a monitor.

A model of a fictitious process is used in the experiment. In this model, natural language was not used to describe the process steps, but a more abstract description using letters and numbers. It is possible that this abstract representation affected some participants’ understanding of the models.

### 5.2.3. Reflection on the results

Most of the participants indicated that they did not have much knowledge of process modeling and that they did not use process models often in their daily practice. After the survey was closed and the results analyzed, it appeared that the average number of correct answers given by the participants was quite high. This meant that the dependent variable ‘task effectiveness’ was consequently also high.

This could indicate that the questions were too easy or the model was too simple The model was deliberately kept fairly simple. For example, there were no crossing edges, the data flow was fairly consistent and only ‘exclusive or gateways’ were used all of which had only two outputs. The number of different symbols used in the models was also limited.

In addition, however, it could also mean that participants took great pains, and thus extra time, to answer the questions correctly. This could mean that the dependent variable ‘task efficiency’ might be a better variable to establish a difference between the interface strategies than ‘task effectiveness’.

It was decided to present all participants with the same sets of questions rather than creating two groups of participants where one group was presented with one set of questions using one interface strategy and the other group would be presented with the set of questions using the other strategy. merging the participants into one group resulted in a slightly more difficult statistical testing process. The number of participants was limited, which could be the reason for not finding a statistically supported difference between the two interface strategies.

When analyzing the results between the global and local questions, it was found that the ‘task efficiency’ for the ‘overview+detail’ interface strategy was higher than for the ‘focus+context’ strategy when asking local questions. Although this was not a main objective of the study, it is potentially an interesting result to investigate further.

## 5.3. Recommendations for practice

The study revealed no significant difference between the results of the two interface strategies. Therefore, based on the results of this study, no recommendation can be made as to which strategy would be better to apply in process models used by laypersons in practice.

A tentative recommendation is that when a user only needs knowledge of certain sub-processes it is better for ‘task efficiency’ to use process models using the 'overview+detail' strategy.

## 5.4. Recommendations for further research

This research has faced some limitations. For example, the number of people who participated in the survey was limited. A study with more participants, divided into two groups, could possibly lead to different results. This would mean a study using the between-groups approach, instead of the within-subjects approach used in this study.

Another limitation was caused by a technical constraint of LimeSurvey. It was not possible to create models where participants could open and close the sub-processes themselves. The research could be repeated using a tool in which it is possible to create models that do allow opening and closing of sub-processes by participants.

Further research could also examine whether using natural language to describe process steps, rather than the more abstract form used in this study, leads differences in understandability.

A number of findings emerged from the survey that could be explored in the future. It would be interesting to investigate whether there is indeed any difference between the interface strategies when used for questions where a user only needs knowledge of a sub-process, compared to questions where knowledge is needed of the overall process.

This study has focused on the vertical modularization using two different interface strategies. Other strategies, like horizontal modularization, were left out of this study. Those strategies could be included in follow-up studies to see whether a significant difference in understandability of models can yet be established.

This experiment has limited vertical modularization of the models to two levels. One level with collapsed sub-processes, and one level deeper where the contents of the sub-processes are shown. It might be interesting to investigate what the results will be when a model consists of more layers with sub-processes.

## References


Blankenship, J., & Dansereau, D. F. (2000, 2000/01/01). The Effect of Animated Node-Link Displays on Information Recall. *The Journal of Experimental Education, 68*(4), 293-308. https://doi.org/10.1080/00220970009600640

Cockburn, A., Karlson, A., & Bederson, B. B. (2009). A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv., 41*(1), Article 2. https://doi.org/10.1145/1456650.1456652

Dikici, A., Turetken, O., & Demirors, O. (2018). Factors influencing the understandability of process models: A systematic literature review. *Information and Software Technology, 93*, 112-129.

Field, A. (2017). Discovering statistics using IBM SPSS statistics, 5th edition.

Figl, K. (2017). Comprehension of Procedural Visual Business Process Models. *Business & Information Systems Engineering, 59*, 41-67.

Figl, K., Koschmider, A., & Kriglstein, S. (2013). Visualising process model hierarchies.

Mendling, J., Strembeck, M., & Recker, J. (2012, 2012/04/01/). Factors of process model comprehension—Findings from a series of experiments. *Decision Support Systems, 53*(1), 195-206. https://doi.org/https://doi.org/10.1016/j.dss.2011.12.013

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The, P. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moody, D. (2006). Dealing with "Map Shock": A Systematic Approach for Managing Complexity in Requirements Modelling. *REFSQ*.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics, 2*(1), 21-33.

Reijers, H., Freytag, T., Mendling, J., & Eckleder, A. (2011, 06/01). Syntax highlighting in business process models. *Decision Support Systems, 51*, 339-349. https://doi.org/10.1016/j.dss.2010.12.013

Reijers, H., & Mendling, J. (2008). *Modularity in Process Models: Review and Effects*. https://doi.org/10.1007/978-3-540-85758-7_5

Reijers, H., & Mendling, J. (2011, 06/01). 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*, 449-462. https://doi.org/10.1109/TSMCA.2010.2087017

Reijers, H. A., Mendling, J., & Dijkman, R. M. (2011). Human and automatic modularizations of process models to enhance their comprehension. *Information Systems, 36*(5), 881-897.

Sanchez-Gonzalez, L., Garcia, F., Ruiz, F., & Piattini, M. (2017). A case study about the improvement of business process models driven by indicators. *Software and systems modeling, 16*(3), 759-788. https://doi.org/10.1007/s10270-015-0482-0

Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research methods for business students* (eighth edition ed.). Pearson education.

Turetken, O., Dikici, A., Vanderfeesten, I., Rompen, T., & Demirors, O. (2019). The Influence of Using Collapsed Sub-processes and Groups on the Understandability of Business Process Models. *Business & Information Systems Engineering, 62*(2), 121-141. https://doi.org/10.1007/s12599-019-00577-4

Winter, M., Pryss, R., Probst, T., Baß, J., & Reichert, M. (2020). Measuring the Cognitive Complexity in the Comprehension of Modular Process Models. *IEEE Transactions on Cognitive and Developmental Systems*, 1-1. https://doi.org/10.1109/TCDS.2020.3032730

# Appendix A – Literature selection and review

## Design and summary of the literature selection process

The defined stages of the PRISMA 2009 Flow Diagram (Moher et al., 2009), as displayed in Figure 24, were used to find and select the literature to be used in the literature review.

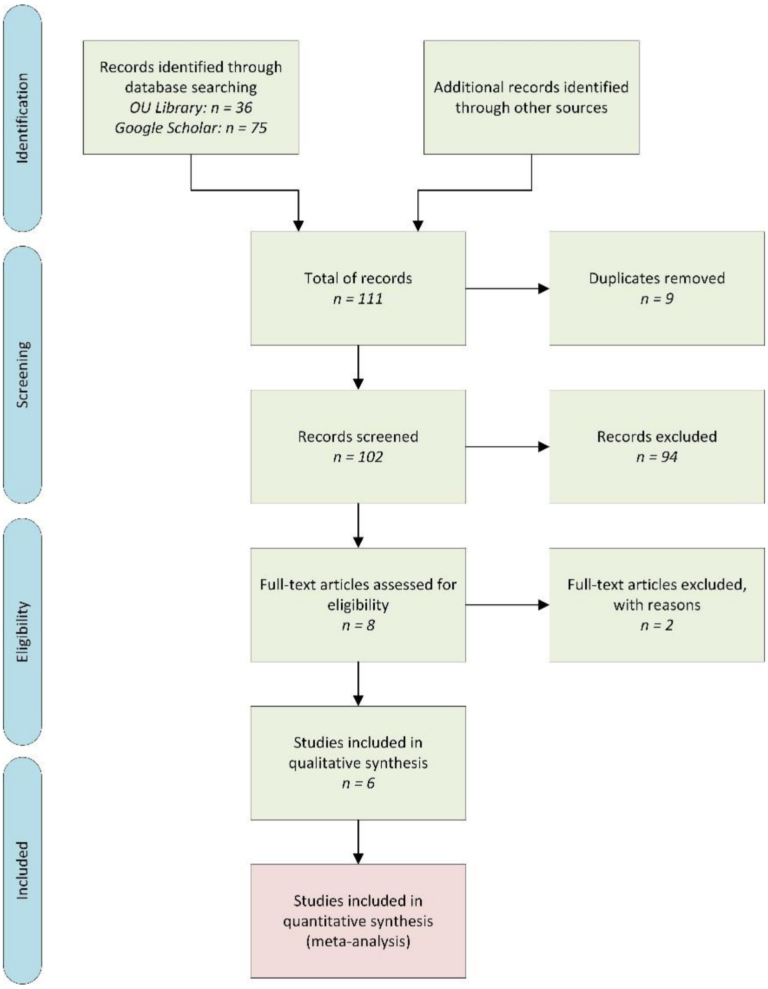


*Figure 24 - PRISMA 2009 Flow Diagram*

### Identification phase

During the identification phase, the Open University (OU) Library and Google Scholar were used to find and select relevant literature. In order to find suitable literature for the study, search criteria were formulated.

Although the search criteria were carefully selected, searching in Google Scholar resulted in too many articles many of which were not related to the subject of this research. This was caused by the fact that the search performed in Google Scholar was on the entire text and was not limited to the title. To make a selection, the first fifty search results, ordered by relevance, of all performed searches within Google Scholar were added to a list: 'My library'. Subsequently, the titles of all remaining articles were scanned. If the title indicated that the article was not of interest for this study, it was removed from the list (E.g.: "Evaluating the surgeon's assistant: Results of a pilot study"). Thereafter, the abstracts of the still remaining articles were examined to see whether they should be added to the potentially relevant articles. The foregoing means that some screening was already done to limit it to meaningful articles. During this phase, no articles were added from alternative sources.

Remaining articles: 111.

### Screening phase

Firstly, in the screening phase, the results from both libraries (OU Library and Google Scholar) were deduplicated. Nine articles were identical, deduplicating resulted in 102 remaining articles.

Secondly, the titles (in the case of search results of the OU Library) and abstracts of all remaining articles were scanned for relevancy to the topic of this study. E.g., articles specifically dealing with declarative business process models were discarded.

As with the search in the OU Library, normally only peer-reviewed publications would be included. However some exceptions were made. These concerned: two master theses, one article from the proceedings of the 21st European Conference on Information Systems and one article which was not yet published. Those were all included because the subject of the articles were closely related to subject of this study.

Remaining articles: 8.

### Eligibility phase

After reading the complete articles in the eligibility phase, two articles were excluded because they covered (parts of) the same research by the same authors.

Remaining articles: 6.

### Included phase

Because a good and limited selection of articles was made in previous phases, this phase would add nothing more and was therefore skipped.

After all the phases have been completed, six articles were included in this study.

## Conducting identification phase in detail

### Searching the OU Library

The conduct of the search, as described in section 2.1, and its results are described in this appendix. Firstly, to identify relevant literature on the subject of modularization during the identification phase, some pilot searches were performed within the OU Library. Table 17 shows the general search criteria used in the OU Library.

*Table 17 - General search terms (OU Library)*

| General search terms | |
|---|---|
| Date of publication | Between 01-01-2000 and 01-10-2020 |
| Content type | All |
| Discipline | Computer Science |
| Language | English |
| Limited to | Peer-reviewed publications |

Results from other sources than the collection in the OU library were added, see Figure 25.

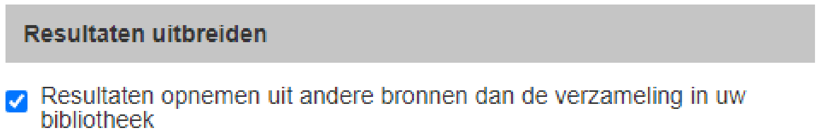


*Figure 25 - Results other sources added*

The search terms as shown in Table 18, were used in the pilot searches.

*Table 18 - Initial search criteria*

| Criterium 1 | AND | Criterium 2 | AND | Criterium 3 | AND | Criterium 4 |
|---|---|---|---|---|---|---|
| Process model | OR | Complexity | OR | Modularization | OR | Understandability |
| Business process model | | Complexness | | Modularisation | | Comprehensibility |
| | | Difficulty | | Decomposition | | Expressiveness |
| | | | | Modularity | | Comprehension |
| | | | | Sub-processes | | |
| | | | | Hierarchy | | |

After some consideration, criterium 2 and 4 were combined because the results showed that terms such as “Difficulty” and “Understandability” were used to denote the same thing. Also, searching for the exact term “Business process model” instead of “Process model” did not contribute to a better result. It unnecessarily limited the results.

To find synonyms, the online dictionaries: merriam-webster.com and dictionary.cambridge.org were used. The initial set of search terms was also expanded using the so-called “building blocks” method. This involved finding alternative terms by scanning for keywords in the titles and abstracts of articles found in the initial search.

Also, comparing the results from searching in the abstract only and the results from searching in the entire article resulted in new search terms. E.g. search term “Understand*”, using a wildcard, was added instead of just “Understandability”. And “Benefits of” was added to search for in the abstract. This resulted in the search criteria as shown in Table 19.

*Table 19 - Search criteria*

| Criterium 1 | AND | Criterium 2 | AND | Criterium 3 |
|---|---|---|---|---|
| Process model | OR | Modulari* | OR | Understand* |
| | | Decomposition | | Comprehens* |
| | | Sub process* | | Expressiveness |
| | | Sub-process* | | Complexity |
| | | Subprocess* | | Complexness |
| | | | | Difficulty |
| | | | | Readability |
| | | | | Benefits of |

Table 20 shows which search terms are covered by the use of a wildcard.

*Table 20 - Use of wildcard in search criteria*

| Use of asterisk in search criteria | Values |
|---|---|
| Modulari* | Modularization |
| | Modularisation |
| | Modularity |
| Sub process* | Sub process |
| | Sub processes |
| Sub-process* | Sub-process |
| | Sub-processes |
| Subprocess* | Subprocess |
| | Subprocesses |
| Understand* | (Difficult to) understand |
| | Understandability |
| Comprehens* | Comprehensibility |
| | Comprehension |

Subsequently, the search criteria, as shown in Table 19, were used in separate searches. This meant separate queries using "process model" (criterium 1) combined with one of the criteria from criterium 2 and one from criterium 3. The search term "process model" was searched for in the abstract, the other search terms were used to search for in the entire text. The number of results of each of those searches are shown in Table 21.

*Table 21 - Search results (OU Library) (d.d. 10-11-2020)*

| Search terms | | | Results |
|---|---|---|---|
| Process model | Modulari* | Understand* | 55 |
| Process model | Decomposition | Understand* | 130 |
| Process model | Sub process* | Understand* | 73 |
| Process model | Sub-process* | Understand* | 73 |
| Process model | Subprocess* | Understand* | 74 |
| | | | |
| Process model | Modulari* | Comprehens* | 42 |
| Process model | Decomposition | Comprehens* | 82 |
| Process model | Sub process* | Comprehens* | 52 |
| Process model | Sub-process* | Comprehens* | 52 |
| Process model | Subprocess* | Comprehens* | 49 |
| | | | |
| Process model | Modulari* | Expressiveness | 16 |
| Process model | Decomposition | Expressiveness | 22 |
| Process model | Sub process* | Expressiveness | 23 |

| | | | |
|---|---|---|---|
| Process model | Sub-process* | Expressiveness | 22 |
| Process model | Subprocess* | Expressiveness | 18 |
| | | | |
| Process model | Modulari* | Complexity | 46 |
| Process model | Decomposition | Complexity | 132 |
| Process model | Sub process* | Complexity | 58 |
| Process model | Sub-process* | Complexity | 58 |
| Process model | Subprocess* | Complexity | 61 |
| | | | |
| Process model | Modulari* | Complexness | 0 |
| Process model | Decomposition | Complexness | 0 |
| Process model | Sub process* | Complexness | 0 |
| Process model | Sub-process* | Complexness | 0 |
| Process model | Subprocess* | Complexness | 0 |
| | | | |
| Process model | Modulari* | Difficulty | 19 |
| Process model | Decomposition | Difficulty | 45 |
| Process model | Sub process* | Difficulty | 22 |
| Process model | Sub-process* | Difficulty | 22 |
| Process model | Subprocess* | Difficulty | 16 |
| | | | |
| Process model | Modulari* | Readability | 6 |
| Process model | Decomposition | Readability | 17 |
| Process model | Sub process* | Readability | 12 |
| Process model | Sub-process* | Readability | 12 |
| Process model | Subprocess* | Readability | 8 |
| | | | |
| Process model | Modulari* | Benefits of | 35 |
| Process model | Decomposition | Benefits of | 62 |
| Process model | Sub process* | Benefits of | 46 |
| Process model | Sub-process* | Benefits of | 46 |
| Process model | Subprocess* | Benefits of | 37 |

This way, it was found that no documents were found with search term “Complexness”. For this reason, “Complexness*” has not been used in further searches. Also, “sub-process” instead of “sub process” did not result in new search results, so “sub-process” was disregarded in the final search term. The final search criteria are shown in Table 22.

*Table 22 - Final set of search criteria*

| Criterium 1 | AND | Criterium 2 | AND | Criterium 3 |
|---|---|---|---|---|
| Process model | OR | Modulari* | OR | Understand* |
| | | Decomposition | | Comprehens* |
| | | Sub process* | | Expressiveness |
| | | Subprocess* | | Complexity |
| | | | | Difficulty |
| | | | | Readability |
| | | | | Benefits of |

The results of the queries where the search term "process model" was searched for in the abstract and the other terms were searched for in the entire document were compared to the results where all search terms were searched for in the abstract. The result of this comparison was used to determine whether the search terms were adequate and whether searching in the abstract was sufficient. After this comparison, it could be concluded that searching in the abstract was sufficient and that searching in the complete text did not add value because it did not result in more articles. The complete query and the (number of) results in the OU Library is included in Appendix B. A total of 36 articles remained from the search in the OU Library.

## Searching Google Scholar

Search engine Google Scholar was primarily used to verify the search results from the OU library. The general search criteria used in Google Scholar are shown in Table 23.

*Table 23 - General search terms (Google Scholar)*

| General search terms | |
|---|---|
| Date of publication | Between 2000 and 2020 |
| Language | English |
| Including patents | Yes |
| Including citations | Yes |

The numbers of search results from Google Scholar are shown in Table 24.

*Table 24 - Search results (Google Scholar) (d.d. 14-11-2020)*

| Search terms | Results |
|---|---|
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Understandability" | 3.100 |
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Comprehensibility" | 1.580 |

| | |
|---|---|
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Comprehension" | 6.850 |
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Expressiveness" | 4.300 |
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Complexity" | 19.400 |
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Complexness" | 17 |
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Difficulty" | 17.400 |
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "Readability" | 4.200 |
| "process model" AND ("Modularization" OR "Modularisation" OR "Modularity" OR "Decomposition" OR "Sub process" OR "Sub processes" OR Sub-process OR Subprocess) AND "benefits of" | 17.300 |

The first fifty search results, ordered by relevance, of all searches within Google Scholar were added to a list: ‘My library’. Deduplicating and screening the titles resulted in 75 articles.

## Conducting the screening phase in detail

In the screening phase, the remaining articles resulting from the queries in the OU Library (36) and Google Scholar (75) were deduplicated and merged. There were nine duplicates, leaving 102 articles. During the screening phase, the titles (in the case of search results of the OU Library) and abstracts of the remaining articles were scanned for relevancy to the topic of this study. Although only peer-reviewed publications were to be included, some exceptions were made, which are explained in Table 25. Publications which were not accessible were discarded. After the screening, eight potentially relevant articles remained.

*Table 25 - Remaining articles after screening phase*

| Date | Author(s) | Title |
|---|---|---|
| 2008 | Reijers, H.A; Mendling, J; Dumas, M ; More… | Modularity in process models: review and effects |
| 2011 | Reijers, H.A; Mendling, J; Dijkman, R.M | Human and automatic modularizations of process models to enhance their comprehension |
| 2013 | Figl, Kathrin; Koschmider, Agnes; Kriglstein, Simone | Visualising process model hierarchies |
| | *This article, found on Google Scholar, is taken from the proceedings of the 21st European Conference on Information Systems. The publication is not found as a result in the OU Library but is included in the eligibility phase because the subject seems closely related to subject of this study* | |

| | | |
|---|---|---|
| 2015 | Rompen, Tessa | The influence of modularity representation on the understandability of business process models |
| | *Although this is a master thesis and therefore not a peer-reviewed article, it is included in the set because the subject of the thesis is very close to the subject this research* | |
| 2016 | Turetken, Oktay; Rompen, Tessa M.P; Vanderfeesten, Irene ; More... | The effect of modularity representation and presentation medium on the understandability of business process models in BPMN |
| 2019 | Turetken, Oktay; Dikici, Ahmet; Vanderfeesten, Irene ; More... | The Influence of Using Collapsed Sub-processes and Groups on the Understandability of Business Process Models |
| 2019 | Baß, Julia | Investigating the Effects of Modularization in Business Process Models on the Cognitive Complexity of Humans |
| | *This document is also a master thesis of which the subject is very close to my research topic* | |
| 2020 | Winter, Michael; Pryss, Rüdiger; Probst, Thomas; Baß, Julia; Reichert, Manfred | Measuring the Cognitive Complexity in the Comprehension of Modular Process Models |
| | *The article has not yet been published. However it is included because of the following reason, stated in the article:*<br>*"This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication."* | |

## Results from the eligibility phase

After reading the complete articles, two more articles were excluded because they covered (parts of) the same research by the same authors. So, in the end six articles were included in this study. Those articles are shown in Table 26.

*Table 26 - Remaining articles after eligibility phase*

| Date | Author(s) | Title |
|---|---|---|
| 2011 | Reijers, H.A; Mendling, J; Dijkman, R.M | Human and automatic modularizations of process models to enhance their comprehension |
| 2013 | Figl, Kathrin; Koschmider, Agnes; Kriglstein, Simone | Visualising process model hierarchies |
| 2015 | Rompen, Tessa | The influence of modularity representation on the understandability of business process models |
| 2019 | Turetken, Oktay; Dikici, Ahmet; Vanderfeesten, Irene ; More... | The Influence of Using Collapsed Sub-processes and Groups on the Understandability of Business Process Models |
| 2019 | Baß, Julia | Investigating the Effects of Modularization in Business Process Models on the Cognitive Complexity of Humans |

| | | |
|---|---|---|
| 2020 | Winter, Michael; Pryss, Rüdiger; Probst, Thomas; Baß, Julia; Reichert, Manfred | Measuring the Cognitive Complexity in the Comprehension of Modular Process Models |

## Appendix B – Queries for literature selection OU Library

“process model” in ABSTRACT, rest of search terms in text:

(((Abstract:("Process model")) AND (Modulari*) AND (Understand*)) OR
((Abstract:("Process model")) AND (Decomposition) AND (Understand*)) OR
((Abstract:("Process model")) AND (“Sub process”*) AND (Understand*)) OR
((Abstract:("Process model")) AND(“Sub-process”*) AND Understand*)) OR
((Abstract:("Process model")) AND (Subprocess*) AND (Understand*))) OR
(((Abstract:("Process model")) AND (Modulari*) AND (Comprehens*)) OR
((Abstract:("Process model")) AND (Decomposition) AND (Comprehens*)) OR
((Abstract:("Process model")) AND (“Sub process”*) AND (Comprehens*)) OR
((Abstract:("Process model")) AND(“Sub-process”*) AND Comprehens*)) OR
((Abstract:("Process model")) AND (Subprocess*) AND (Comprehens*))) OR
(((Abstract:("Process model")) AND (Modulari*) AND (Expressiveness)) OR
((Abstract:("Process model")) AND (Decomposition) AND (Expressiveness)) OR
((Abstract:("Process model")) AND (“Sub process”*) AND (Expressiveness)) OR
((Abstract:("Process model")) AND(“Sub-process”*) AND Expressiveness)) OR
((Abstract:("Process model")) AND (Subprocess*) AND (Expressiveness))) OR
(((Abstract:("Process model")) AND (Modulari*) AND (Complexity)) OR
((Abstract:("Process model")) AND (Decomposition) AND (Complexity)) OR
((Abstract:("Process model")) AND (“Sub process”*) AND (Complexity)) OR
((Abstract:("Process model")) AND(“Sub-process”*) AND Complexity)) OR
((Abstract:("Process model")) AND (Subprocess*) AND (Complexity))) OR
(((Abstract:("Process model")) AND (Modulari*) AND (Difficulty)) OR
((Abstract:("Process model")) AND (Decomposition) AND (Difficulty)) OR
((Abstract:("Process model")) AND (“Sub process”*) AND (Difficulty)) OR
((Abstract:("Process model")) AND(“Sub-process”*) AND Difficulty)) OR
((Abstract:("Process model")) AND (Subprocess*) AND (Difficulty)) OR
(((Abstract:("Process model")) AND (Modulari*) AND (Readability)) OR
((Abstract:("Process model")) AND (Decomposition) AND (Readability)) OR
((Abstract:("Process model")) AND (“Sub process”*) AND (Readability)) OR
((Abstract:("Process model")) AND(“Sub-process”*) AND Readability)) OR
((Abstract:("Process model")) AND (Subprocess*) AND (Readability)) OR
(((Abstract:("Process model")) AND (Modulari*) AND (“benefits of”)) OR
((Abstract:("Process model")) AND (Decomposition) AND (“benefits of”)) OR
((Abstract:("Process model")) AND (“Sub process”*) AND (“benefits of”)) OR
((Abstract:("Process model")) AND(“Sub-process”*) AND “benefits of”)) OR
((Abstract:("Process model")) AND (Subprocess*) AND (“benefits of”)))
**Results: 337**

All search terms in ABSTRACT:

(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(Understand*))) OR
((Abstract:("Process model")) AND (Abstract:(Decomposition)) AND (Abstract:( Understand*))) OR
((Abstract:("Process model")) AND (Abstract:(“Sub process”*)) AND (Abstract:( Understand*))) OR
((Abstract:("Process model")) AND (Abstract:(Subprocess*)) AND (Abstract:( Understand*)))) OR
(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(Comprehens*))) OR
((Abstract:("Process model")) AND (Abstract:(Decomposition)) AND (Abstract:(Comprehens*))) OR
((Abstract:("Process model")) AND (Abstract:(“Sub process”*)) AND (Abstract:(Comprehens*))) OR
((Abstract:("Process model")) AND (Abstract:(Subprocess*)) AND (Abstract:(Comprehens*)))) OR

(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(Expressiveness))) OR
((Abstract:("Process model")) AND (Abstract:(Decomposition)) AND (Abstract:(Expressiveness))) OR
((Abstract:("Process model")) AND (Abstract:(“Sub process”*)) AND (Abstract:(Expressiveness))) OR
((Abstract:("Process model")) AND (Abstract:(Subprocess*)) AND (Abstract:(Expressiveness)))) OR
(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(Complexity))) OR
((Abstract:("Process model")) AND (Abstract:(Decomposition)) AND (Abstract:(Complexity))) OR
((Abstract:("Process model")) AND (Abstract:(“Sub process”*)) AND (Abstract:(Complexity))) OR
((Abstract:("Process model")) AND (Abstract:(Subprocess*)) AND (Abstract:(Complexity)))) OR
(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(Difficulty))) OR
((Abstract:("Process model")) AND (Abstract:(Decomposition)) AND (Abstract:(Difficulty))) OR
((Abstract:("Process model")) AND (Abstract:(“Sub process”*)) AND (Abstract:(Difficulty))) OR
((Abstract:("Process model")) AND (Abstract:(Subprocess*)) AND (Abstract:(Difficulty)))) OR
(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(Readability))) OR
((Abstract:("Process model")) AND (Abstract:(Decomposition)) AND (Abstract: Readability))) OR
((Abstract:("Process model")) AND (Abstract:(“Sub process”*)) AND (Abstract:(Readability))) OR
((Abstract:("Process model")) AND (Abstract:(Subprocess*)) AND (Abstract:(Readability)))) OR
(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(“benefits of”))) OR
(((Abstract:("Process model")) AND (Abstract:(Modulari*)) AND (Abstract:(“benefits of”))) OR
((Abstract:("Process model")) AND (Abstract:(Decomposition)) AND (Abstract:(“benefits of”))) OR
((Abstract:("Process model")) AND (Abstract:(“Sub process”*)) AND (Abstract:(“benefits of”))) OR
((Abstract:("Process model")) AND (Abstract:(Subprocess*)) AND (Abstract:(“benefits of”))))

**Results: 38 (2 double results)**

*Table 27 - Search results (OU Library - Abstract)*

| Date | Author(s) | Title |
|---|---|---|
| jan-19 | Turetken, Oktay; Dikici, Ahmet; Vanderfeesten, Irene ; More… | The Influence of Using Collapsed Sub-processes and Groups on the Understandability of Business Process Models |
| sep-16 | Turetken, Oktay; Rompen, Tessa M.P; Vanderfeesten, Irene ; More… | The effect of modularity representation and presentation medium on the understandability of business process models in BPMN |
| jun-13 | Zugal, Stefan; Soffer, Pnina; Haisjackl, Cornelia ; More… | Investigating expressiveness and understandability of hierarchy in declarative business process models |
| mrt-16 | Milani, Fredrik; Dumas, Marlon; Ahmed, Naved ; More… | Modeling families of business process variants: A decomposition driven method |
| 2011 | Reijers, H.A; Mendling, J; Dijkman, R.M | Human and automatic modularizations of process models to enhance their comprehension |
| 2008 | Lendek, Z; Babuška, R; De Schutter, B | Distributed Kalman filtering for cascaded systems |
| feb-16 | Milani, Fredrik; Dumas, Marlon; Matulevicius, Raimundas ; More… | Criteria and Heuristics for Business Process Model Decomposition Review and Comparative Evaluation |
| sep-19 | Nguyen, Hoang; Dumas, Marlon; ter Hofstede, Arthur H.M ; More… | Stage-based discovery of business process models from event logs |

| | | |
|---|---|---|
| 2008 | Reijers, H.A; Mendling, J; Dumas, M ; More… | Modularity in process models: review and effects |
| 2012 | Lyhne, Svend; Georgiou, Andrew; Marks, Anne ; More… | Towards an understanding of the information dynamics of the handover process in aged care settings—A prerequisite for the safe and effective use of ICT |
| okt-15 | Kalenkova, Anna A; van der Aalst, Wil M. P; Lomazova, Irina A ; More… | Process mining using BPMN: relating event logs and process models |
| jul-13 | Song, M; Yang, H; Siadat, S.H ; More… | A comparative study of dimensionality reduction techniques to enhance trace clustering performances |
| sep-18 | Debois, Søren; Debois, Søren; Hildebrandt, Thomas T ; More… | Replication, refinement & reachability: complexity in dynamic condition-response graphs |
| apr-15 | La Rosa, Marcello; Dumas, Marlon; Ekanayake, Chathura C ; More… | Detecting approximate clones in business process model repositories |
| dec-19 | Baker, Jeff; Singh, Harminder | The roots of misalignment: Insights on strategy implementation from a system dynamics perspective |
| okt-14 | Prackwieser, Christoph; Buchmann, Robert; Grossmann, Wilfried ; More… | Overcoming Heterogeneity in Business Process Modeling with Rule-Based Semantic Mappings |
| apr-10 | Kerber, Florian; van der Schaft, Arjan | Compositional analysis for linear control systems |
| dec-13 | Montali, Marco; Maggi, Fabrizio M; Chesani, Federico ; More… | Monitoring business constraints with the event calculus |
| 2017 | Ilahi, Latifa; Martinho, Ricardo; Ghannouchi, Sonia Ayachi | BPFlexTemplate: a software tool to derive flexible process model templates |
| sep-17 | Zhu, Feng; Yao, Yiping; Tang, Wenjie ; More… | A hierarchical composite framework of parallel discrete event simulation for modeling complex adaptive systems |
| sep-15 | Casola, Gioele; Yoshikawa, Satoshi; Nakanishi, Hayao ; More… | Systematic retrofitting methodology for pharmaceutical drug purification processes |
| dec-14 | García-Bañuelos, Luciano; Dumas, Marlon; La Rosa, Marcello ; More… | Controlled automated discovery of collections of business process models |
| 2019 | Obregon, Josue; Song, Minseok; Jung, Jae-Yoon | InfoFlow: Mining Information Flow Based on User Community in Social Networking Services |
| jun-12 | Eidsvik, Jo; Finley, Andrew O; Banerjee, Sudipto ; More… | Approximate Bayesian inference for large spatial datasets using predictive process models |
| 2014 | Kalenkova, A.A; Lomazova, I.A; Aalst, van der, W.M.P ; More… | Process model discovery : a method based on transition system decomposition |
| nov-17 | Bremer, Jens; Goyal, Pawan; Feng, Lihong ; More… | POD-DEIM for efficient reduction of a dynamic 2D catalytic reactor model |

| | | |
|---|---|---|
| jan-12 | Carmona, Josep; Carmona, Josep | Projection approaches to process mining using region-based techniques |
| okt-15 | Huang, Jiaqi; Li, Yupeng; Chu, Xuening ; More… | An integrated top-down design process evaluation approach of complex products and systems based on hierarchical design structure matrix |
| jul-16 | Leopold, Henrik; Mendling, Jan; Gunther, Oliver | Learning from Quality Issues of BPMN Models from Industry |
| jan-15 | Skiborowski, Mirko; Harwardt, Andreas; Marquardt, Wolfgang | Efficient optimization-based design for the separation of heterogeneous azeotropic mixtures |
| mrt-20 | Ghiasi, Shadi; Greco, Alberto; Barbieri, Riccardo ; More… | Assessing Autonomic Function from Electrodermal Activity and Heart Rate Variability During Cold-Pressor Test and Emotional Challenge |
| mei-10 | ZHANG, JIANFU; FENG, PINGFA; WU, ZHIJUN ; More… | ACTIVITY BASED CIM MODELING AND TRANSFORMATION FOR BUSINESS PROCESS SYSTEMS |
| 2011 | Ren, Qian; Banerjee, Sudipto; Finley, Andrew O ; More… | Variational Bayesian methods for spatial data analysis |
| jul-05 | Slaats, Tijs; Schunselaar, Dennis M. M; Maggi, Fabrizio M ; More… | The Semantics of Hybrid Process Models |
| sep-12 | Johannsen, Florian; Leist, Susanne | Wand and Weber's Decomposition Model in the Context of Business Process Modeling |
| jul-19 | Sonnenberg, Christoph; Bannert, Maria | Using Process Mining to examine the sustainability of instructional support: How stable are the effects of metacognitive prompting on self-regulatory behavior? |

# Appendix C – Interface strategies

In this appendix examples are provided of the four interface strategies defined by Cockburn et al. (2009) when vertical modularization is used in a process model.

## Overview+detail

The interface approach called 'overview+detail' uses a spatial separation between focused and contextual views. An example of this approach is shown in Figure 24.


*Figure 26 - Overview+detail (Power Point)*

## Focus+context

The 'focus+context' strategy aims at minimizing the seam between views by displaying the focus within the context. An example of the 'focus+context' approach is given in Figure 27.

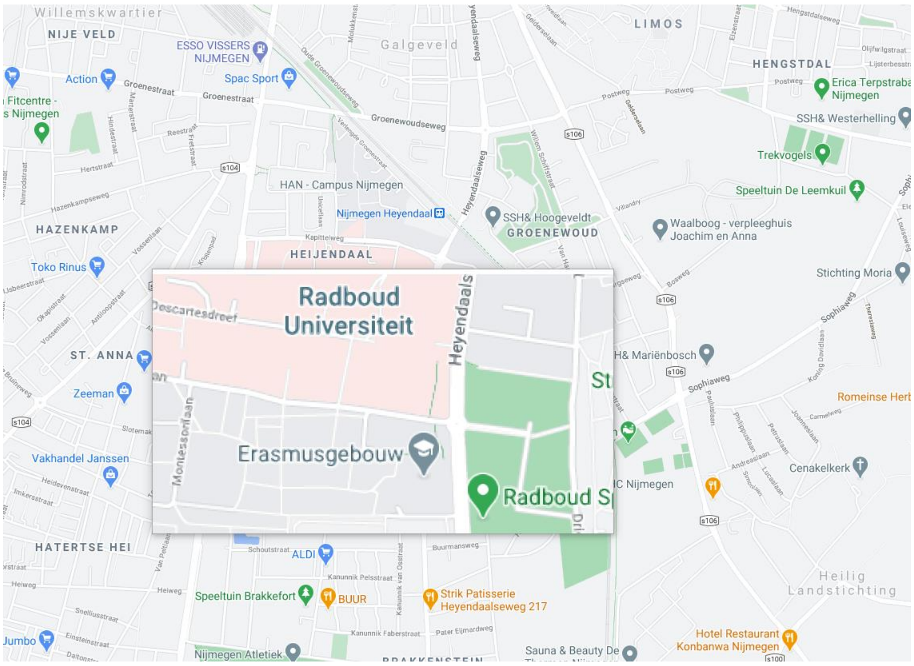


*Figure 27 - Focus+context (Google Maps)*

## Zooming

The approaches 'zooming', uses a separation between the views of the context and the focus. A user sees only one view, focus or context, at a time.

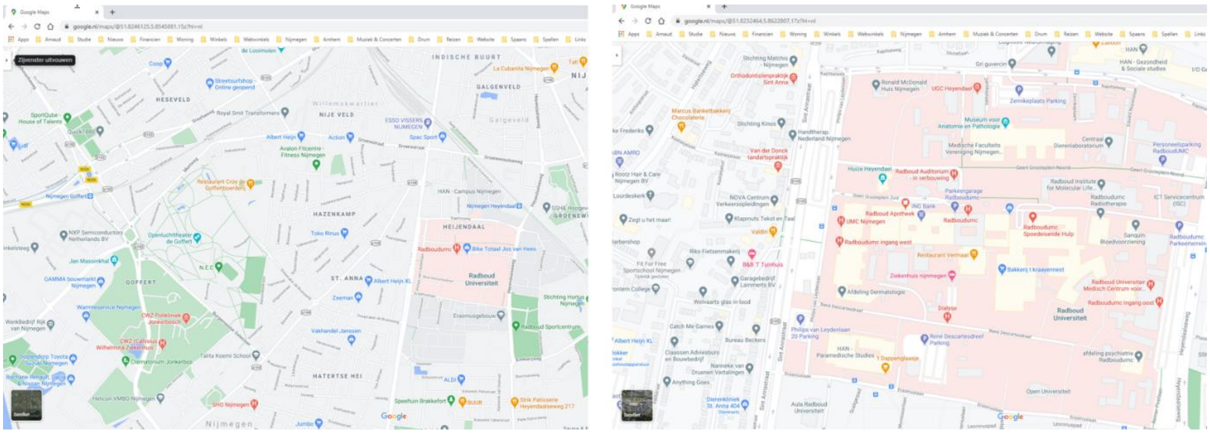

*Figure 28 - Zooming (two separate screens) (Google Maps)*

## Cue-based

The ‘cue-based’ strategy changes the view of a model to highlight or suppress items, see Figure 29.

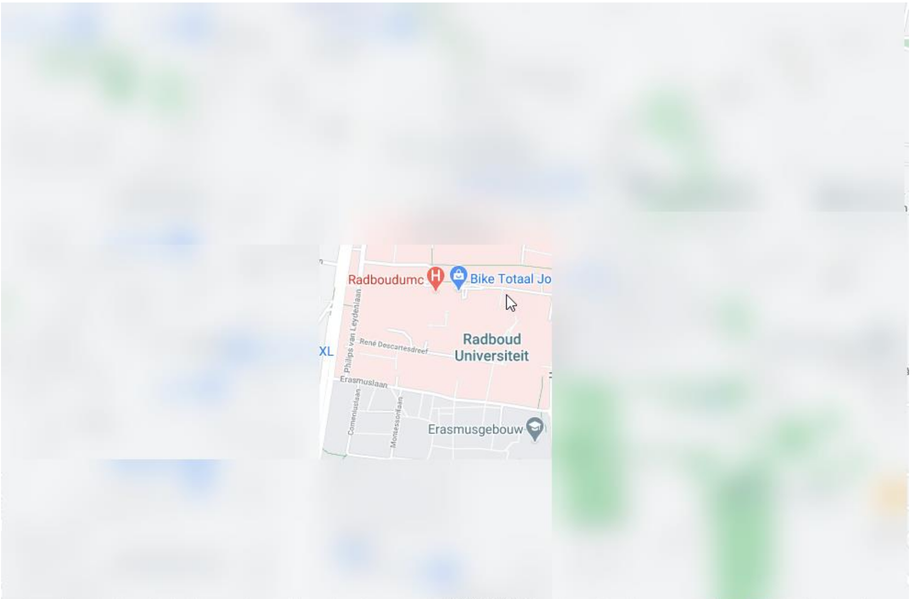


*Figure 29 - Cue-based (Blurring) (Google Maps)*

# Appendix D – Design of questionnaire and experiment

## Questionnaire

The questionnaire will be presented to participants online prior to the actual experiment. They will be asked if and how often they use process modeling and asked to assess their theoretical knowledge. Their educational background will also be asked. The questions and possible answers, see Table 28, were taken from the study by Reijers and Mendling (2011). Because the questionnaire are presented to Dutch speaking participants, the questions will be translated into Dutch.

*Table 28 - Questionnaire*

| **Question** | **Possible Answers** |
|---|---|
| How often do you use business process modeling in practice? | ○ I never use business process modeling in practice<br>○ I sometimes use business process modeling in practice<br>○ I regularly use business process modeling in practice, but not every day<br>○ I use business process modeling in practice every day |
| How do you assess your theoretical knowledge on business process modeling? | ○ I have no theoretical knowledge<br>○ I have rather weak theoretical knowledge<br>○ I have mediocre theoretical knowledge<br>○ I have rather strong theoretical knowledge<br>○ I have strong theoretical knowledge |
| What is you highest level of education? | ○ Secondary School (Middelbare school)<br>○ Post-secondary vocational education (MBO)<br>○ Higher vocational education (HBO)<br>○ University education (WO)<br>○ Postgraduate (Postdoctoraal)<br>○ Other |

These ratings are subjective but will provide some insight into how much knowledge of and experience with process models and process modeling the participant has got. It was decided not to carry out a competency test because it might be time-consuming and could also cause participants to drop out. Moreover, this experiment is not about understanding the exact syntax of BPMN, but about the difference between representing sub-processes.

## Experiment

### Process model

The experiment will make use of a fictitious process. This is done to ensure that familiarity with the process does not play a role in the outcome of the experiment. The model will contain vertical modularity. Both the 'overview+detail' and 'focus+context' interface strategy will be used to unfold the sub-processes. The model is based on Model L from Reijers and Mendling (2011). The number of activities and the number of nodes are taken into account when creating the model. By using vertical modularization, the number of activities is reduced from 76 to 30 and the number of nodes from 101 to 44. This meets the guidelines given by Sanchez-Gonzalez et al. (2017) in terms of understandability regarding the number of activities, no more than 31, and nodes, no more than 58.

It is also chosen to designate processes with letters and numbers rather than text, as was also done by Turetken et al. (2019). The reason for using abstract letters and numbers is to avoid confusing participants with texts that would mean nothing in the fictional model.

The model used in the experiment, see Figure 30, will contain two hierarchical levels. This means the model consists of a main level, with collapsed sub-processes, and one level deeper in which the content of the sub-processes are shown. The process model contains six collapsed sub-processes.

No other indicators were taken into account.

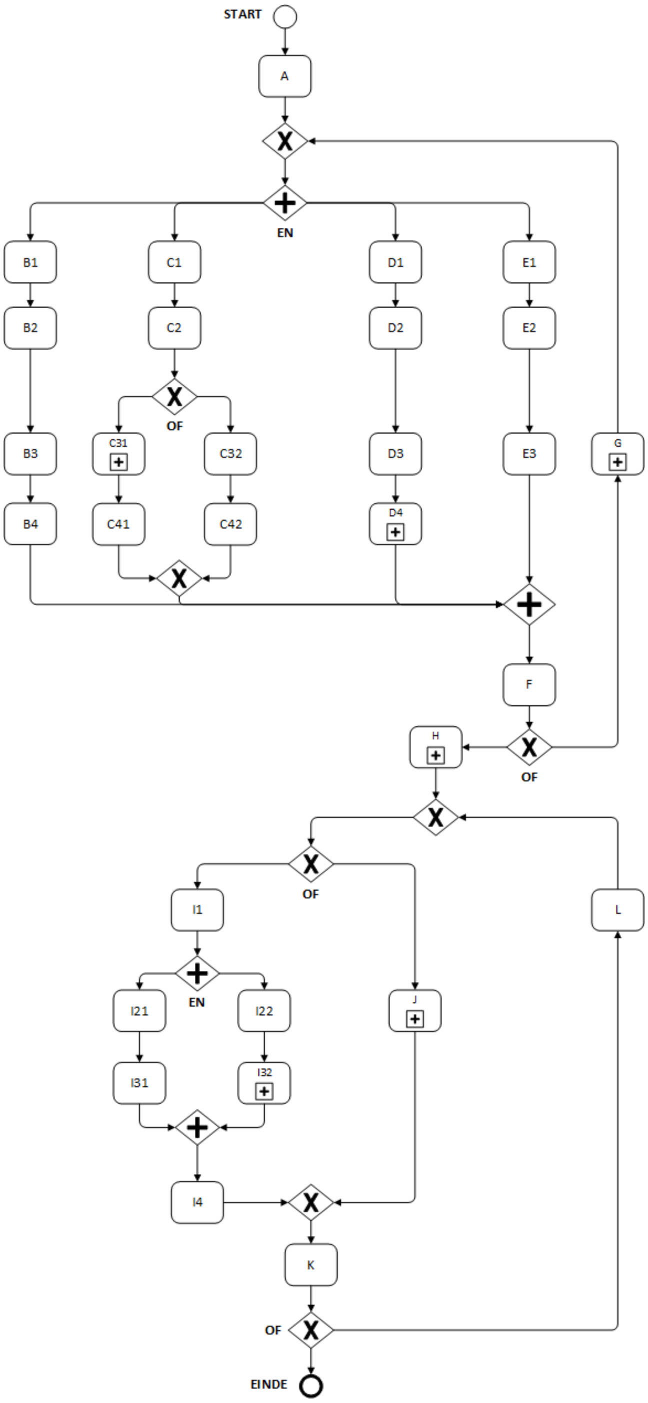


*Figure 30 - Fictitious process model used in experiment*

## Sub-processes

To be able to answer the questions, participants must use the models with unfolded sub-processes. In the next sections, the models are depicted using the 'overview+detail' as well as the 'focus+context' interface strategy.

## Sub-processes using 'overview+detail' interface strategy

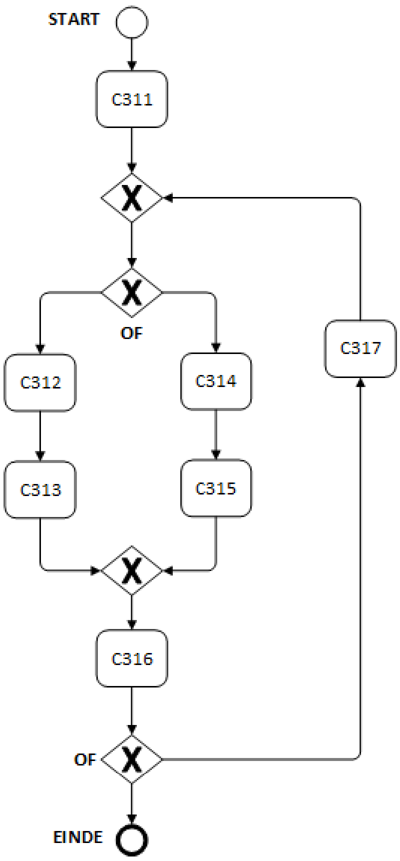


*Figure 31 - Sub-processes C31 ('overview+detail')*

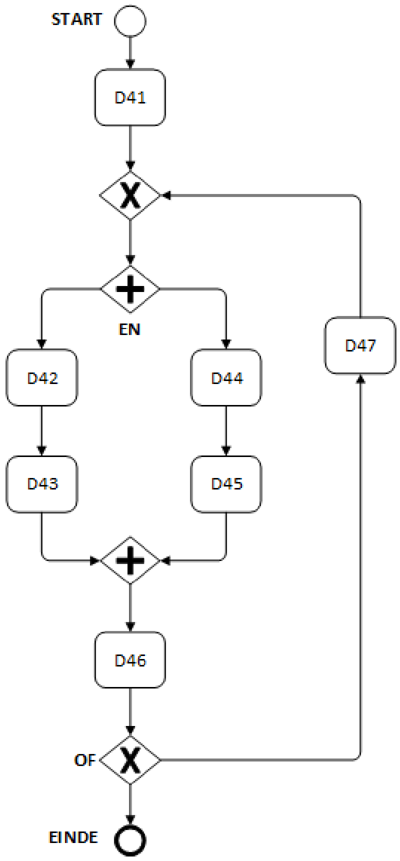


*Figure 32 - Sub-processes D4 ('overview+detail')*

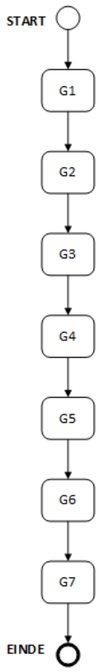


*Figure 33 - Sub-processes G ('overview+detail')*

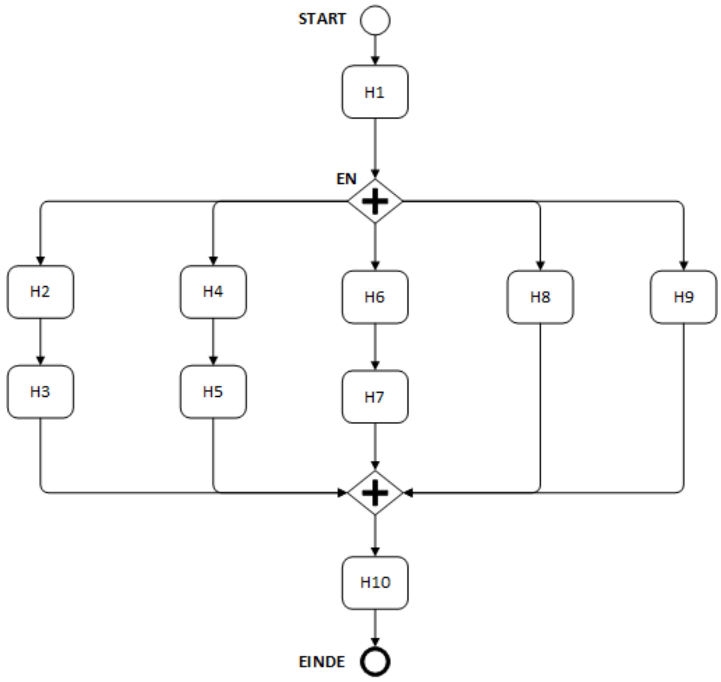


*Figure 34 - Sub-processes H ('overview+detail')*

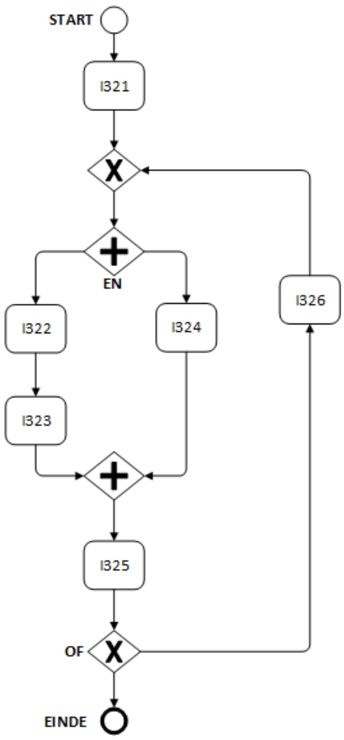


*Figure 35 - Sub-processes I32 ('overview+detail')*

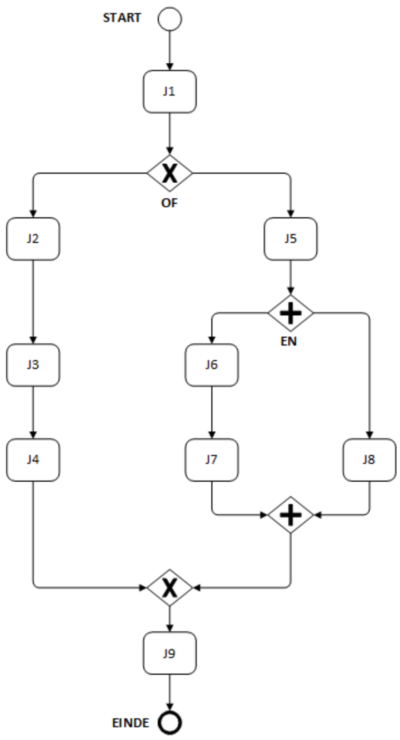


*Figure 36 - Sub-processes J ('overview+detail')*

# Sub-processes using 'focus+context' interface strategy

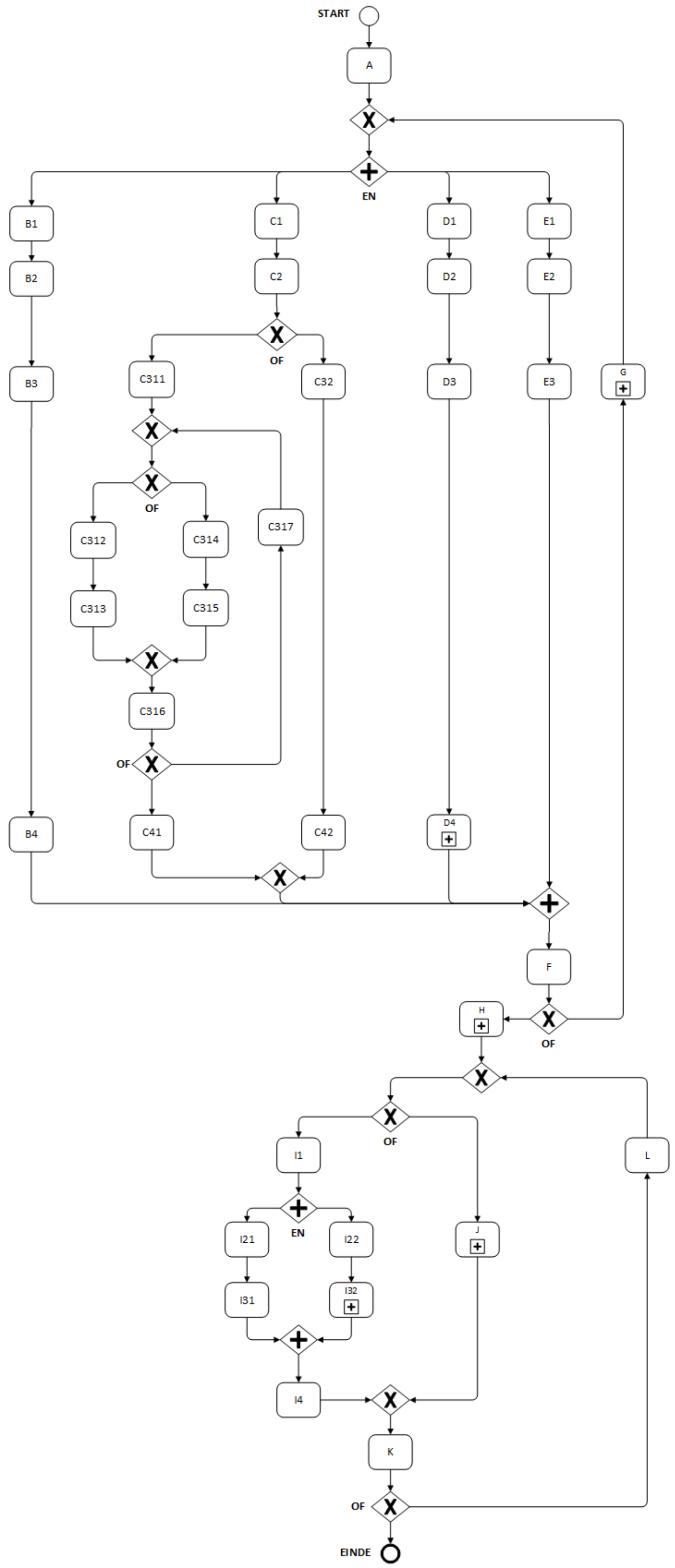


*Figure 37 - Sub-processes C31 ('focus+context')*

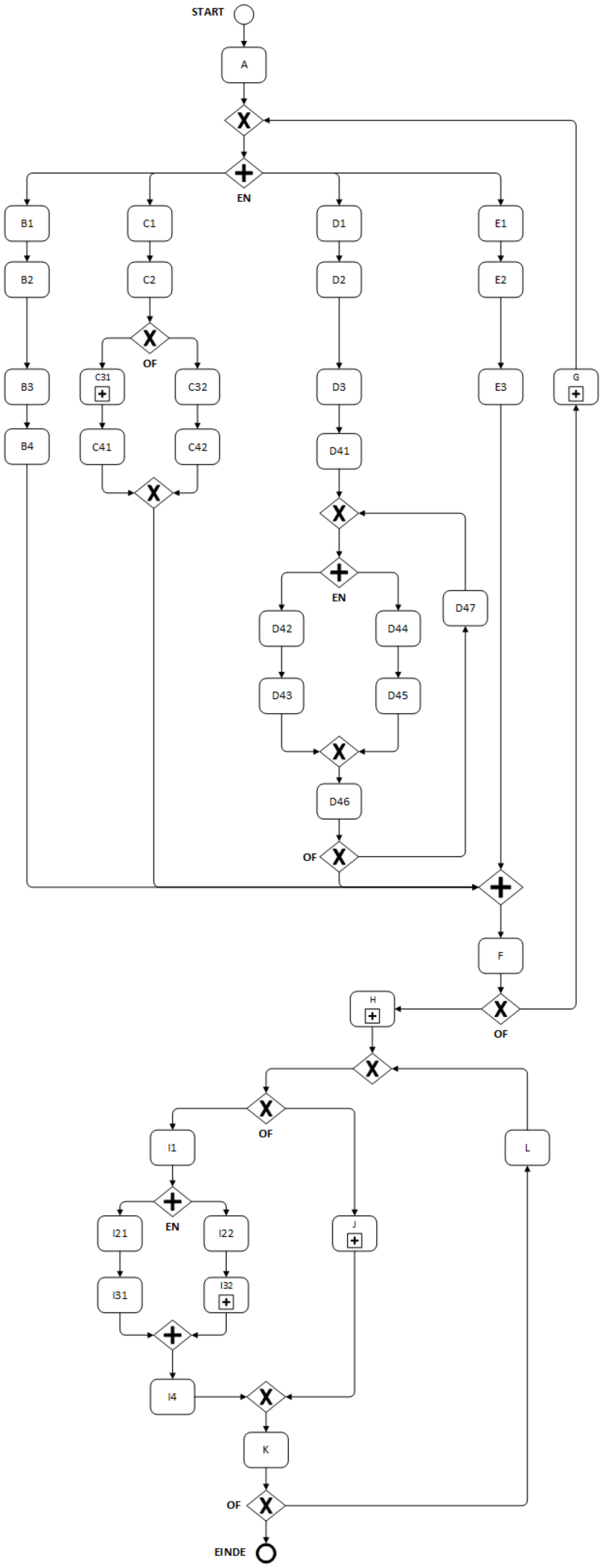


*Figure 38 - Sub-processes D4 ('focus+context')*

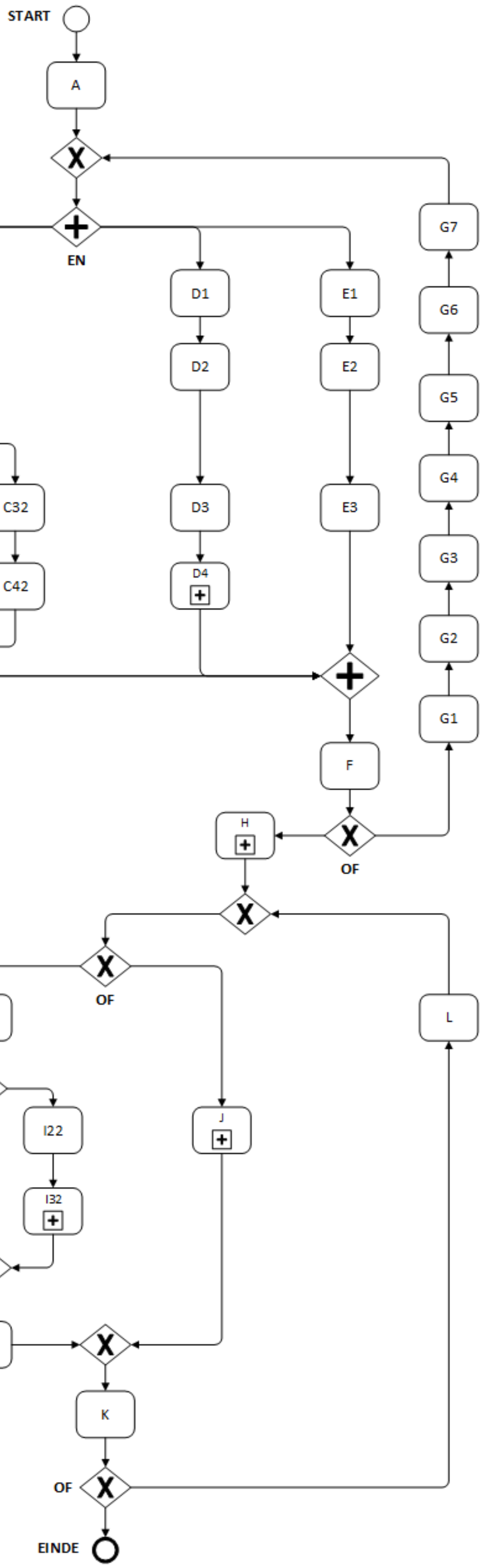


*Figure 39 - Sub-processes G ('focus+context')*

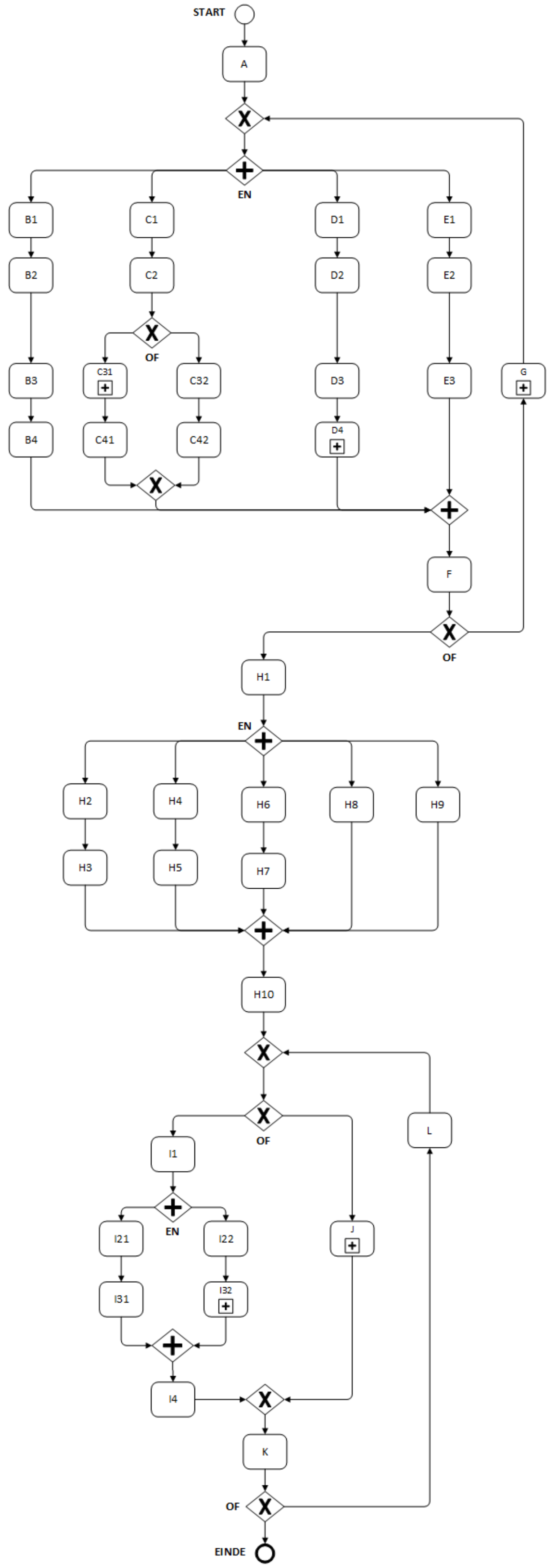


*Figure 40 - Sub-processes H ('focus+context')*

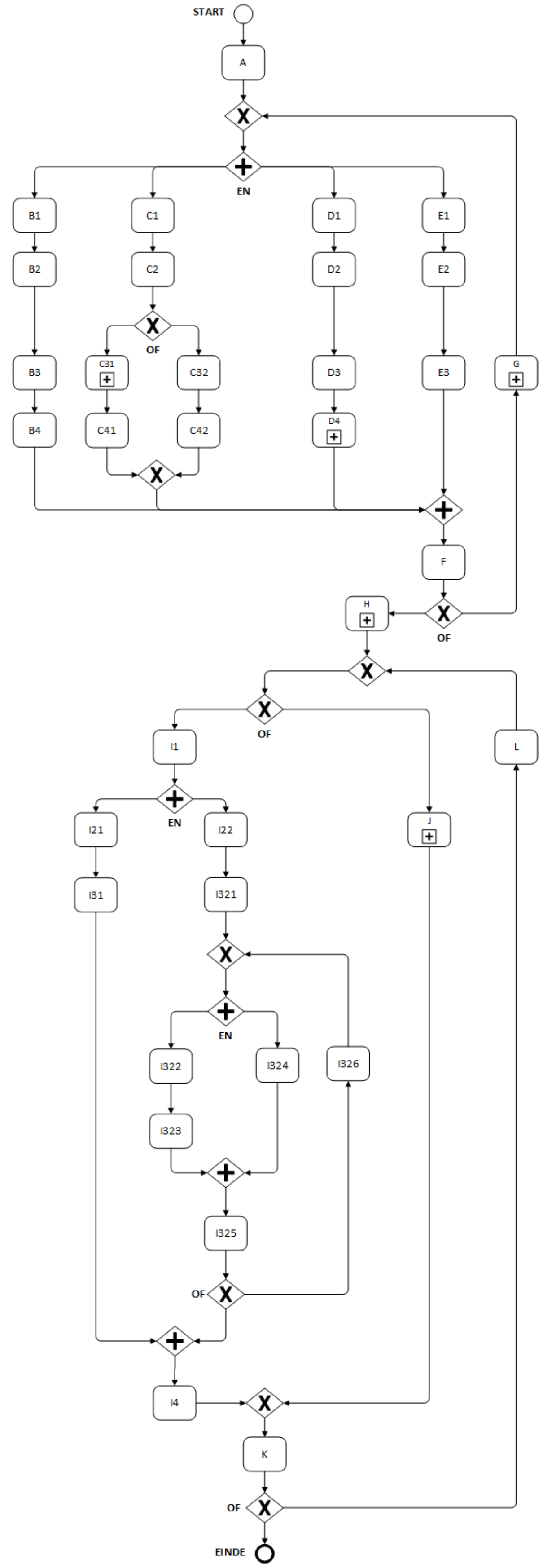


*Figure 41 - Sub-processes I32 ('focus+context')*

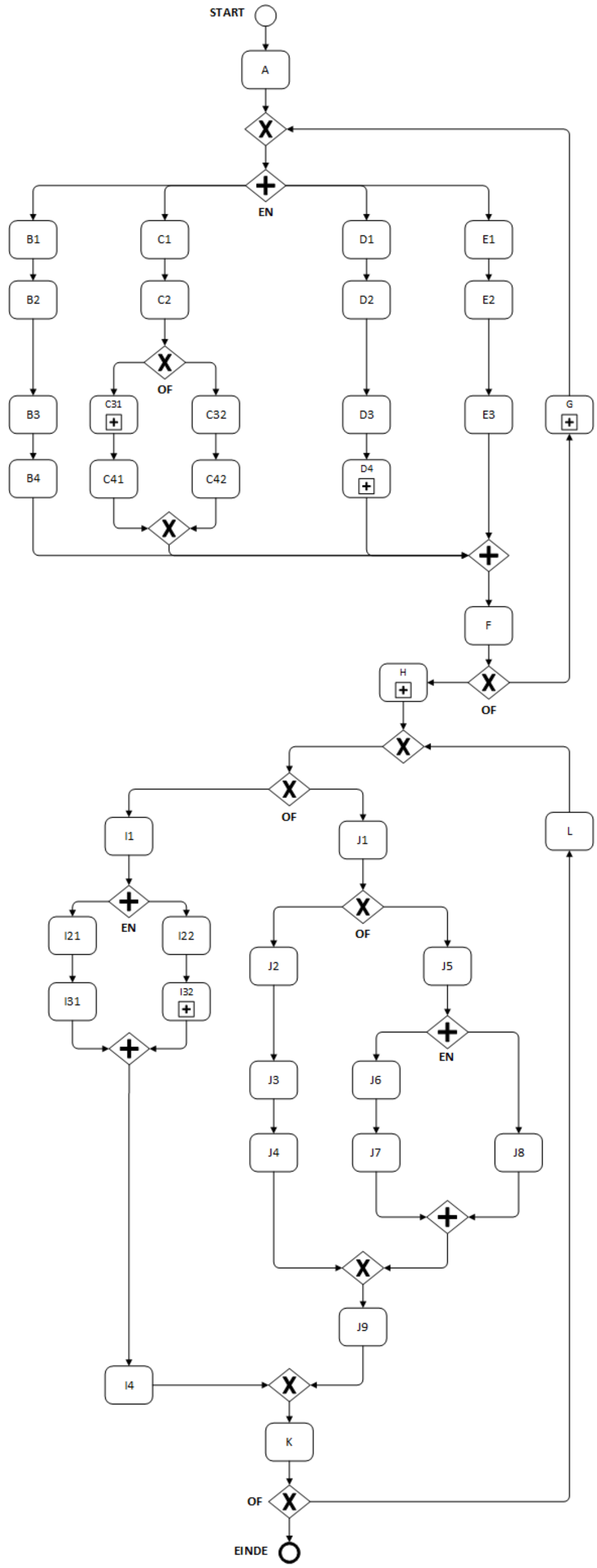


*Figure 42 - Sub-processes J ('focus+context')*

## Questions

Two sets of questions were prepared. One set for the ‘overview+detail’ and another set for the ‘focus+context’ interface strategy. Although the questions differ per interface strategy, care was taken to ensure that the degree of difficulty of the questions was the same. This way, the results of both strategies can be compared.

The questions had to be drawn up carefully. For some questions overview of the whole process is needed (global). For others only information of a specific part of the process represented in a sub-process is needed (local). Factors, such as concurrency, exclusiveness, order and repetition as mentioned by H. Reijers et al. (2011) were also taken into account in drafting the questions.

The questions are distributed among the two question sets, ‘overview+detail’ and ‘focus+context’, in such a way that both sets consist of the same number of global and local questions. The aforementioned factors are also equally distributed across the question sets. This distribution ensures that both question sets consist of similar questions, making the results comparable.

Because all participants will be Dutch speaking people, the questions, as shown in Table 29, will be translated into Dutch for use in the experiment.

*Table 29 - Questions per interface strategy*

| | **Questions ‘overview+detail’** | **Answer** | **Global/ Local** | **Concurrency/Exclusiveness/ Order/Repetition** |
|---|---|---|---|---|
| 1 | If the process has gone through process step B3, does the process then always also go through process step G3? | No | Global | Order |
| 2 | If the process passes through process step C31, does it always also pass through process step C317? | No | Local | Order / Exclusiveness |
| 3 | Can the process pass through process step D47 more than once? | Yes | Local | Repetition |
| 4 | Can process steps H2 and H7 be executed simultaneously? | Yes | Local | Concurrency |
| 5 | If the process has passed through process step J2, can it pass through process step I324 again? | Yes | Global | Repetition |
| 6 | Can process steps E3 and I324 be performed simultaneously? | No | Global | Concurrency |
| | **Questions ‘focus+context’** | **Answer** | **Global/ Local** | **Concurrency/Exclusiveness/ Order/Repetition** |
| 1 | If process step G5 has been passed, does the process always pass through process step B3? | Yes | Global | Order |
| 2 | If the process enters step D4, does it always go through process step D47? | No | Local | Order / Exclusiveness |
| 3 | Can the process pass through process step C317 multiple times? | Yes | Local | Repetition |

| | | | | |
|---|---|---|---|---|
| 4 | Can process steps H2 and H10 be executed simultaneously? | No | Local | Concurrency |
| 5 | If the process has gone through process step I322, can it also go through process step J5 ? | Yes | Global | Repetition |
| 6 | Can process steps J7 and C1 be performed simultaneously? | No | Global | Concurrency |

As described in section 3.1, the first group of participants is first presented with the questions using models with the 'overview+detail' interface strategy and subsequently with the questions using models with the 'focus+context' interface strategy. The second group of participants will be presented with the questions in reverse order.

When answering the questions, the participants will be presented with the complete process model with all sub-processes collapsed on the left side of the screen. On the right side of the screen the sub-process will be unfolded using the 'overview+detail' interface strategy or 'focus+context' interface strategy depending on the question. Both models are shown simultaneously, due to limitations of the LimeSurvey tool.

An example of question 1 of the ‘overview+detail’ interface strategy is shown in Figure 43.

Question 1: If the process has gone through process step B3, does the process then always also go through process step G3?

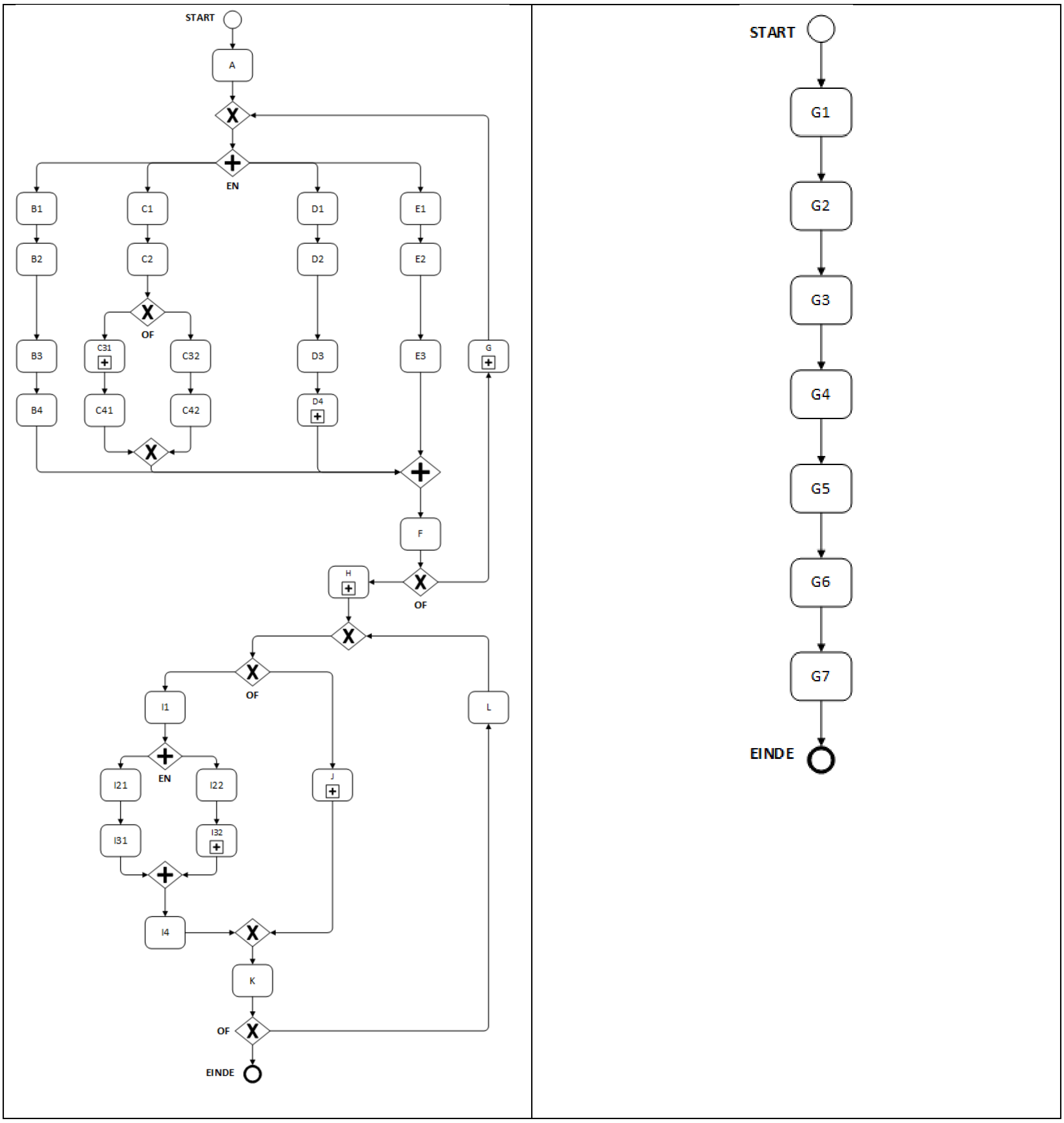


*Figure 43 - Question 1 'overview+detail'*

Figure 44 shows the models which are presented to the participants while answering question 1 of the ‘focus+context’ interface strategy.

Question 1: If process step G5 has been passed, does the process always pass through process step B3?

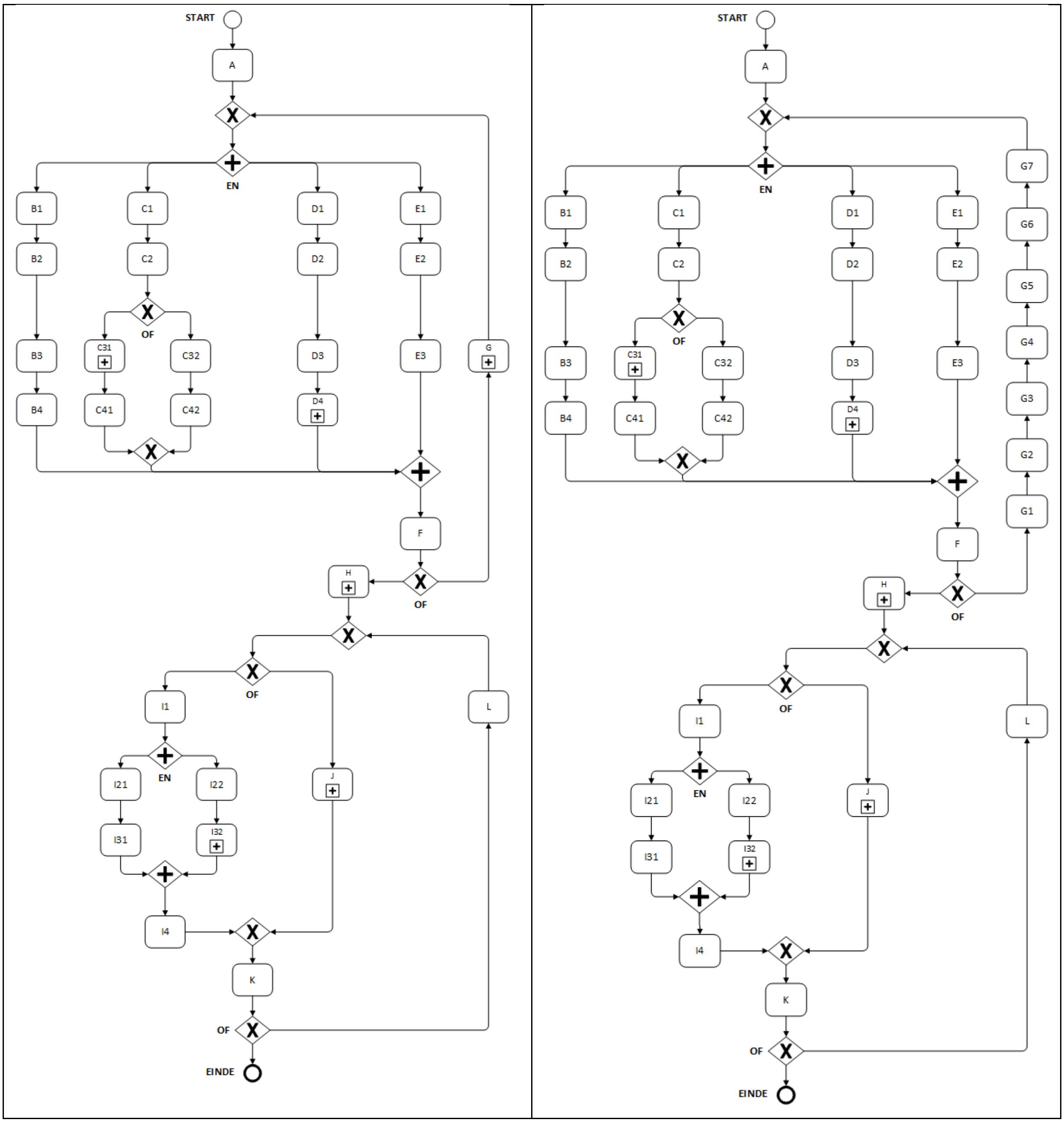


*Figure 44 - Question 1 'focus+context'*

# Appendix E – Learning Effect

The data from the two participant groups were analyzed to see if there was a learning effect. Participant group 1 was first presented with the set of questions about the models using the 'overview+detail' interface strategy and then with questions about the models using the 'focus+context' interface strategy. For participant group 2, the order of the questions was reversed.

The analysis was done by comparing the number of correct answers ('task effectiveness') and the number of correct answers per hour ('task efficiency') for each interface strategy within each group as well as between both groups.

## Task effectiveness

Figure 45 shows the histogram of the number of correct answers for the 'overview+detail' strategy per participant group, whereas Figure 46 shows the histogram of the number of correct answers for the 'focus+context' strategy per participant group.

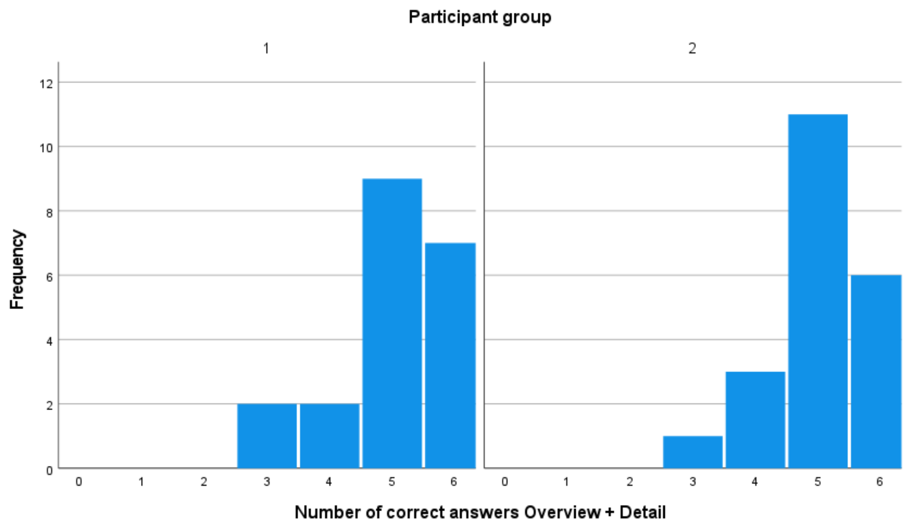


*Figure 45 - Histogram of the number of correct answers for the 'overview+detail' strategy*

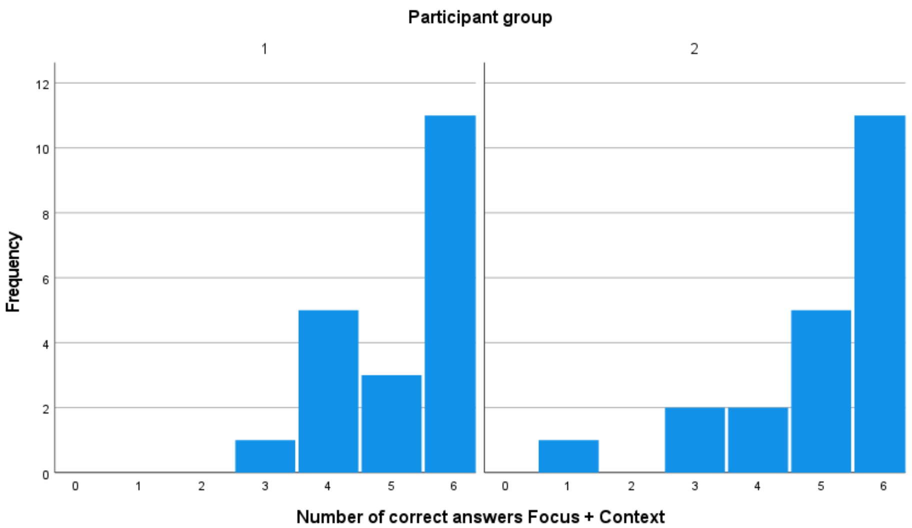


*Figure 46 - Histogram of the number of correct answers for the 'focus+context' strategy*

The statistics of the number of correct answers are presented in Table 30.

*Table 30 - Number of correct answers per participant group*

| Participant group | | Number of correct answers ('overview+detail') | Number of correct answers ('focus+context') |
|---|---|---|---|
| 1 | Mean | 5,05 | 5,20 |
| | N | 20 | 20 |
| | Std. Deviation | 0,945 | 1,005 |
| 2 | Mean | 5,05 | 5,05 |
| | N | 21 | 21 |
| | Std. Deviation | 0,805 | 1,359 |
| Total | Mean | 5,05 | 5,12 |
| | N | 41 | 41 |
| | Std. Deviation | 0,865 | 1,187 |

The mean of correct answers for participant group 1 for the questions with the 'overview+detail' strategy, which they got as the first set, is 5,05. This is exactly the same as for participant group 2, who got the 'overview+detail' interface strategy as the second set.

For the questions with 'focus+context' participant group 1 scored a mean of 5,20. They got these questions as the second set and scored slightly better than participant group 2 (mean = 5,05), who got the 'focus+context' questions as the first set of questions.

To test whether the order in which the questions were presented to the participants caused a significant difference in the results per interface strategy, the Independent-Samples median test was used, as the participants are divided into two separate groups. The results of the test are shown in Table 31 and Figure 47.

*Table 31 - Independent-Samples median test task effectiveness (participant group)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers 'overview+detail' are the same across categories of participant group. | Independent-Samples Median Test | 0,915[c] | Retain the null hypothesis. |
| The medians of number of correct answers 'focus+context' are the same across categories of participant group. | Independent-Samples Median Test | [d] | Unable to compute. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.
c. Yates's Continuity Corrected Asymptotic Sig.
d. All test fields values are less than or equal to the median.

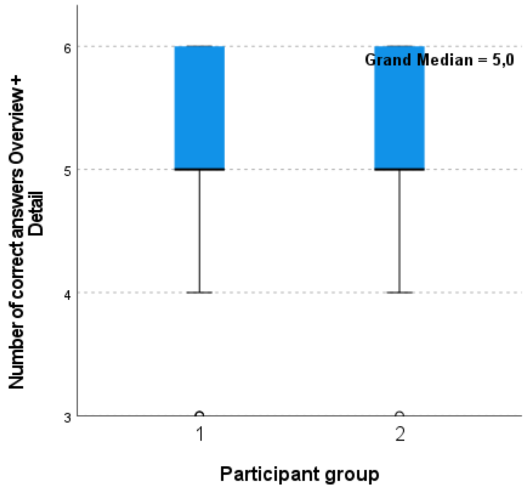


*Figure 47 - Test task effectiveness 'overview+detail'*

For the ‘focus+context’ strategy it was not possible to carry out the test. For the questions with the ‘overview+detail’ strategy, no significant difference (p = 0,915) was found between the two participant groups.

Within the participant groups, it can be seen that participant group 1 scored less on the ‘overview+detail’ strategy (mean = 5,05) questions, which they got as the first set, than on the ‘focus+context’ questions (mean = 5,20), which they got as the second set of questions. Participant group 2 made the ‘focus+context’ (mean = 5,05) questions, which they got as the first set, on average as good as the ‘overview+detail’ questions (mean = 5,05), which they got as the second set.

Although a slight difference could be distinguished for the results between both groups when looking at the ‘focus+context’ questions and between both sets of questions within participant group 1, no statistical significant difference could be established.

## Task efficiency

Figure 48 shows the number of correct answers per hour for the 'overview+detail' strategy per participant group. The number of correct answers for the 'focus+context' strategy per participant group is given in Figure 49.

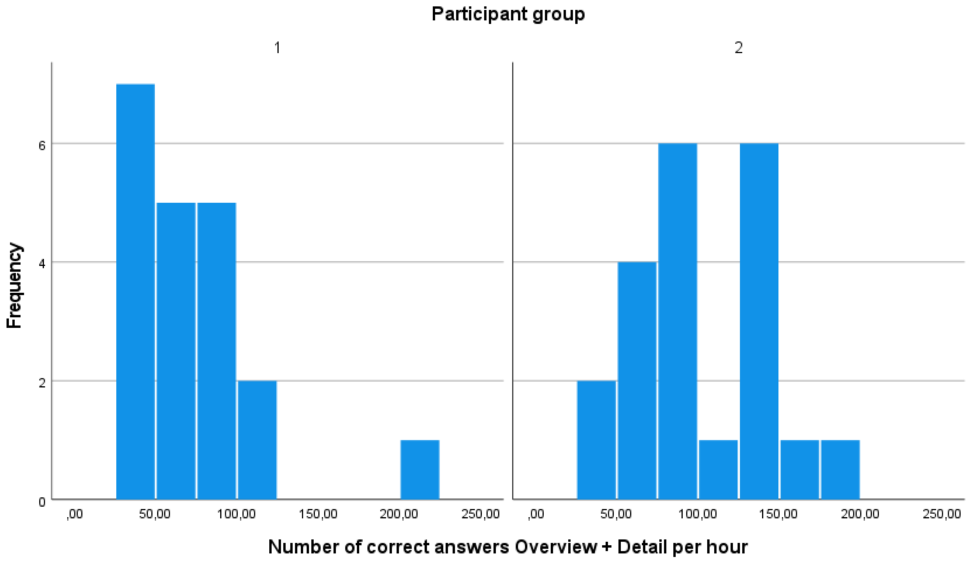


*Figure 48 - Histogram of the number of correct answers per hour 'overview+detail' strategy*

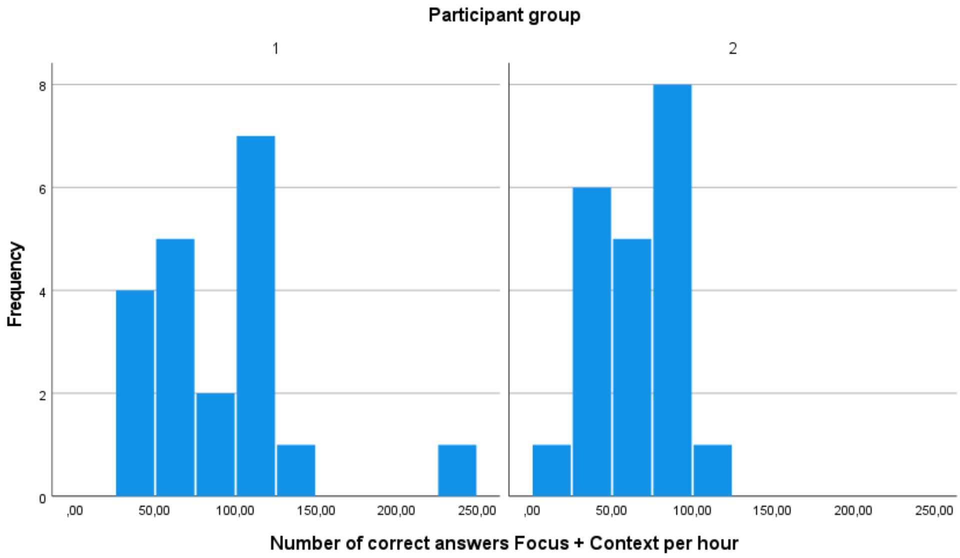


*Figure 49 - Histogram of the number of correct answers per hour 'focus+context' strategy*

Table 32 shows the statistics of the number of correct answers per hour given by participant group 1 and 2.

*Table 32 - Number of correct answers per hour per participant group*

| Participant group | | Number of correct answers per hour ('overview+detail') | Number of correct answers per hour ('focus+context') |
|---|---|---|---|
| 1 | Mean | 72,05 | 87,94 |
| | N | 20 | 20 |
| | Std. Deviation | 41,240 | 63,019 |
| 2 | Mean | 102,17 | 63,02 |
| | N | 21 | 21 |
| | Std. Deviation | 41,847 | 24,724 |
| Total | Mean | 87,48 | 75,18 |
| | N | 41 | 41 |
| | Std. Deviation | 43,769 | 38,299 |

Participant group 1 answered fewer questions correctly on average in an hour (mean = 72,05) for the questions with the 'overview+detail' strategy, which they got as the first set, than participant group 2 (mean = 102,17), who got the 'overview+detail' interface strategy as the second set.

For the questions with 'focus+context' participant group 1 scored a mean of 87,94 correct answers per hour. They got these questions as the second set and scored better on average than participant group 2 (mean = 63,02), who got the 'focus+context' questions as the first set.

To see whether the order in which the questions were presented to the participants caused a significant difference in the 'task efficiency' per interface strategy, the Independent-Samples median test was used. Table 33 and Figure 50 and Figure 51 show the results of the tests.

*Table 33 - Independent-Samples median test task efficiency (participant group)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers per hour 'overview+detail' are the same across categories of participant group. | Independent-Samples Median Test | 0,042[c] | Reject the null hypothesis. |
| The medians of number of correct answers per hour 'focus+context' are the same across categories of participant group. | Independent-Samples Median Test | 0,642[c] | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.
c. Yates's Continuity Corrected Asymptotic Sig.

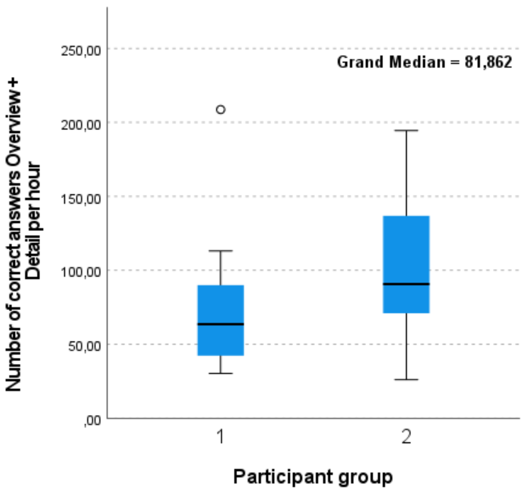


Figure 50 - Test task efficiency 'overview+detail'

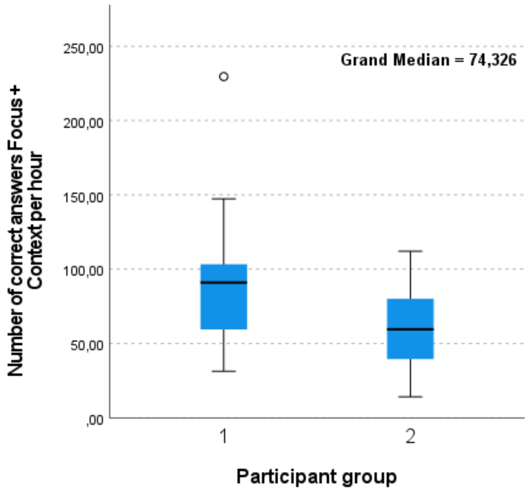


Figure 51 - Test task efficiency 'focus+context'

For the 'focus+context' strategy, no significant difference (p = 0,642) could be established. However, for the questions with the 'overview+detail' strategy, a significant difference (p = 0,042), although small, was found between the two participant groups with regard to 'task efficiency'.

Participant group 1 scored less on the 'overview+detail' strategy (mean = 72,05) questions, which they got as the first set, than on the 'focus+context' questions (mean = 87,94), which they got as the second set of questions. Participant group 2 answered fewer questions correctly per hour for the 'focus+context' strategy (mean = 63,02), which they got as the first set, than for the 'overview+detail' questions (mean = 102,17), which they got as the second set.

Although not statistically substantiated, both participant groups got a better score (mean) on the second set of questions. It follows from the Independent-Samples median test that participant group 2 scored significantly better (p = 0,042) than participant group 1 on questions with the 'overview+detail' interface strategy. Participant group 2 received these questions as second set while participant group got these questions as first set. It can therefore be cautiously concluded that a learning effect occurred.

# Appendix F – Confounding variables

This appendix presents the results of the self-assessment of the three confounding variables, 'theoretical modeling competency', 'practical modeling competency' and 'education'.

## Theoretical modeling competency

The distribution of participants across the different levels of theoretical knowledge is shown in Table 34. None of the participants indicated that he/she has a strong theoretical knowledge.

*Table 34 - Theoretical knowledge*

| | **Frequency** | **Percentage** | **Cumulative Percentage** |
|---|---|---|---|
| I have no theoretical knowledge | 12 | 29,3 | 29,3 |
| I have rather weak theoretical knowledge | 10 | 24,4 | 53,7 |
| I have mediocre theoretical knowledge | 13 | 31,7 | 85,4 |
| I have rather strong theoretical knowledge | 6 | 14,6 | 100,0 |
| I have strong theoretical knowledge | 0 | 0 | 100,0 |
| Total | 41 | 100,0 | |

## Practical modeling competency

The participants were asked to self-assess their use of process models in daily practice. This way the confounding variable 'practical modeling competency' was operationalized. Table 35 shows the results of this self-assessment.

*Table 35 - Practical experience*

| | **Frequency** | **Percentage** | **Cumulative Percentage** |
|---|---|---|---|
| I never use business process modeling in practice | 15 | 36,6 | 36,6 |
| I sometimes use business process modeling in practice | 17 | 41,5 | 78,0 |
| I regularly use business process modeling in practice, but not every day | 7 | 17,1 | 95,1 |
| I use business process modeling in practice every day | 2 | 4,9 | 100,0 |
| Total | 41 | 100,0 | |

## Education

The confounding variable 'education' was operationalized by asking the participants after their highest education. The results are shown in Table 36.

*Table 36 - Highest education*

| | Frequency | Percentage | Cumulative Percentage |
|---|---|---|---|
| Secondary School (Middelbare school) | 1 | 2,4 | 2,4 |
| Post-secondary Vocational Education (MBO) | 1 | 2,4 | 4,9 |
| Higher Vocational Education (HBO) | 12 | 29,3 | 34,1 |
| University Education (WO) | 20 | 48,8 | 82,9 |
| Postgraduate (Postdoctoraal) | 5 | 12,2 | 95,1 |
| Other (Overige) | 2 | 4,9 | 100,0 |
| Total | 41 | 100,0 | |

# Appendix G – Analysis of interface strategy

This appendix shows the results of the non-parametric Wilcoxon signed-rank test for both the number of correct answers ('task effectiveness') as well as the number of correct answers per hour provided by the participants ('task efficiency'). The results are obtained using SPSS.

## Task effectiveness

The Wilcoxon signed-rank test is used to check whether there is a significant difference in the result for the number of correct answers. Because the value p = 0,453 is greater than the critical value of 0,050 the null hypothesis of the Wilcoxon signed-rank test is retained, see Table 37. Thus, it is concluded that there is no significant difference in number of correct answers between the 'overview+detail' and 'focus+context' interface strategies.

*Table 37 - Hypothesis test summary for task effectiveness*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The median of differences between number of correct answers 'overview+detail' and number of correct answers 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,453 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

The result of the test is shown in Table 38.

*Table 38 - Wilcoxon signed-rank test summary for task effectiveness*

**Related-Samples Wilcoxon Signed Rank Test Summary**

| | |
|---|---|
| Total N | 41 |
| Test Statistic | 161,000 |
| Standard Error | 30,627 |
| Standardized test Statistic | 0,751 |
| Asymptotic Sig. (2-sided test) | 0,453 |

Figure 52 shows the result of the Wilcoxon signed rank test.

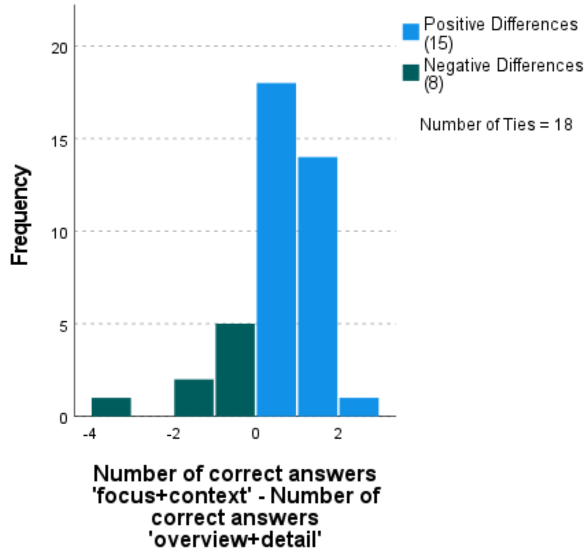


*Figure 52 - Wilcoxon signed-rank test for task effectiveness*

## Task efficiency

The number of correct answers per hour is analyzed to check whether there is a significant difference between both interface strategies. The null hypothesis is not rejected, because the significance value of p = 0,236 is greater than the critical value of 0,050. Therefore, no significant difference between the interface strategies is detected with regard to 'task efficiency'. The following tables and figure show the test summaries of the Wilcoxon signed-rank test.

*Table 39 - Hypothesis test summary for task efficiency*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The median of differences between number of correct answers per hour 'overview+detail' and number of correct answers per hour 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,236 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 40 - Wilcoxon signed-rank test summary for task efficiency*

**Related-Samples Wilcoxon Signed Rank Test Summary**

| | |
|---|---|
| Total N | 41 |
| Test Statistic | 339,000 |
| Standard Error | 77,170 |
| Standardized test Statistic | -1,186 |
| Asymptotic Sig. (2-sided test) | 0,236 |

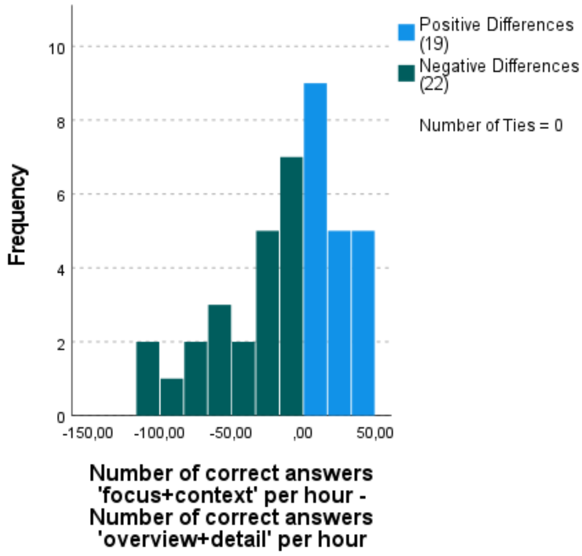


*Figure 53 - Wilcoxon signed-rank test for task efficiency*

# Appendix H – Analysis of confounding variables

Three confounding variables, 'theoretical modeling competency', 'practical modeling competency' and 'education', were defined and analyzed. This appendix shows the results of testing whether any of these variables has a significant impact on the understanding of the process models. This understandability is operationalized by number of correct answers ('task effectiveness') as well as number of correct answers per hour ('task efficiency').

## Theoretical modeling competency

To see whether the theoretical knowledge of participants causes a significant difference in the results per interface strategy, the Independent-Samples median test is used. This test is used because the participants are divided into four separate groups based on the self-assessed theoretical knowledge. The participants were offered five answers, none of them chose the option 'I have strong theoretical knowledge', as shown in Appendix F.

### Task effectiveness

The statistics related to theoretical knowledge are shown in Table 41.

*Table 41 - Case summaries number of correct answers (theoretical knowledge)*

| **Theoretical knowledge** | | **Number of correct answers 'overview+detail'** | **Number of correct answers 'focus+context'** |
|---|---|---|---|
| I have no theoretical knowledge | Mean | 4,75 | 4,25 |
| | Median | 5,00 | 4,00 |
| | Minimum | 3 | 3 |
| | Maximum | 6 | 6 |
| I have rather weak theoretical knowledge | Mean | 5,10 | 5,30 |
| | Median | 5,00 | 6,00 |
| | Minimum | 5 | 1 |
| | Maximum | 6 | 6 |
| I have mediocre theoretical knowledge | Mean | 5,08 | 5,62 |
| | Median | 5,00 | 6,00 |
| | Minimum | 3 | 4 |
| | Maximum | 6 | 6 |
| I have rather strong theoretical knowledge | Mean | 5,50 | 5,50 |
| | Median | 5,50 | 5,50 |
| | Minimum | 5 | 5 |
| | Maximum | 6 | 6 |
| Total | Mean | 5,05 | 5,12 |
| | Median | 5,00 | 6,00 |
| | Minimum | 3 | 1 |
| | Maximum | 6 | 6 |

The Independent-Samples median test is carried out to check if theoretical knowledge has a significant impact on the 'task effectiveness' per interface strategy.

The test summary of the Independent-Samples median test for the 'overview+detail' strategy is shown in Table 42, Table 43 and Figure 54.

*Table 42 - Hypothesis test task effectiveness (theoretical knowledge, 'overview+detail')*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers 'overview+detail' are the same across categories of theoretical knowledge. | Independent-Samples Median Test | 0,203 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 43 - Indep.-Samples median test task effectiveness (theoretical knowledge, 'overview+detail')*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 5,000 |
| Test Statistic | 4,606[a,b] |
| Degree Of Freedom | 3 |
| Asymptotic Sig. (2-sided test) | 0,203 |

a. More than 20% of the cells have expected values less than five.
b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

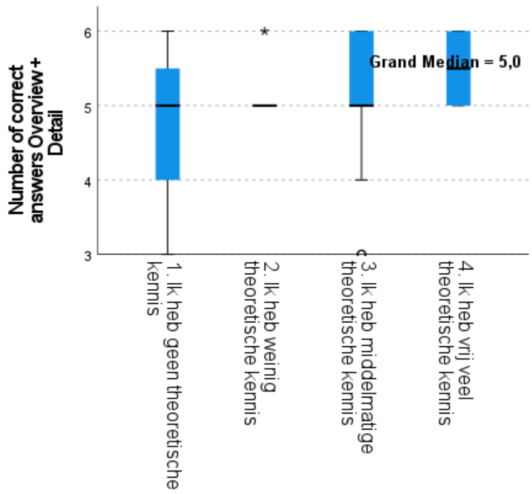


*Figure 54 - Indep.-Samples median test task effectiveness (theoretical knowledge, 'overview+detail')*

Because all test fields are less than or equal to the median the test could not be performed for the 'focus+context' strategy, see Table 44.

*Table 44 - Hypothesis test task effectiveness (theoretical knowledge, 'focus+context')*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers 'focus+context' are the same across categories of theoretical knowledge. | Independent-Samples Median Test | [a] | Unable to compute |

a. All test field values are less than or equal to the median.

Although the test could not be carried out for the 'focus+context strategy, the test for the 'overview+detail' strategy shows that there is no significant difference (p = 0,203) in number of correct answers given by participants with different levels of theoretical knowledge.

## Task efficiency

Besides the task effectiveness, the 'task efficiency' is also analyzed using the Independent-Samples median test. The statistics are shown in Table 45.

*Table 45 - Case summaries number of correct answers per hour (theoretical knowledge)*

| Theoretical knowledge | | Number of correct answers per hour 'overview+detail' | Number of correct answers per hour 'focus+context' |
|---|---|---|---|
| I have no theoretical knowledge | Mean | 68,02 | 55,99 |
| | Median | 55,54 | 53,90 |
| | Minimum | 26,09 | 14,20 |
| | Maximum | 137,96 | 100,86 |
| I have rather weak theoretical knowledge | Mean | 113,68 | 89,69 |
| | Median | 98,65 | 83,22 |
| | Minimum | 62,25 | 37,04 |
| | Maximum | 208,68 | 229,57 |
| I have mediocre theoretical knowledge | Mean | 95,53 | 77,46 |
| | Median | 84,43 | 78,45 |
| | Minimum | 30,34 | 34,97 |
| | Maximum | 155,54 | 115,24 |
| I have rather strong theoretical knowledge | Mean | 69,62 | 84,43 |
| | Median | 66,53 | 79,42 |
| | Minimum | 38,78 | 34,02 |
| | Maximum | 110,81 | 147,36 |
| Total | Mean | 87,48 | 75,18 |
| | Median | 81,86 | 74,33 |
| | Minimum | 26,09 | 14,20 |
| | Maximum | 208,68 | 229,57 |

*Table 46 - Hypothesis test task efficiency (theoretical knowledge, 'overview+detail')*

The results of the test for the 'overview+detail strategy are shown in the next tables and figure.

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers per hour 'overview+detail' are the same across categories of theoretical knowledge. | Independent-Samples Median Test | 0,099 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 47 - Indep.-Samples median test task efficiency (theoretical knowledge, 'overview+detail')*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 81,862 |
| Test Statistic | 6,272[a,b] |
| Degree Of Freedom | 3 |
| Asymptotic Sig. (2-sided test) | 0,099 |

a. More than 20% of the cells have expected values less than five.
b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

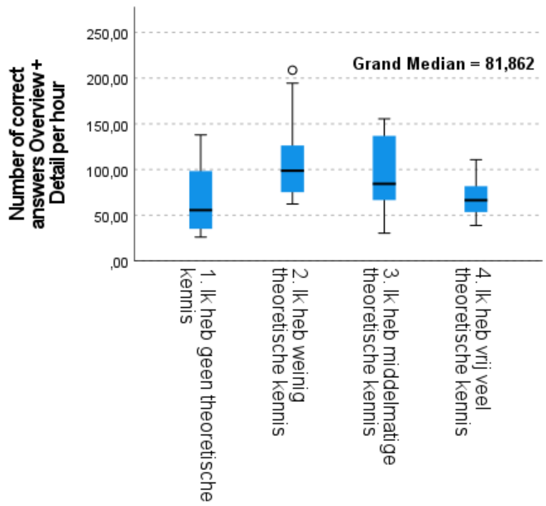


*Figure 55 - Indep.-Samples median test task efficiency (theoretical knowledge, 'overview+detail')*

The same test is performed for the 'focus+context' interface strategy, see next tables and figure.

*Table 48 - Hypothesis test task efficiency (theoretical knowledge, 'focus+context')*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers per hour 'focus+context' are the same across categories of theoretical knowledge. | Independent-Samples Median Test | 0,618 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 49 - Indep.-Samples median test task efficiency (theoretical knowledge, 'focus+context')*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 74,326 |
| Test Statistic | 1,787[a,b] |
| Degree Of Freedom | 3 |
| Asymptotic Sig. (2-sided test) | 0,618 |

a. More than 20% of the cells have expected values less than five.
b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

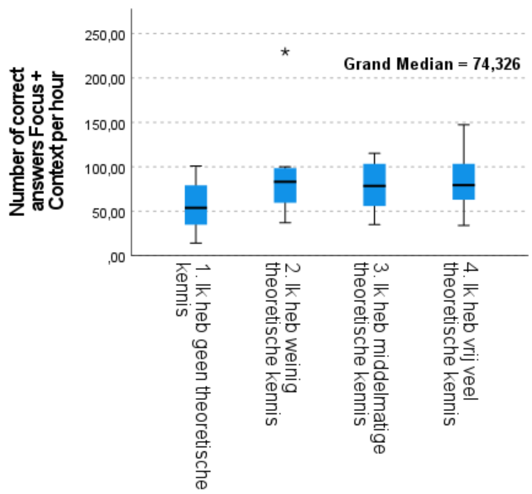


*Figure 56 - Indep.-Samples median test task efficiency (theoretical knowledge, 'focus+context')*

Both tests show that there is no significant difference (α = 0,050) in the number of correct answers per hour given by participants with different levels of theoretical knowledge of business process modeling.

The conclusion of this study is that the 'theoretical modeling competency' does not play a significant role in the use of process models for either interface strategy.

## Practical modeling competency

The Independent-Samples median test is used to find out whether the participants' use of process models in their daily practice causes a significant difference in the results. This test is used because the participant are divided into four groups by using the self-assessed practical experience. The summary of the answers is given in Table 50.

*Table 50 - Case summaries number of correct answers (practical experience)*

| Practical experience | | Number of correct answers 'overview+detail' | Number of correct answers 'focus+context' |
|---|---|---|---|
| I never use business process modeling in practice | Mean | 4,73 | 4,67 |
| | Median | 5,00 | 5,00 |
| | Minimum | 3 | 3 |
| | Maximum | 6 | 6 |
| I sometimes use business process modeling in practice | Mean | 5,12 | 5,47 |
| | Median | 5,00 | 6,00 |
| | Minimum | 3 | 4 |
| | Maximum | 6 | 6 |
| I regularly use business process modeling in practice, but not every day | Mean | 5,43 | 5,00 |
| | Median | 5,00 | 6,00 |
| | Minimum | 5 | 1 |
| | Maximum | 6 | 6 |
| I use business process modeling in practice every day | Mean | 5,50 | 6,00 |
| | Median | 5,50 | 6,00 |
| | Minimum | 5 | 6 |
| | Maximum | 6 | 6 |
| Total | Mean | 5,05 | 5,12 |
| | Median | 5,00 | 6,00 |
| | Minimum | 3 | 1 |
| | Maximum | 6 | 6 |

## Task effectiveness

The summary of the Independent-Samples median test for the ‘task effectiveness’ in case of using the ‘overview+detail’ interface strategy, is shown in Table 51, Table 52 and Figure 57.

*Table 51 - Hypothesis test task effectiveness (practical experience, ‘overview+detail’)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers ‘overview+detail’ are the same across categories of practical experience. | Independent-Samples Median Test | 0,289 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 52 - Indep.-Samples median test task effectiveness (practical experience, ‘overview+detail’)*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 5,000 |
| Test Statistic | 3,754[a,b,c] |
| Degree Of Freedom | 3 |
| Asymptotic Sig. (2-sided test) | 0,289 |

a. At least one cell has an expected value less than one.
b. More than 20% of the cells have expected values less than five.
c. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

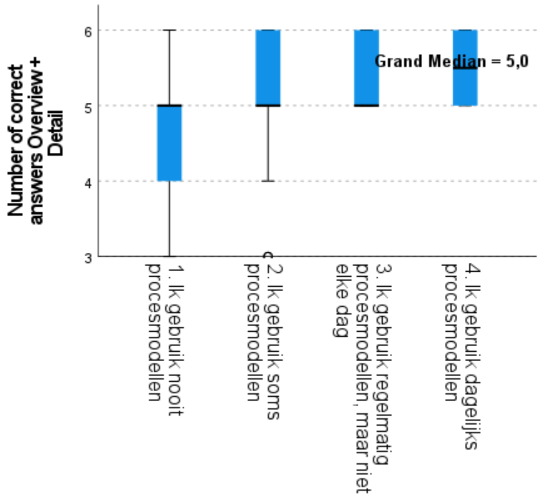


*Figure 57 - Indep.-Samples median test task effectiveness (practical experience, 'overview+detail')*

The same Independent-Samples median test is carried out for the results using the ‘focus+context’ interface strategy. In this case the test cannot be performed because all test fields are less or equal to the median, see Table 53.

*Table 53 - Hypothesis test task effectiveness (practical experience, 'focus+context')*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|
| The medians of number of correct answers 'focus+context' are the same across categories of practical experience. | Independent-Samples Median Test | [a] | Unable to compute |

a. All test field values are less than or equal to the median.

As far as can be concluded from the tests, there is no significant difference in number of correct answers given by participants considering their different levels of practical use of process models.

## Task efficiency

The 'task efficiency' is also analyzed using the Independent-Samples median test. The statistics are shown in Table 54.

*Table 54 - Case summaries number of correct answers per hour (practical experience)*

| **Practical experience** | | **Number of correct answers per hour 'overview+detail'** | **Number of correct answers per hour 'focus+context'** |
|---|---|---|---|
| I never use business process modeling in practice | Mean | 78,61 | 67,83 |
| | Median | 75,28 | 63,06 |
| | Minimum | 26,09 | 14,20 |
| | Maximum | 137,96 | 147,36 |
| I sometimes use business process modeling in practice | Mean | 102,07 | 81,31 |
| | Median | 90,72 | 74,33 |
| | Minimum | 37,22 | 33,07 |
| | Maximum | 208,68 | 229,57 |
| I regularly use business process modeling in practice, but not every day | Mean | 84,07 | 87,56 |
| | Median | 81,86 | 84,51 |
| | Minimum | 61,64 | 59,61 |
| | Maximum | 144,89 | 115,24 |
| I use business process modeling in practice every day | Mean | 41,95 | 34,49 |
| | Median | 41,95 | 34,49 |
| | Minimum | 30,34 | 34,02 |
| | Maximum | 53,56 | 34,97 |
| Total | Mean | 87,48 | 75,18 |
| | Median | 81,86 | 74,33 |
| | Minimum | 26,09 | 14,20 |
| | Maximum | 208,68 | 229,57 |

The results of the test for the ‘overview+detail’ strategy are shown in next tables and Figure 58.

*Table 55 - Hypothesis test task efficiency (practical experience, ‘overview+detail’)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers per hour ‘overview+detail’ are the same across categories of practical experience. | Independent-Samples Median Test | 0,437 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 56 - Indep.-Samples median test task efficiency (practical experience, ‘overview+detail’)*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 81,862 |
| Test Statistic | 2,761[a,b,c] |
| Degree Of Freedom | 3 |
| Asymptotic Sig. (2-sided test) | 0,437 |

a. At least one cell has an expected value less than one.
b. More than 20% of the cells have expected values less than five.
c. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

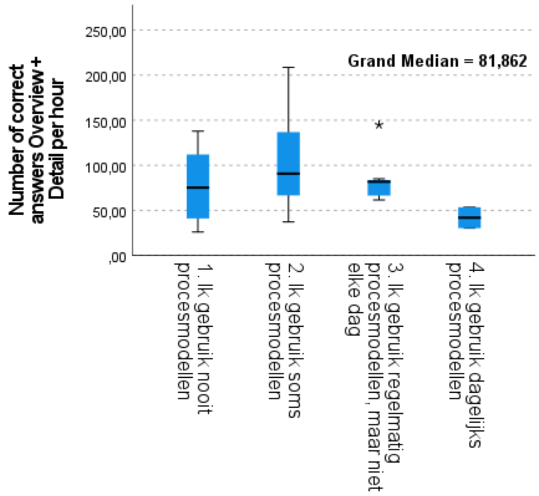


*Figure 58 - Indep.-Samples median test task efficiency (practical experience, 'overview+detail')*

The results of the test for the ‘focus+context’ strategy are shown in the next tables (Table 57 and Table 58) and Figure 59.

*Table 57 - Hypothesis test summary task efficiency (practical experience, 'focus+context')*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers per hour 'focus+context' are the same across categories of practical experience. | Independent-Samples Median Test | 0,335 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 58 - Indep.-Samples median test task efficiency (practical experience, 'focus+context')*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 74,326 |
| Test Statistic | 3,389[a,b,c] |
| Degree Of Freedom | 3 |
| Asymptotic Sig. (2-sided test) | 0,335 |

a. At least one cell has an expected value less than one.
b. More than 20% of the cells have expected values less than five.
c. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

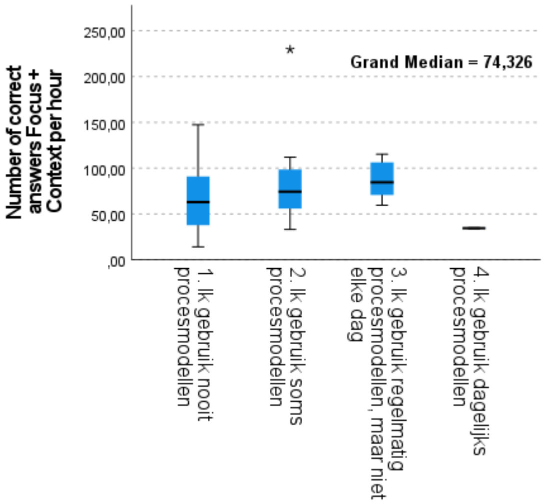


*Figure 59 - Indep.-Samples median test task efficiency (practical experience, 'focus+context')*

It follows from the tests that for both interface strategy 'overview+detail' and 'focus+context' there is no significant difference in number of correct answers per hour given by participants, taking into account their different levels of practical use of business process modeling. Therefore, it is concluded that the 'practical modeling competency' does not play a significant role when using process models for either interface strategy.

### Education

The third confounding variable 'education' is also checked to see if a different highest level of education leads to a significant difference in understanding of process models. To check this, the Independent-Samples median test is used. For this test the participants are divided into six groups by using the default education levels. Two participants choose this option 'Other', see also Appendix F. A summary of the statistics is shown in Table 59.

*Table 59 - Case summaries number of correct answers (education)*

| Education | | Number of correct answers 'overview+detail' | Number of correct answers 'focus+context' |
|---|---|---|---|
| Secondary School (Middelbare school) | Mean | 4,00 | 6,00 |
| | Median | 4,00 | 6,00 |
| | Minimum | 4 | 6 |
| | Maximum | 4 | 6 |
| Post-secondary vocational education (MBO) | Mean | 5,00 | 6,00 |
| | Median | 5,00 | 6,00 |
| | Minimum | 5 | 6 |
| | Maximum | 5 | 6 |
| Higher vocational education (HBO) | Mean | 4,92 | 4,58 |
| | Median | 5,00 | 5,00 |
| | Minimum | 3 | 1 |
| | Maximum | 6 | 6 |
| University education (WO) | Mean | 5,10 | 5,50 |
| | Median | 5,00 | 6,00 |
| | Minimum | 3 | 3 |
| | Maximum | 6 | 6 |
| Postgraduate (Postdoctoraal) | Mean | 5,20 | 4,80 |
| | Median | 5,00 | 5,00 |
| | Minimum | 5 | 4 |
| | Maximum | 6 | 6 |
| Other (Overige) | Mean | 5,50 | 4,50 |
| | Median | 5,50 | 4,50 |
| | Minimum | 5 | 1 |
| | Maximum | 6 | 6 |
| Total | Mean | 5,05 | 5,12 |
| | Median | 5,00 | 6,00 |
| | Minimum | 3 | 1 |
| | Maximum | 6 | 6 |

## Task effectiveness

The test summary of the 'task effectiveness' for the 'overview+detail' strategy is shown in Table 60, Table 61 and Figure 60.

*Table 60 - Hypothesis test task effectiveness (education, 'overview+detail')*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers 'overview+detail' are the same across categories of highest education. | Independent-Samples Median Test | 0,786 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 61 - Indep.-Samples median test task effectiveness (education, 'overview+detail')*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 5,000 |
| Test Statistic | 2,439[a,b,c] |
| Degree Of Freedom | 3 |
| Asymptotic Sig. (2-sided test) | 0,786 |

a. At least one cell has an expected value less than one.
b. More than 20% of the cells have expected values less than five.
c. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

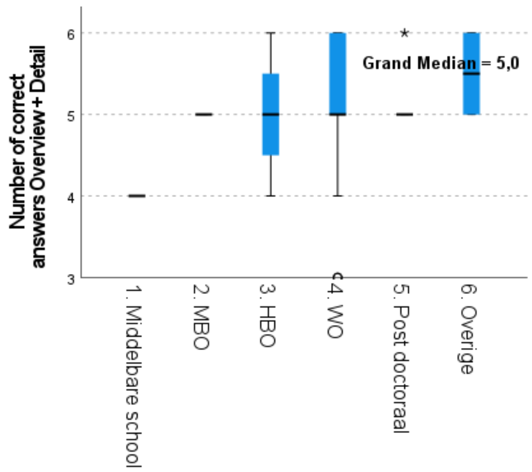


*Figure 60 - Indep.-Samples median test task effectiveness (education, 'overview+detail')*

The test for strategy ‘focus + context’ could not be carried out because all test values are less than or equal to the median, see Table 62.

*Table 62 - Hypothesis test task effectiveness (education, ‘focus+context’)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|
| The medians of number of correct answers ‘focus+context’ are the same across categories of highest education. | Independent-Samples Median Test | [a] | Unable to compute |

a. All test field values are less than or equal to the median.

Although not all tests could be carried out successfully, no significant difference in number of correct answers considering the different levels of highest education of the participants could be established.

## Task efficiency

The ‘task efficiency’ is analyzed using the Independent-Samples median test as well. The overall statistics are shown in Table 63.

*Table 63 - Case summaries number of correct answers per hour (education)*

| Education | | Number of correct answers per hour ‘overview+detail’ | Number of correct answers per hour ‘focus+context’ |
|---|---|---|---|
| Secondary School (Middelbare school) | Mean | 137,96 | 59,30 |
| | Median | 137,96 | 59,30 |
| | Minimum | 137,96 | 59,30 |
| | Maximum | 137,96 | 59,30 |
| Post-secondary vocational education (MBO) | Mean | 75,28 | 94,15 |
| | Median | 75,28 | 94,15 |
| | Minimum | 75,28 | 94,15 |
| | Maximum | 75,28 | 94,15 |
| Higher vocational education (HBO) | Mean | 81,39 | 73,85 |
| | Median | 74,83 | 67,51 |
| | Minimum | 34,05 | 33,07 |
| | Maximum | 144,89 | 147,36 |
| University education (WO) | Mean | 96,69 | 79,65 |
| | Median | 83,02 | 79,87 |
| | Minimum | 30,34 | 34,02 |
| | Maximum | 208,68 | 229,57 |
| Postgraduate (Postdoctoraal) | Mean | 79,63 | 72,22 |
| | Median | 84,43 | 75,76 |

| | | | |
|---|---|---|---|
| | Minimum | 26,09 | 31,36 |
| | Maximum | 38,78 | 98,52 |
| Other (Overige) | Mean | 87,48 | 44,26 |
| | Median | 81,86 | 44,26 |
| | Minimum | 26,09 | 14,20 |
| | Maximum | 38,78 | 74,33 |
| Total | Mean | 87,48 | 75,18 |
| | Median | 81,86 | 74,33 |
| | Minimum | 26,09 | 14,20 |
| | Maximum | 208,68 | 229,57 |

The results of the test for the 'overview+detail' strategy are shown in Table 64, Table 65 and Figure 61.

*Table 64 - Hypothesis test task efficiency (education, 'overview+detail')*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers per hour 'overview+detail' are the same across categories of highest education. | Independent-Samples Median Test | 0,452 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 65 - Indep.-Samples median test task efficiency (education, 'overview+detail')*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 81,862 |
| Test Statistic | 4,712[a,b,c] |
| Degree Of Freedom | 5 |
| Asymptotic Sig. (2-sided test) | 0,452 |

a. At least one cell has an expected value less than one.
b. More than 20% of the cells have expected values less than five.
c. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

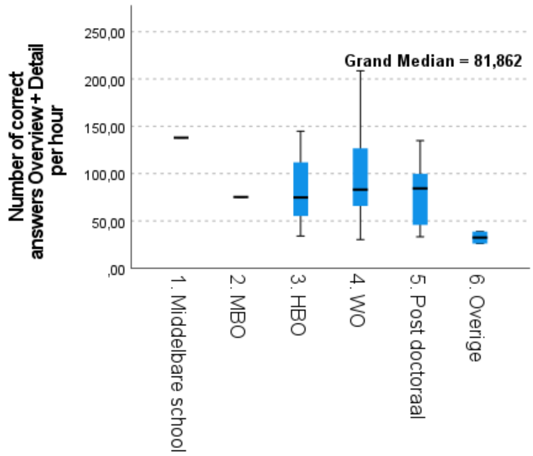


*Figure 61 - Indep.-Samples median test task efficiency (education, 'overview+detail')*

The results for the test of the ‘focus+context’ strategy is shown in the next tables (Table 66 and Table 67 ) and Figure 62.

*Table 66 - Hypothesis test task efficiency (education, ‘focus+context)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The medians of number of correct answers per hour ‘focus+context’ are the same across categories of highest education | Independent-Samples Median Test | 0,452 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

*Table 67 - Indep.-Samples median test task efficiency (education, ‘focus+context’)*

**Independent-Samples Median Test Summary**

| | |
|---|---|
| Total N | 41 |
| Median | 74,326 |
| Test Statistic | 4,712[a,b,c] |
| Degree Of Freedom | 5 |
| Asymptotic Sig. (2-sided test) | 0,452 |

a. At least one cell has an expected value less than one.
b. More than 20% of the cells have expected values less than five.
c. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

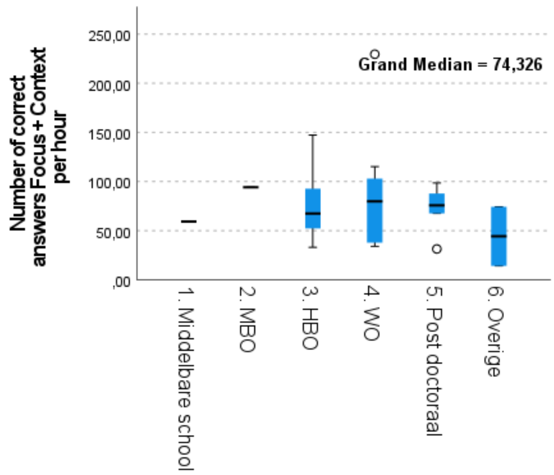


*Figure 62 - Indep.-Samples median test task efficiency (education, 'focus+context')*

It is concluded that the difference in the highest educational level of the participants does not result in a significant difference in the understanding of process models.

# Appendix I – Local and global questions

As seen in Appendix D, the degree of difficulty of the questions for the 'focus+context' and 'overview+detail' strategies was taken into account. The distribution of 'local' and 'global' questions over both strategies was also taken into account. For each strategy, three questions from the local category and three from the global category were included. Table 68 shows which question is regarded as local and which questions is regarded as global.

*Table 68 - Local and global questions*

| | Questions 'overview+detail' | | | Questions 'focus+context' | |
|---|---|---|---|---|---|
| 1 | If the process has gone through process step B3, does the process then always also go through process step G3? | Global | 1 | If process step G5 has been passed, does the process always pass through process step B3? | Global |
| 2 | If the process passes through process step C31, does it always also pass through process step C317? | Local | 2 | If the process enters step D4, does it always go through process step D47? | Local |
| 3 | Can the process pass through process step D47 more than once? | Local | 3 | Can the process pass through process step C317 multiple times? | Local |
| 4 | Can process steps H2 and H7 be executed simultaneously? | Local | 4 | Can process steps H2 and H10 be executed simultaneously? | Local |
| 5 | If the process has passed through process step J2, can it pass through process step I324 again? | Global | 5 | If the process has gone through process step I322, can it also go through process step J5 ? | Global |
| 6 | Can process steps E3 and I324 be performed simultaneously? | Global | 6 | Can process steps J7 and C1 be performed simultaneously? | Global |

## Test local questions

In order to check whether there is a significant difference between results of the two interface strategies results when only the local questions are taken into account, the Wilcoxon signed-rank test is used. The outcome of this test is shown in Table 69.

*Table 69 - Hypothesis test task effectiveness (local questions)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The median of differences between number of correct answers (local) 'overview+detail' and number of correct answers (local) 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,157 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

From the outcome of this test, it is concluded that there is no significant difference (p = 0,157) between the 'task effectiveness' for the 'overview+detail' and the 'focus+context' interface strategies when local questions are asked.

Table 70 shows the result of the Wilcoxon signed-rank test when testing the 'task efficiency' between both interface strategies.

*Table 70 - Hypothesis test task efficiency (local questions)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.$^{a,b}$ | Decision |
|---|---|---|---|
| The median of differences between number of correct answers per hour (local) 'overview+detail' and number of correct answers per hour (local) 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,004 | Reject the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

The null hypothesis is rejected (p = 0,004), this means that a significant difference is observed between the number of correct answers per hour for the different interface strategies. In this case, it could mean that participants answered more local questions per hour correctly when using the 'overview+detail' strategy than when using the 'focus+context' strategy. In other words, the 'task efficiency' for the 'overview+detail' interface strategy is higher than for the 'focus+context' strategy when asking local questions.

## Test global questions

The same tests as for the local questions are performed for the global questions.

The outcome of the Wilcoxon signed-rank test for testing the 'task effectiveness' is given in Table 71.

*Table 71 - Hypothesis test task effectiveness (global questions)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.$^{a,b}$ | Decision |
|---|---|---|---|
| The median of differences between number of correct answers (global) 'overview+detail' and number of correct answers (global) 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,653 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

No significant difference between the 'task effectiveness' of the 'overview+detail' and the 'focus+context' interface strategy (p = 0,653) is observed when asking global questions.

Table 72 shows the result of the Wilcoxon signed-rank test.

*Table 72 - Hypothesis test task efficiency (global questions)*

**Hypothesis Test Summary**

| Null Hypothesis | Test | Sig.[a,b] | Decision |
|---|---|---|---|
| The median of differences between number of correct answers per hour (global) 'overview+detail' and number of correct answers per hour (global) 'focus+context' equals 0. | Related-Samples Wilcoxon Signed Rank Test | 0,946 | Retain the null hypothesis. |

a. The significance level is 0,050.
b. Asymptotic significance is displayed.

In the case of answering global questions, no significant difference (p = 0,946) is observed between the 'task efficiency' for the different interface strategies. This means that participants answered the same number of questions per hour correctly when using 'overview+detail' strategy and the 'focus+context' strategy.