

# Journal of the Association for Information Systems

JAIS 

Research Article

## The Influence of Notational Deficiencies on Process Model Comprehension

**Kathrin Figl**

Vienna University of Economics and Business  
kathrin.figl@gmail.com

**Jan Mendling**

Vienna University of Economics and Business  
contact@mendling.com

**Mark Strembeck**

Vienna University of Economics and Business  
mark.strembeck@wu.ac.at

### Abstract

*Visual process models are helpful when analyzing and improving complex organizational processes. However, the symbol sets used in different modeling notations vary in perceptual discriminability, visual expressiveness, and semantic transparency; such factors are likely to influence a notation's perception and cognitive effectiveness. In this paper, we investigate whether the basic symbol sets in visual process modeling languages influence comprehension and cognitive load of process models. For this purpose, we analyze four different symbol sets in an experiment with model comprehension tasks carried out by 136 participants. Our results indicate that notational deficiencies concerning perceptual discriminability and semiotic clarity have measurable effects on comprehension, cognitive load, and the time needed to understand the models.*

**Keywords:** *Process Modeling, Notational Design, Comprehension, Cognitive Effectiveness.*

---

\* Jeffrey Parsons was the accepting senior editor. This article was submitted on 11<sup>th</sup> April 2011 and went through three revisions.

Volume 14, Issue 6, pp. 312-338, June 2013

## 1. Introduction

Business process models play an important role in documenting organizational processes. They help capture the operational procedures that need to be supported by an organization's information system. For this reason, business process models are instrumental in defining software system requirements. By the same token, many errors in software systems can be traced to conceptual issues stemming from the requirement phase (Darke & Shanks, 1997). Low-quality requirements models tend to make system development altogether less efficient (in terms of time, cost, and effort) and less effective (in terms of the quality of results) (Charette, 2005). Therefore, defining comprehensible, consistent, and unambiguous business process models is highly relevant to information systems engineering. This is especially true since it can help to correct errors early on in the software development process, when it is much easier and less expensive compared with later stages.

Many visual modeling languages have been specifically developed to support communication among project participants (Aranda, Ernst, Horkoff, & Easterbrook, 2007). This implies that using these languages should be easy both for those creating models and those reading them. Such a modeling language should have precise syntax, well-defined semantics, and a suitable visual notation (Mendling, 2008). While syntax and semantics of many of these languages are grounded in Petri net concepts (Lohmann, Verbeek, & Dijkman, 2009), it is remarkable that the design of their visual notation is hardly approached in a scientific manner. In this paper, we focus on the visual notation of business process modeling languages. The notation of a process modeling language can be identified through a set of symbols that visually represent the underlying abstract concepts. The term "process modeling notation" therefore focuses solely on this visual aspect of a process modeling language. Research in this area of conceptual modeling is important due to the great wealth and variety of existing process modeling languages: since its emergence in the 1970s in the context of office automation systems, process modeling has grown to become one of the most important areas of conceptual modeling (Dumas, Aalst, & Hofstede, 2005; Melão & Pidd, 2000).

Among the wide range of existing languages are Event-driven Process Chains (EPCs) (Keller, Nüttgens, & Scheer, 1992; Scheer, 2000), Unified Modeling Language (UML) Activity Diagrams (Object Management Group, 2011a), Yet Another Workflow Language (YAWL) (van der Aalst & ter Hofstede, 2005), and Business Process Model and Notation (BPMN) (Object Management Group, 2011b). To discuss the usability of these languages, we turn to cognitive research: it explores how the human mind processes information, creates knowledge, and solves problems. Relevant cognitive theories include the cognitive load theory (Sweller, 1988), the cognitive fit theory (Vessey, 1991), the cognitive dimensions framework for notational systems (Green & Petre, 1996), and the theory of multimedia learning (Mayer, 2001).

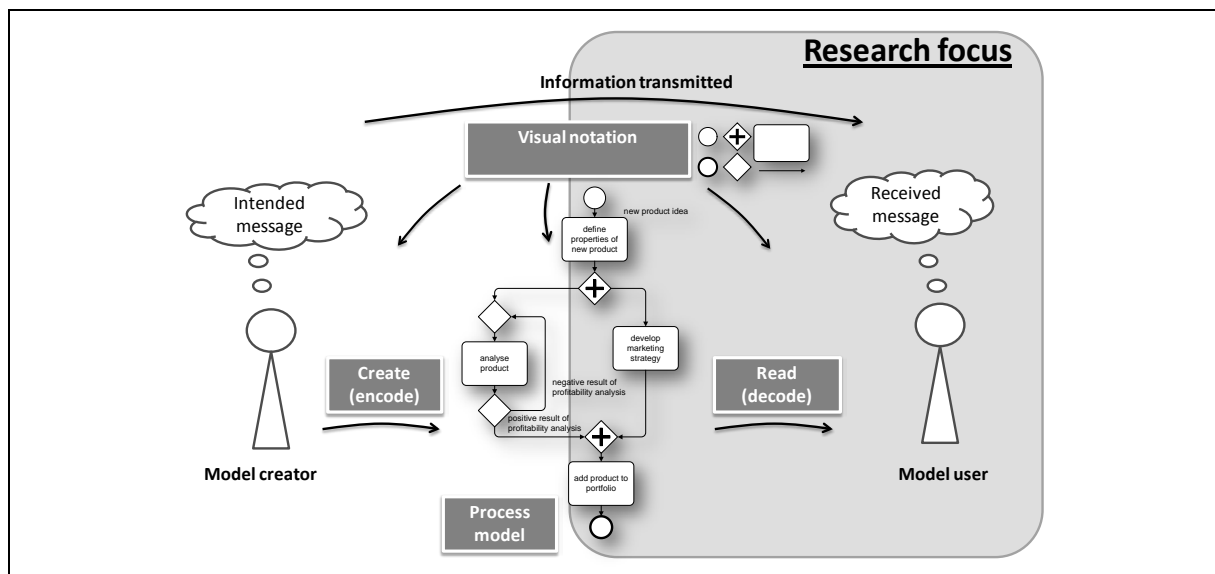
Indeed, cognitive aspects have been found to play an important role in assessing the efficiency and effectiveness of modeling languages, which includes their modularity (Reijers, Mendling, & Dijkman, 2011), ontological differences (Recker, Rosemann, Indulska, & Green, 2009), learnability (Recker & Dreiling, 2007), and control-flow representation (Sarshar & Loos, 2005). However, the role of visual notation in this context has not yet been studied thoroughly, even though there are strong indications that it is crucial; for instance, in literature on the perceptual effectiveness of notations (Moody, 2009) and the efficiency of information search and problem solving (Larkin & Simon, 1987). Furthermore, empirical evidence suggests that notational differences influence comprehension of data models (Hitchman, 2002; Nordbotten & Crosby, 1999). First conceptual analyses based on the principles identified by Moody (2009) point to significant differences in several modeling notations (Figl, Derntl, Rodriguez, & Botturi, 2010; Figl, Mendling, Strembeck, & Recker, 2010; Genon, Amyot, & Heymans, 2010; Genon, Heymans, & Amyot, 2010; Moody, Heymans, & Matulevicius, 2010). However, all these works remain on a qualitative level. It has not yet been investigated how strongly notational deficiencies impair model understanding, and whether the effect is statistically significant. The quality of a notation is of particular importance for business process modeling. It has been identified as a critical success factor for modeling success, most notably regarding model quality and user satisfaction (Bandara, Gable, & Rosemann, 2005). There is anecdotal evidence that process

modeling projects fail because the notation is not accepted by certain groups of stakeholders (Rosemann, 2006). This paper investigates the relative strengths and weaknesses of the symbol sets of EPCs, UML, YAWL, and BPMN. It has so far not been clarified which symbols should be preferred and why; we address this research gap. We build on research into the “physics of notations” (Moody, 2009), which integrates different cognitive theories into an overarching framework to discuss which properties of notational elements are desirable from a cognitive perspective. Giving experimental insight into the significance of visual notations for process model understanding, our findings are directly applicable to research and practice. The importance of visual notations suggests that the syntactic characteristics of a modeling language and its visual representation have to be studied separately. Furthermore, our results can serve as guidance when selecting the notation for process modeling projects.

## 2. Theoretical Background

### 2.1. Factors Influencing Model Understanding

A model’s design aims to draw the viewer’s attention to those components that are crucial for understanding and cognitive inferring (Scaife & Rogers, 1996). Our study focuses on a model user’s understanding of a process model. Figure 1 shows the model user and highlights their ability to read and decode a process model. The model depicted uses the visual notation of BPMN, a popular process modeling language. In essence, languages like BPMN are specific kinds of graphs. Tasks, which define elementary pieces of work, are captured as nodes. The arcs of the graph describe temporal and logical dependencies among these tasks. There are usually two types of routing elements: the first node type is used for specifying decision points towards alternative branches and corresponding merge nodes, and the second type indicates splits into parallel branches of execution and corresponding points of synchronization. Finally, many process modeling notations define symbols that signal the start and the end of a process.



**Figure 1. The Influence of Visual Notation on Creation and Reading (Understanding) of Models (Adapted from Moody, 2009).**

Whether a model user can read and decode a process model in an efficient and effective way also depends on the visual representation of the model. Definitions of the term “model understandability” as the ease with which the model can be understood (Canfora, Garc, Piattini, Ruiz, & Visaggio, 2005; Moody, 1998) emphasize the relevance of cognitive load theory in this context. Understanding the complex control flow logic of process models is a task likely to demand high cognitive effort. Cognitive activities such as visual perception, attention, short and long-term memory processing, reasoning,

and problem solving have to be performed. However, humans have only limited information-processing capabilities (Vessey, 1991), and this, in turn, means that understanding should be facilitated by keeping the cognitive load for model users low.

According to the cognitive load theory, high cognitive load during problem-solving exercises impairs learning and knowledge acquisition (Sweller, 1988). The theory differentiates between three types of cognitive load: intrinsic, extraneous, and germane. Intrinsic cognitive load is determined by the complexity of information (i.e., the amount of elements, and their relations and interactions). Therefore, very complex and large models are likely to increase cognitive load and may adversely affect understanding (Gruhn & Laue, 2009; Nordbotten & Crosby, 1999). In contrast to intrinsic cognitive load, extraneous cognitive load is influenced by the way information is represented (Kirschner, 2002). Different types of representation can impact the relative difficulty of a task, depending on the different levels of cognitive load involved (Kotovsky, Hayes, & Simon, 1985). Finally, germane cognitive load refers to instructional information that helps a person solve a particular task. While the cognitive load devoted to learning and understanding (i.e. germane cognitive load) should be promoted, extraneous cognitive load should be kept low.

There are different options for reducing extraneous cognitive load in process modeling, and a variety of quality aspects must be kept in mind (Siau & Tan, 2005). Table 1 gives an overview of some factors that could influence the cognitive load involved in model understanding, with a focus on notational aspects. Researching notational aspects is particularly appealing because they can be modified more easily than other comprehension factors, such as the theoretical knowledge of model users. Also model inherent factors like size and complexity do matter (Reijers & Mendling, 2011), but these can hardly be modified when a particular aspect of a domain has to be represented.

This paper extends research into model understanding by a study on the influence of notational elements. According to Moody (2009), a visual notation consists of a set of graphical symbols (visual vocabulary), a set of compositional rules (visual grammar), and definitions of the meaning of each symbol (visual semantics). Process modeling notations use different vocabularies to visualize concepts and elements of different types. Cognitive load theory helps illustrate the effect of notation on cognitive effectiveness. If the same information is modeled using different notations, the resulting models imply a comparable intrinsic cognitive load. However, considerable variations in their symbol sets might imply differences in extraneous cognitive load, and consequently in the performance of understanding (Chandler & Sweller, 1996). To determine which aspects of a notation can cause additional extraneous cognitive load, a recent framework for the cognitively effective design of modeling notations proposes nine principles (Moody, 2009): semiotic clarity, graphic economy, perceptual discriminability, visual expressiveness, dual coding, semantic transparency, cognitive fit, complexity management, and cognitive integration. These criteria allow us to look at individual symbols from various angles. When discussing symbol sets which offer different symbols for the same set of semantic concepts, only six of the principles have to be considered. The criteria cognitive fit, complexity management, and cognitive integration would be relevant for an overall evaluation of process modeling notations, but not for investigating only the symbol sets. The relevant principles are discussed in Section 2.2.

**Table 1. Factors Influencing the Cognitive Load Involved in Understanding Models**

Source of cognitive load		Influence factors on cognitive load	
Extraneous cognitive load	Notational design level	Symbol	Semiotic clarity, visual expressiveness, semantic transparency
		Symbol set	Graphic economy, perceptual discriminability, visual expressiveness, semiotic clarity
		Primary notation	Graphic economy, dual coding, cognitive fit, complexity management, cognitive integration
	(Process) Model level	Secondary notation	Model layout (edge crossings, modularity), textual labels
		Inherent factors	Size, density, structuredness, control flow and structure, complexity
Intrinsic cognitive load	Domain level		Complexity
	User level		Familiarity and expertise with domain, notation, cognitive abilities

## 2.2. Influencing Factors on the Level of Symbols and Symbol Sets

Moody uses the term “symbol set” for the visual vocabulary of a modeling notation, which comprises graphical elements such as lines, areas, and spatial relationships (Moody & Hillegersberg, 2008). Table 1 distinguishes between “symbol set” and “symbol” because some symbol characteristics can only be defined in relation to other symbols (McDougall, Curry, & Bruijn, 1999). For instance, if and to what extent a symbol is understood intuitively (criterion of semantic transparency) can be determined by looking at the symbol alone, but, when we want to investigate perceptual discriminability of symbols, we must analyze the entire symbol set. The subsequent paragraphs discuss the criteria relevant for symbol choice as proposed and named by Moody (2009), combining them with other systems of symbol characteristics such as that of McDougall et al. (1999).

### 2.2.1. Semiotic Clarity and Graphic Economy

The principle of semiotic clarity underlines the importance of a good fit between the graphical symbols used in a visual notation and the semantic concepts they refer to. The concept of semiotic clarity extends Wand and Weber’s (1993) theory on ontological clarity and completeness to the area of visual syntaxes (Moody & Hillegersberg, 2008). Anomalies such as symbol redundancy (several symbols represent the same concept), overload (one symbol represents more than one concept), symbol excess and deficit (there are graphical symbols without a correspondence to a semantic construct or vice versa) should be avoided, since they lead to ambiguity and unnecessary cognitive load for the user (Moody & Hillegersberg, 2008). To achieve cognitive effectiveness, graphics must be constrained. Consequently, notations also highlight specific aspects of information at the expense of others (Green & Petre, 1996). The principle of graphic economy demands a reasonable balance between the expressiveness of a notation and the number of its symbols.

### 2.2.2. Perceptual Discriminability

The perceptual discriminability of symbols looks at how easy it is for a user to distinguish between different symbols and to visually recognize differences between them. This strongly depends on the amount of visual variables (e.g., size, color, shape) in which symbols differ (also referred to as visual distance). If symbols are highly unique with regard to their visual representation, they are likely to “pop out” and are easy to spot in a model (Moody & Hillegersberg, 2008). Low perceptual discriminability can lead to misunderstanding. For instance, rectangles and diamonds in ER diagrams are easily confused with each other (Nordbotten & Crosby, 1999). On the other hand, if different symbols in a notation are similar (e.g. in color or shape), they are likely to be recognized as belonging together in accordance with the “Gestalt law of similarity” (Wertheimer, 1938).



### 2.2.3. Visual Expressiveness

Modeling notations that exploit the full range of visual variables (spatial dimensions like horizontal and vertical, as well as shape, size, color, brightness, orientation, and texture) have higher visual expressiveness. Research results about the optimal visual complexity of symbols are inconsistent: some researchers recommend keeping symbols as simple as possible, others argue that complexity and detail can make them easier to use (McDougall et al., 1999).

### 2.2.4. Dual Coding

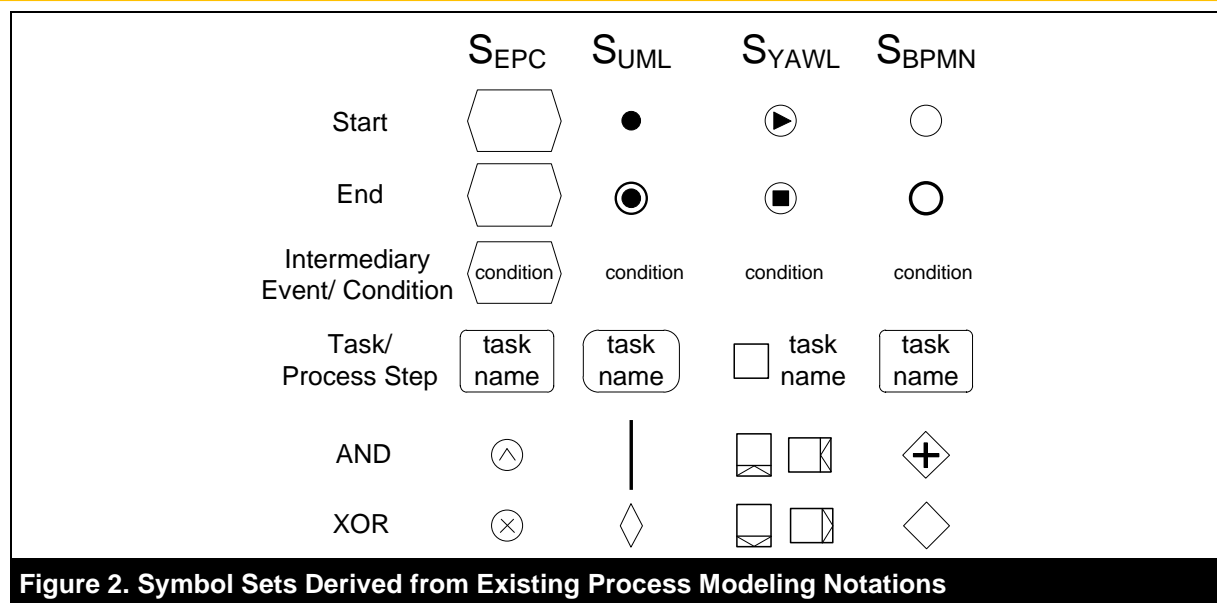
The dual coding principle refers to the visual combination of text and graphical representation. It is based on theories of short-term memory and learning. The dual coding theory (Paivio, 1991) postulates, for instance, that visual information (e.g., pictures) and verbal information (e.g., texts) are stored and processed differently via separate mental channels that do not compete with each other. According to the contiguity principle of cognitive multimedia learning theory (Mayer, 2001), learning outcome is higher if text and pictures are presented next to each other. Therefore, text and symbols belonging together should also be placed near each other in visual models.

### 2.2.5. Semantic Transparency

According to Moody (2009), semantic transparency allows for an easy association of graphic symbols and their corresponding meaning. Similarly, McDougall et al. (1999, p. 489) refers to semantic distance as the continuum of “the closeness of the relationship between the symbol and what it is intended to present”. Icons, for example, are easily associated with their referent real-world concepts, because there is a direct link between visual appearance and meaning (Mendling, Recker, & Reijers 2010). In comparison, symbols have a rather distant relationship with their meaning that is described as arbitrary (McDougall et al., 1999). Additionally, the symbols used in process modeling can be characterized as abstract and not concrete since they mainly use features such as shapes and arrows (McDougall et al., 1999).

## 2.3. Comparing Symbol Sets of Different Process Modeling Notations

In general, process modeling notations “tend to emphasize diverse aspects of processes, such as task sequence, resource allocation, communications, and organizational responsibilities” (Soffer & Wand, 2007, p. 176). Nevertheless, when it comes to visualizing process flows, most techniques share a basic set of consensual elements. Figure 2 depicts selected symbol sets that are considered in this paper and derived from the process modeling notations of EPC, UML Activity Diagrams, YAWL, and BPMN.  $S_{EPC}$  uses hexagons as start and end nodes.  $S_{UML}$  uses filled circles instead. The end node has an additional surrounding circle.  $S_{YAWL}$  employs an audio player metaphor, the start being a circle with a right-pointing arrowhead and the end showing a circle with a rectangle in it.  $S_{BPMN}$  uses circles, with the end node having a thicker outline than the start node. In  $S_{EPC}$ , the hexagons are also used to represent any intermediate events; those conditions of the sequence flow are only represented by text without any symbol in the other notations. All notations use rectangles for capturing tasks. The routing elements are similar in  $S_{EPC}$ : AND nodes define parallel execution using a circle with the logical symbol for “and” in it. XOR nodes represent alternative branches with a circle and x symbol.  $S_{UML}$  has quite different symbols for these concepts: AND is depicted as a filled bar, while XOR is represented by a diamond-shaped symbol.  $S_{YAWL}$  uses small rectangles with inscribed triangles. In the AND node, the triangle points inward, in the XOR node outward.  $S_{BPMN}$  employs diamond symbols for both node types, using a plus symbol for the AND. The previously identified aspects permit us to discuss potential notational deficiencies in terms of visual representation of the four symbol sets.



### 2.3.1. Semiotic Clarity

Concerning semiotic clarity,  $S_{EPC}$  is the only set that does not use explicit start and end symbols; models start and end via events instead. The representation of events lacks semiotic clarity because the event symbol is overloaded. In  $S_{EPC}$ , event symbols represent start and end events as well as intermediate events (such as conditions) during a process. Such overloading may result in problems with respect to model clarity.

### 2.3.2. Perceptual Discriminability

Concerning the discriminability of routing elements,  $S_{UML}$  distinguishes very clearly between concurrency and alternative branching by using significantly different symbols. The AND and the XOR elements differ in terms of shape and line strength. At first sight,  $S_{YAWL}$  appears weak in discriminating routing elements: their semantics are determined not only by the symbols themselves but also by the position on the split/join block along with the number of incoming and outgoing arcs to the connected tasks. In essence, the AND and XOR symbols only differ in the orientation of the small triangle in this block. In the symbol sets  $S_{BPMN}$  and  $S_{EPC}$ , split and join nodes as well as decision and merge nodes share the same shape, distinguished by a symbol. Discriminability of start and end nodes in  $S_{UML}$  and  $S_{BPMN}$  is weak: all symbols are circles, and only line thickness differs. It is worth to note that none of these popular notations discriminates explicitly between split and join nodes. Whether a node is a split or a join has to be inferred from the number of incoming and outgoing arcs.

### 2.3.3. Visual Expressiveness

The visual expressiveness of all symbol sets is somewhat limited: mostly shape and size (arrows, different quadrangles, circles) are used to distinguish between symbols. No color is specified in the definition of the notations. However, several tools have augmented EPCs and BPMN with color, but not according to a particular standard. Most notably, events and functions are often shown in red and green, following the style used by ARIS Business Architect software tool. Still, even for this variant of the EPC notation, the routing elements are all uniformly shaded grey such that they cannot be distinguished by color.

### 2.3.4. Semantic Transparency

Most elements of the four symbol sets are very abstract such that there is little transparency and intuition concerning the semantic meaning of the symbols. The only exemptions are the start and end symbols in  $S_{YAWL}$ . Because they make use of an audio player metaphor, these elements should be easy to identify and intuitive as to their semantics.

### 2.3.5. Dual Coding

The dual coding principle states that using text and symbols together supports comprehension better than using either on their own. All notations investigated in this paper support this principle. However, compared to the other symbol sets, which place text inside task symbols,  $S_{YAWL}$  locates text and symbols next to each other.

Altogether, we observe that the notational deficiencies of the four symbol sets identified above are of different importance to the understanding of a process model (see Table 2).  $S_{UML}$  and  $S_{BPMN}$  merely suffer from deficiencies with their start and end elements. This deficiency has only a local impact because most process models have one start and one end element, and they can easily be identified based on the overall visual layout of a model. Also the interpretation of the control flow is not affected. Therefore, we classify these notations as having minor deficiencies. In general, we could identify only minor differences in the notations' support of semantic transparency, dual coding, and visual expressiveness. However, we found global deficiencies with semiotic clarity and perceptual discriminability.  $S_{EPC}$  shows problems with semiotic clarity of events. They are used as start and end symbols, but also for defining conditions. Accordingly, since events are distributed over all parts of  $S_{EPC}$  models, these deficiencies have a global impact on the model. In  $S_{YAWL}$ , the deficiencies relate to the perceptual discriminability of routing symbols. As they can be used at any place in the model, this deficiency has a global impact on  $S_{YAWL}$  models. Furthermore, the weak discriminability of the routing elements has a direct impact on the understanding of the behavior of a process model. Therefore, we consider this to be another serious deficiency that we observe for the symbol sets of the four business process modeling notations. In summary, we identified global deficiencies for the symbol set  $S_{EPC}$  concerning semiotic clarity and for the symbol set  $S_{YAWL}$  concerning perceptual discriminability.

**Table 2. Characterization of Notational Deficiencies**

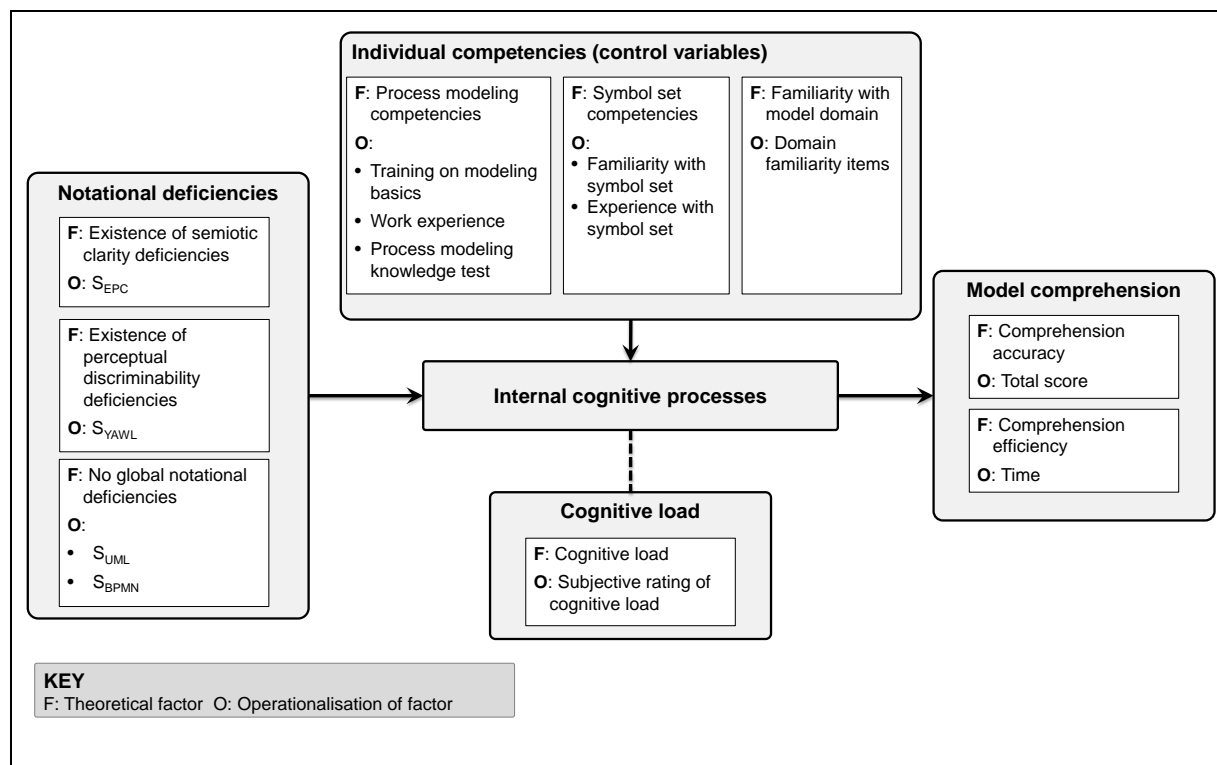
Symbol set	Type of deficiency	Span of deficiencies
$S_{UML}$ and $S_{BPMN}$	Perceptual discriminability (start and end)	Local impact at start and end (Absence of global deficiencies)
$S_{EPC}$	Semiotic clarity (events)	Global impact across model (Existence of global deficiencies)
$S_{YAWL}$	Perceptual discriminability (AND, XOR)	Global impact across model (Existence of global deficiencies)

## 3. Hypotheses Development

Against the theoretical background discussed above, this paper argues that the visual design of a process modeling notation's symbol set impacts the users' ability to understand a process that is modeled with the respective symbol set. The limited capacity of human working memory constitutes a bottleneck for cognitive activities involved in understanding process models, and the way information is represented via a specific symbol set may place extra cognitive load on the user. These considerations lead to our research model as depicted in Figure 3. The model argues that the existence of global notational deficiencies (in our case: semiotic clarity and perceptual discriminability deficiencies) in a process model negatively affect users' internal cognitive processes when trying to understand the model. We conceptualize the presence of global notational deficiencies based on weaknesses of symbols which can be used at different positions in the process model. This presumable effect of global deficiencies is backed up by cognitive load theory. In the case of local notational deficiencies, only elements on the border of the process model are affected. Global deficiencies are potentially relevant for a greater number of elements in solving comprehension tasks. Since cognitive load depends on the number of elements users need to pay attention to at the same time (Kirschner, 2002), the extent to which a specific model can be affected by a deficiency determines the basic cognitive load involved in a comprehension task.



Both identified types of deficiencies (i.e., concerning semiotic clarity and perceptual discriminability) presumably increase the extraneous cognitive load. Prior research has demonstrated that problems with semiotic clarity such as construct excess (Bodart, Patel, Sim, & Weber, 2001; Gemino & Wand, 2005) or construct overload (Shanks, Tansley, Nuredini, Tobin, & Weber, 2008) may increase cognitive load and lead to comprehension problems. On the other hand, lack of perceptual discriminability suggests that it will be more difficult and slower for model readers to perceptually process and differentiate the different visual components of a model which, again, leads to an increased cognitive load (Moody, 2009). Therefore, the experimental variation of extraneous cognitive load is expected to influence speed and accuracy of understanding (Moody, 2004). The effect on users' cognitive comprehension processes can also be described in terms of comprehension effectiveness (comprehension score), efficiency (time) and cognitive load (subjective rating of cognitive load). Additionally, individual competencies of users concerning process modeling in general and each specific symbol set along with the domain of a model in particular (Lowe, 1989; Winn, 1993) may improve understanding and information extraction when reading a model. Therefore, we consider them relevant control variables in the research model.



**Figure 3. Research Model**

Based on this research model, we state the main research hypothesis: that the existence of global notational deficiencies with respect to the semiotic clarity and perceptual discriminability of a particular symbol set influences the cognitive effectiveness of a process modeling notation. By the same token, cognitively inefficient design will impair comprehension of a process model due to increased cognitive load for the user.

**H1a:** *Semiotic clarity deficiencies of a symbol set negatively affect process model comprehension accuracy.*

**H1b:** *Perceptual discriminability deficiencies of a symbol set negatively affect process model comprehension accuracy.*

A visual notation that communicates the meaning of a process more efficiently should enable model viewers to understand models faster and to solve comprehension tasks more rapidly due to lower cognitive load. Therefore, we hypothesize:

**H2a:** *Semiotic clarity deficiencies of a symbol set negatively affect process model comprehension efficiency.*

**H2b:** *Perceptual discriminability deficiencies of a symbol set negatively affect process model comprehension efficiency.*

Furthermore, we expect model viewers to subjectively experience additional cognitive load caused by notational deficiencies. Maes and Poels (2007, p. 708) define the perceived ease of understanding a model as “the degree to which a person believes that using a conceptual modeling script ... would be free of mental effort”. In addition to process comprehension and the time needed, which can be quantified more easily, we are also interested in measuring subjective cognitive load. This is of specific interest because we argue that further implications for comprehension and time efficiency are actually caused by differences in cognitive load. Therefore, we state the next hypotheses:

**H3a:** *Semiotic clarity deficiencies of a symbol set negatively affect the users’ subjective cognitive load.*

**H3b:** *Perceptual discriminability deficiencies of a symbol set negatively affect the users’ subjective cognitive load.*

We use an experimental design to test these hypotheses.

## 4. Research Method

### 4.1. Experimental Design

To test our hypotheses, we chose an experiment that allows for controlling external factors that might distort any impact of the alternative symbol sets on how well users understand processes to assure the internal validity of our study. We measured the cognitive effectiveness of process modeling notations using three dimensions, as detailed in our research model: (1) model comprehension, (2) perceived cognitive load and (3) time taken for comprehension tasks. The outcome of our main dependent variable “model comprehension” is cognitive per se and can thus only be measured indirectly (e.g. using problem-solving tasks or comprehension tests) (Gemino & Wand, 2004). These two levels of measurements have also been referred to as “deep-level understanding” and “surface-level-understanding” (Moody, 2004). Because our research focuses on the effect of symbol design in models that include equivalent information, surface-level model comprehension tasks are most appropriate. In comparison to deep-level tasks that may interact with existing knowledge schemas of participants, surface-level tasks measure comprehension of models more directly (Parsons & Cole, 2005). Another reason for choosing surface-level comprehension tasks is that the general interpretability of models is the basis for a variety of more specific tasks such as process analysis or redesign (Burton-Jones, Wand, & Weber, 2009). The comprehension questions we used in our experiment were directly related to the models and therefore have high face validity, which assured construct validity.

The experiment followed a between-groups design. The main factor symbol set has four levels: the symbol sets derived from the process modeling notations EPC, UML Activity Diagrams, YAWL, and BPMN. This paper therefore falls within what is referred to as intra-grammar research (Gemino & Wand, 2004).

To manipulate the variable “symbol set” in the experiment, we used a limited set of symbols containing the main symbols of the process modeling notations investigated (Mending, 2008). To allow for high experimental equivalence (Parsons & Cole, 2005) among different study groups we constructed the models as follows: we transferred the models from one symbol set to another by

exchanging symbols without adhering to specific syntax restrictions of different modeling notations (see Figure 4). This meant that the  $S_{EPC}$  models did not follow a strict alternation of tasks and events; otherwise, the  $S_{EPC}$  models would have been larger than the others. Indeed, this EPC syntax rule has been criticized as it leads to unnecessarily large models. In practice, such trivial events are therefore often omitted. This practice is also adopted in various research papers (e.g., Bögl, Schreff, Pomberger, & Weber, 2009; Dollmann et al., 2011; Gersch, Hewing, & Schöler 2011). Because the number of elements is considered a significant confounding variable (Reijers & Mendling, 2011), we had to maintain a constant quantity of symbols in all versions of the models. This in turn meant that the routing element symbols in the  $S_{YAWL}$  models were not attached to other activities. To mitigate further confounding, layout was kept similar with the help of superimposing printing techniques. Also font size was the same for each model in all notations.

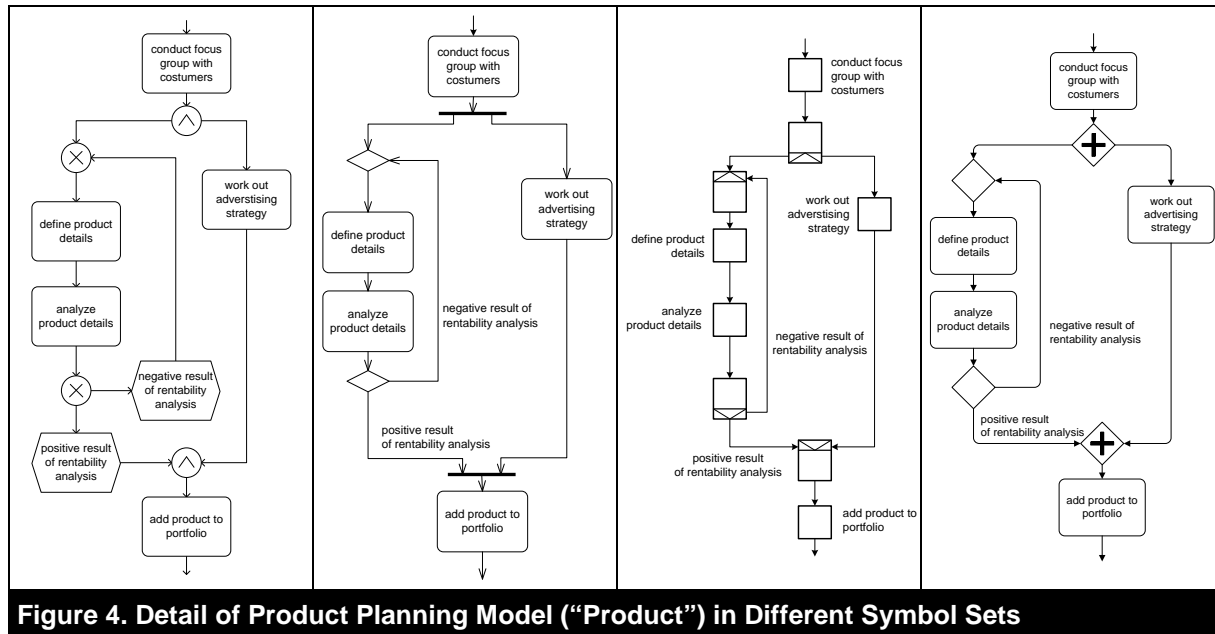


Figure 4. Detail of Product Planning Model ("Product") in Different Symbol Sets

## 4.2. Materials

We used a paper questionnaire with four different sections for the experiment (please contact authors for more information about this). The first section comprised questions about the participants' demographic data, academic qualifications and modeling experience. We also asked about their experience with the four process modeling notations, which served as a basis for the symbol sets investigated in the study. Additionally, we used the 3-item scale of Recker and Dreiling (2007) to determine the familiarity of the study group with the respective process modeling notation.

To complement the participants' subjective ratings of their experience with process models with an objective measurement based on Mendling and Strembeck (2008), the second section of the questionnaire included a test with eight items on general knowledge about process modeling.

The third section, which was followed immediately by the practical part of the questionnaire, was a tutorial on the relevant process modeling notation, which was specifically tailored to inform participants about the meaning of each symbol and covered everything the participants needed to know to perform the subsequent comprehension tasks.

The last section of the questionnaire displayed three different models and the corresponding comprehension tasks. We employed a small model on curriculum development ("curriculum", 10 tasks) and a large model on an email election process ("election", 21 tasks); each included 12 comprehension questions. The comprehension task used for the third model, on product planning ("product", 8 tasks), was a comparison of the model with a text and the identification of deviations.

The domains were chosen because they do not require specialized knowledge. We believe that designing example models instead of using real-world process models stemming from practice provided a controlled setting for testing theory-based assumptions. In the comprehension questions, participants had a choice of “right”, “wrong” or “I don’t know”. For the smaller model, the questions addressed all tasks approximately twice, and in the larger models once. The option “I don’t know” was included to reduce the probability of guessing. To measure subjective cognitive load, we included a 7-point single-item measure accompanying each question as proposed by Marcus, Cooper, and Sweller (1996).

### 4.3. Participants

We chose to focus on business school students because they are the future users of business process models. The population of interest for final data collection comprised participants who were already familiar with modeling and participants without prior modeling knowledge. In this way, we take expert-novice differences (see, for example, Petre, 1995) into account and the sample resembles potential users in practice. Therefore, participants were recruited from five different classes from information systems and business curricula with and without prior training in modeling. To assure sufficient motivation during the experiment, participants received approximately 5 percent course credit for this task.

A total of 188 students (100 males, 88 females) participated in the study, resulting in 45-50 participants per group. According to a first screening, experience with the process modeling notations was unbalanced in the dataset. This was because some notations are more widespread than others and are included in curricula to different degrees. Therefore, we decided to use only those datasets, in which participants did not have any prior experience (i.e., they had never modeled or read a model) with the symbol set used in their experimental group to avoid an experimental bias of prior experience. One hundred and thirty-six datasets remained, resulting in 28-40 participants per group. Of all participants, 29 percent had prior training in modeling basics at school or university. Table 3 summarizes other key demographics. To screen for possible differences between the experimental groups’ demographic variables, we calculated ANOVAs in addition to a visual check (also for familiarity with the model domains, which is not included in the table). Results did not hint at differences, except for the process modeling test score, which we include as a covariate in the analyses.

**Table 3. Participants’ Demographic Data**

	S <sub>UML</sub> (n=29)		S <sub>BPMP</sub> (n=39)		S <sub>YAWL</sub> (n=40)		S <sub>EPC</sub> (n=28)		Total (n=136)		ANOVA
	Mean/ number	SD/%	Mean/ number	SD/%	Mean/ number	SD/%	Mean/ number	SD/%	Mean/ number	SD/%	
Age	21.07	2.28	21.87	4.28	22.15	3.45	20.32	2.07	21.46	3.32	n.s.
Gender											
Male	10	35%	20	51%	21	53%	13	46%	64	47%	
Female	19	66%	19	49%	19	48%	15	54%	72	53%	
Highest degree completed											n.s.
High school	27	93%	26	67%	25	63%	24	86%	102	75%	
Bachelor	-	-	11	28%	11	28%	4	14%	26	19%	
Master	22	7%	2	5%	3	8%	-	-	7	5%	
Work experience in the IT sector	2	7%	5	13%	8	20%	1	4%	16	12%	n.s.
Work experience with process models	-	-	3	8%	4	10%	-	-	7	5%	n.s.
Training on modeling basics at university or school	4	14%	8	21%	13	33%	4	14%	29	21%	n.s.
Process modeling test score	0.33	0.25	0.43	0.30	0.52	0.27	0.33	0.26	0.41	0.27	F <sub>df=3,127</sub> =3.80, p=0.01

## 4.4. Procedures

To avoid any order effects (e.g., due to fading attention), we used two different samplings of the material in which models and comprehension questions were arranged in a different sequence. Due to the use of different scramblings, we could also ensure that participants could not copy correct answers from others sitting next to them. Participants were randomly assigned to one of the 8 different questionnaires (four treatments in two different sampling versions each). Subjects were allowed to spend as much or as little time as desired on the questionnaire. On average, the experiment took about 50 minutes to complete.

## 5. Results

### 5.1. Reliability and Validity Assessment

Reliability is typically approximated based on the internal consistency of a measurement instrument. To be able to consider items as unidimensional and combine them in an index, Cronbach's  $\alpha$  should equal or be greater than 0.7 (Nunnally & Bernstein, 1994). In our experiment, Cronbach's  $\alpha$  for the general knowledge test on process modeling is 0.74. Additionally, we calculated Cronbach's  $\alpha$  for the 12 cognitive load items of the comprehension tasks relating to the models "curriculum" ( $\alpha=0.96$ ) and "election" ( $\alpha=0.95$ ). The results suggest that reliability is adequate. Deletion of any item produced no marked effect on the reliability score. In light of these results, we retained all items.

### 5.2. Tests of Hypotheses

In this section, we report the results of testing our sets of hypotheses. For each main hypothesis and the three comprehension tasks (small model: "curriculum", large model: "election", text-model comparison: "product"), we ran a univariate ANCOVA with "perceptual discriminability deficiencies" and "semiotic clarity deficiencies" as independent factor, respectively. In addition to using two types of notational deficiencies as predictors, prior training in modeling and prior knowledge about process modeling were used as model covariates. Depending on the hypothesis to be tested, the dependent variables were process "model comprehension", "time taken", and "subjective cognitive load".

We use the variables perceptual discriminability deficiencies and semiotic clarity deficiencies to group the different treatments. This means, the existence or absence of such a deficiency puts a symbol set in one group or another. Accordingly, we get:

- Symbol sets with deficiencies in perceptual discriminability:  $S_{YAWL}$
- Symbol sets with deficiencies in semiotic clarity:  $S_{EPC}$
- Symbol sets without global deficiencies:  $S_{BPMN}$  and  $S_{UML}$

In order to assess whether it is reasonable to combine the symbol sets  $S_{BPMN}$  and  $S_{UML}$  in one reference group, we first checked if they differ in any of the dependent variables. We conducted the same ANCOVAs as for testing our hypotheses, but interpreted Beta instead of Alpha error significance levels. In that case, the Beta error was of higher interest, because we wanted to rule out that we incorrectly accept a no-difference hypothesis for the two symbols sets without global deficiencies. In line with our expectations the lowest Beta error was 0.322. Because beta levels are commonly set at  $p=.20$  (Ellis, 2010), the results lend support to the anticipated no-difference hypothesis for  $S_{BPMN}$  and  $S_{UML}$  and validate their combination in one reference group.

Table 4 and Table 5 give an overview of the results of the ANCOVA. Figures 5, 6, and 7 show descriptive statistics for the dependent measures (score on comprehension questions, time taken, and subjective cognitive load).

We first discuss hypotheses H1a and H1b on comprehension accuracy. When analyzing the corresponding tasks, ANCOVA results indicate that the factor perceptual discriminability deficiencies



significantly influences comprehension of the large model ( $F=6.69$ ,  $p=0.01$ ) and tends to influence comprehension of the small model ( $F=4.53$ ,  $p=0.06$ ). Therefore, hypothesis H1a, which predicted that perceptual discriminability deficiencies of a symbol set negatively affect process model comprehension, is supported. H1b is partly supported because semiotic clarity deficiencies did have a statistically significant influence on comprehension of the large model ( $F=7.03$ ,  $p=0.01$ ), but not on comprehension of the small model.

Comprehension in the third task (text-model comparison) was neither affected by perceptual discriminability deficiencies nor by semiotic clarity deficiencies.

Hypotheses H2a and H2b relate to answer time. Regarding the time taken for the comprehension questions, the factor perceptual discriminability deficiencies has a significant effect for the task with the large model ( $F=7.02$ ,  $p=0.01$ ) and the text-model comparison ( $F=5.17$ ,  $p=0.03$ ). For the symbol sets with perceptual discriminability deficiencies, time taken by participants increased by an average of over 1 minute compared to the symbol sets without global notational deficiencies. This supports hypothesis H2a. H2b is not accepted because semiotic clarity deficiencies were only found to have a trendwise influence on the time in one of three tasks (text-model comparison).

Concerning hypotheses H3a and H3b on subjective cognitive load, the participants' perceptions trendwise differ depending on the presence of semiotic clarity deficiencies in all three tasks (small model:  $F=2.89$ ,  $p=0.09$ , large model:  $F=5.01$ ,  $p=0.07$ , text-model comparison:  $F=2.89$ ,  $p=0.09$ ). Participants indicated that symbol sets with semiotic clarity deficiencies imposed higher cognitive load than symbol sets without global deficiencies. There was also a significant effect of perceptual discriminability deficiencies on cognitive load for the tasks with the small model ( $F=3.97$ ,  $p=0.05$ ) as well as the large model ( $F=5.26$ ,  $p=0.02$ ). These results are in line with our expectation that existence of global deficiencies would influence the cognitive load for users, which supports hypothesis H3a and H3b.

**Table 4. Experimental Results: Influence of Perceptual Discriminability Deficiencies**

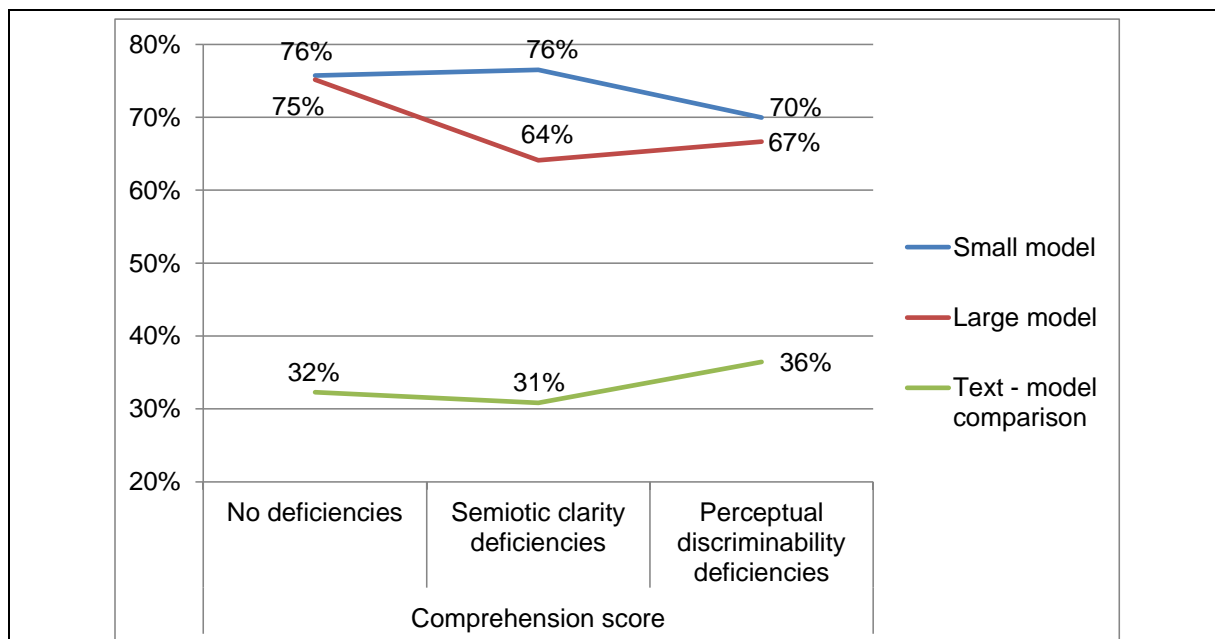
		Effect	Type III sum of squares	F ( $df_{Hypothesis}$ - $df_{Error}$ )	Significance	Partial eta squared
Comprehension accuracy (Total Score)	Small Model	Training on modeling basics	0.13	3.98 (1; 95)	<b>0.05</b>	0.04
		Perceptual discriminability deficiencies	0.12	4.53 (1; 95)	<b>0.06</b>	0.04
	Large Model	Perceptual discriminability deficiencies	0.19	6.69 (1; 94)	<b>0.01</b>	0.07
Comprehension efficiency (Time)	Large Model	Perceptual discriminability deficiencies	48.83	7.02 (1; 86)	<b>0.01</b>	0.08
	Text-Model Comparison	Perceptual discriminability deficiencies	32.24	5.17 (1; 84)	<b>0.03</b>	0.06
Subjective cognitive load	Small Model	Training on modeling basics	5.19	4.24 (1; 86)	<b>0.04</b>	0.05
		Perceptual discriminability deficiencies	4.86	3.97 (1; 86)	<b>0.05</b>	0.04
	Large Model	Process modeling knowledge	5.00	3.65 (1; 82)	<b>0.06</b>	0.04
		Perceptual discriminability deficiencies	7.21	5.26 (1; 83)	<b>0.02</b>	0.06

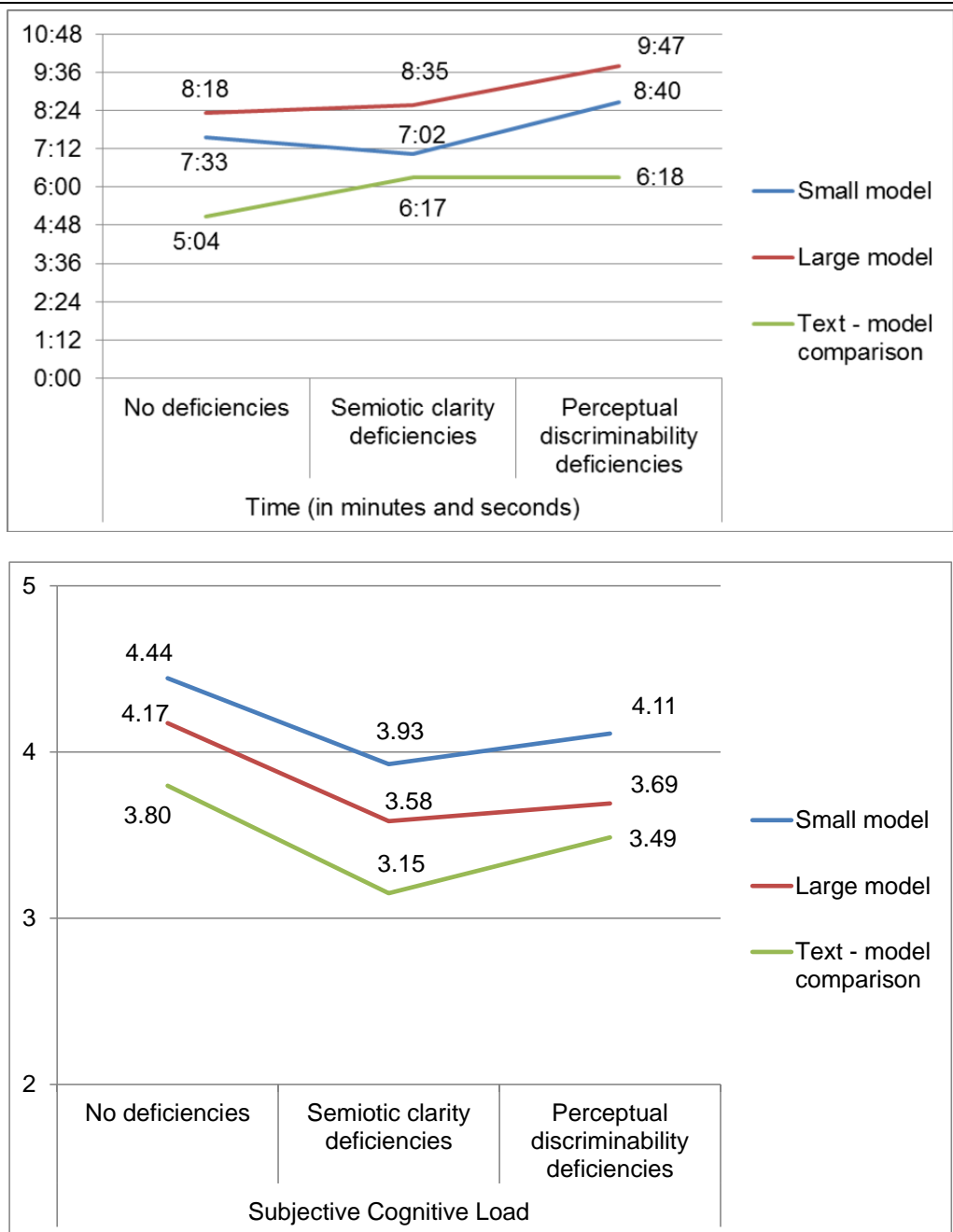
**Table 5. Experimental Results: Influence of Semiotic Clarity Deficiencies**

		Effect	Type III sum of squares	F (df <sub>Hypothesis</sub> -df <sub>Error</sub> )	Significance	Partial eta squared
Comprehension accuracy (total score)	Large Model	Training on modeling basics	0.13	4.34 (1; 84)	<b>0.04</b>	0.05
		Semiotic clarity deficiencies	0.22	7.03 (1; 84)	<b>0.01</b>	0.08
Comprehension efficiency (time)	Text - Model Comparison	Semiotic clarity deficiencies	27.35	3.47 (1; 79)	<b>0.07</b>	0.04
Subjective cognitive load	Small Model	Semiotic clarity deficiencies	4.30	2.89 (1; 81)	<b>0.09</b>	0.03
	Large Model	Semiotic clarity deficiencies	5.01	3.27 (1; 79)	<b>0.07</b>	0.04
	Text - Model Comparison	Semiotic clarity deficiencies	6.90	2.89 (1; 82)	<b>0.09</b>	0.03

We also investigated the effect of the different covariates. For selected tasks, we found a significant influence of the covariates “prior knowledge about process modeling” and “training on modeling basics” on comprehension accuracy and subjective cognitive load. Participants with less training and participants with less prior knowledge considered the cognitive load to be heavier and performed worse in the comprehension tasks. Overall, the effect of the covariates is consistent with previous findings (Reijers & Mendling, 2011).

Altogether, the results for the hypotheses are in line with our expectations: that the notational deficiencies influence process comprehension, subjective cognitive load, and time taken. The impact of notational deficiencies on comprehension seems to be slightly more pronounced for complex tasks because this effect was most reflected in the task with the large model “election”.

**Figure 5. The Influence of Notational Deficiencies on Comprehension Accuracy (Total Score)**



**Figure 5. The Influence of Notational Deficiencies on Comprehension Accuracy (Total Score) (cont)**

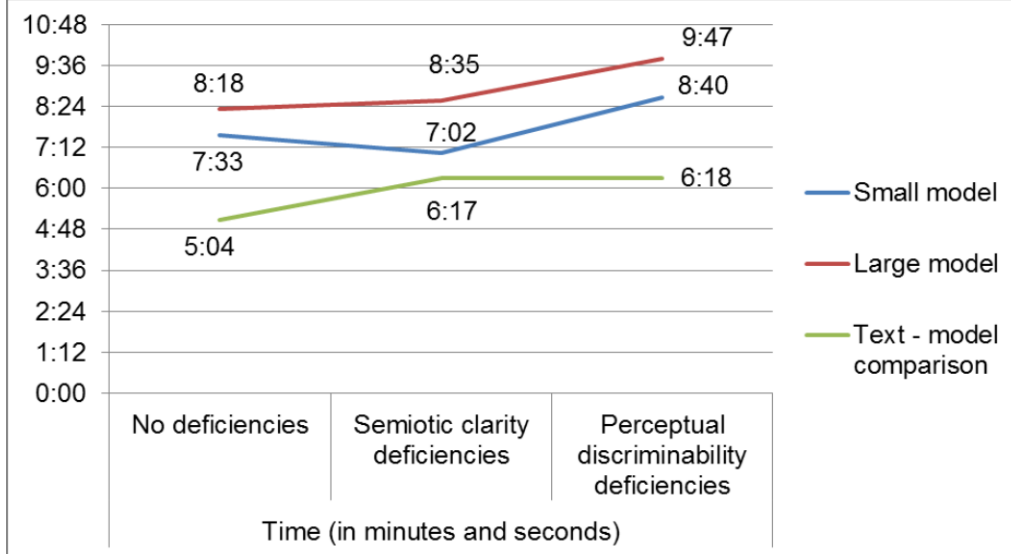


Figure 6. The Influence of Notational Deficiencies on Comprehension Efficiency (Time)

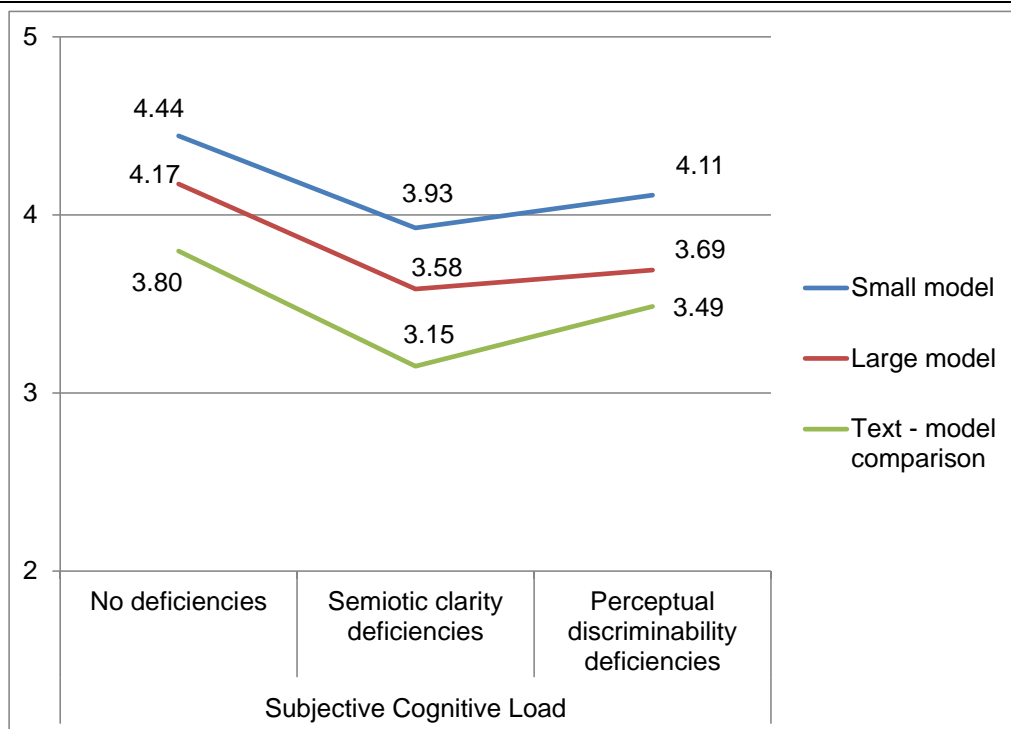


Figure 7. The Influence of Notational Deficiencies on Subjective Cognitive Load (Scale: 1-7)

## 6. Discussion

Our empirical study set out to test six hypotheses about the effects of notational deficiencies on process model comprehension. Our results are in line with the assumption that notational weaknesses may lead to increased cognitive load, which can in turn hamper comprehension. In a controlled experiment we used four different symbol sets as treatments. According to an analysis of the symbol sets, we identified global notational deficiencies in two symbol sets—perceptual discriminability deficiencies in one and semiotic clarity deficiencies in the other. The symbol sets for

that we had identified severe notational deficiencies indeed performed worse than the other symbol sets with regard to comprehension, time taken, and subjective cognitive load.

While we find that the dependent variables were affected similarly by both types of notational deficiencies—comprehension was lowered, time taken was longer, and subjective cognitive load was higher—effects were not the same for all three tasks. Because the effect of deficiencies on comprehension was strongest in the “election” task, the largest process model (which included 21 tasks), the cognitively inefficient design of symbol sets may prove problematic, especially when adding further factors that elevate cognitive load for users, such as increased complexity of processes. For the third task (text model comparison), we received no significant effect on comprehension, but on time and subjective cognitive load. Apparently, this task could be accurately solved by focusing on the textual content of the models, such that notation had not the major impact. It is interesting to see that even though the comprehension performance was not significantly affected, still the time to inspect the model and the perceived cognitive load was. This is in line with cognitive theory, which stresses the impact of visual representation on visual and working memory processing.

The findings now highlight that global deficiencies negatively affect process model comprehension. The cognitive load theory can assist in providing a suitable explanation as to why there was a measurable impact of global notational deficiencies on comprehension: apparently, cognitively inefficient symbol design impairs process model understanding as it increases extraneous cognitive load. However, cognitive load theory fails to explain all observations of the experiment. More specifically, we would like to discuss two possible interpretations as to why effects were not very pronounced. Firstly, even though different symbols were used, the representational paradigm of the process models remained unchanged (i.e., graph-based flow charts). Therefore, the experiment was unlikely to obtain huge differences between groups, such as a 100 percent solution rate for comprehension questions compared to 0 percent. This is backed by the theory of cognitive fit (Vessey, 1991), which predicts larger differences between completely different representations (e.g., text vs. models) depending on their fit with the comprehension tasks.

Secondly, we investigated notational deficiencies of symbol sets that have all been developed in academically informed initiatives and standardization committees. Many of the parties involved in the design could build on extensive practical experience. Most notably, BPMN was explicitly developed to consolidate existing process modeling languages. It has been noted that process modeling languages tend to mature with respect to their representational capabilities (Rosemann, Recker, Indulska, & Green, 2006), and it is interesting to note that such a tendency is also observable for the notational deficiencies of the four symbol sets. Moreover, it is noteworthy that we obtained significant results for our hypotheses. Thus, we think that our results support the fundamental proposition we sought to test in our research: our findings highlight the relevance of notational design in the context of conceptual models.

## 6.1. Implications for Research

The work presented in this paper has three major implications for research, in particular for the “physics of notations” and the design of experiments on model comprehension.

Our research adds to the emerging body of knowledge on notational design. The results support the usefulness of the “physics of notations” (Moody, 2009) as a framework for discussing strengths and weaknesses of a particular type of notation. However, this framework does not provide measurement instruments. Therefore, the notational analysis in this paper is restricted to qualitative classes of deficiencies. In particular, we distinguish notational deficiencies that are of local and global relevance for an individual model and which relate to perceptual discriminability and semiotic clarity. Up until now, there has been a lack of theoretical insight into which types of deficiencies impact comprehension, time, cognitive load, or any combination of them. Investigating this research question will likely require a closed-up inspection of the understanding process, either using think-aloud techniques or cognitive as much as visual monitoring. Moreover, the paper encourages the exploration of additional deficiency types to promote the design of more understandable modeling languages, as we focused on two examples, which were found in existing notations. Furthermore, it



would be highly desirable for future research to develop scales to measure each aspect of notational deficiency on a ratio level. Furthermore, the development of mechanisms to aggregate these individual measurements in a valid and reliable manner would be welcome.

Our research informs work on modeling language evaluation altogether. First, the results of this paper suggest that future research into (process) model understanding should consider symbol sets and syntactical aspects in isolation when investigating the relative superiority of different (process) modeling languages. For instance, the comparison of EPCs and Petri nets reported in (Sarshar & Loos, 2005) only assesses the overall superiority of a language. This is problematic because the results cannot be traced to either the characteristics of the symbol set or the formal syntactical rules. The two modeling languages in (Sarshar & Loos, 2005) significantly differ in both aspects, such that the relative importance of each factor for the observed effect remains unclear. Therefore, research should either study notational, syntactical, and semantic aspects of modeling languages independently or integrate them as separate treatments.

Second, our results emphasize the importance of visual discriminability. This has implications for how complex a symbol set of a notation can be. Clearly, it is easier to achieve an overall good discriminability of a notation such as EPCs with six notational elements than for BPMN with its more than 30 event types alone. It would be of fundamental interest to find a quantitative measure for visual discriminability of a symbol set, such that a maximum number of elements could be related to a certain threshold value of it.

## 6.2. Implications for Practice

The results presented in this paper are also relevant for business process modeling practice, in particular with reference to language selection and symbol set design. In addition, we believe our findings have important implications for the development of domain-specific modeling languages in general.

Business process models are an important source for the specification of software systems requirements. In fact, errors that are conducted in an early stage of an information system's development process, and are detected in a late phase of the development process, frequently cause severe problems and often even cause a development project to fail entirely (Charette, 2005; Stepanek, 2005). In this context, the selection of a business process modeling language is an important decision when setting up a process modeling project. Usability is an important dimension for selecting a process modeling tool and language due to its significance for process modeling success (Bandara et al., 2005). Our research suggests that particular attention should be paid to the usability of the symbol set. Our results confirm recommendations brought forth by the "physics of notations" (Moody, 2009). In this regard, it seems recommendable to use symbol sets without global notational deficiencies, as for instance offered by BPMN which performed significantly better in our experiment. Furthermore, the notational deficiencies of YAWL and EPCs along with the weaker performance of participants working with them in the experiment warrants improvements of the notations symbol sets. Interestingly, they have complementary strengths: while YAWL has suitable start and end elements, EPCs have nicely distinguishable routing elements. Our experimental results suggest a reworking of the symbol sets of both notations in order to improve usability. To assure adequate perceptual discriminability of symbols, inter-symbol similarity-rating matrices could for instance be used (Geiselman, Landee, & Christen, 1982).

However, we want to underline that we do not intend to make direct evaluations of the notations as we used a limited subset of the symbols offered by each notation and had to slightly deviate from the original notation proposals for the models in our experiment. In addition, modeling tools often level notational weaknesses off (e.g. by using color to increase visual discriminability of symbols). Modeling languages also differ in semantics and in the complexity of the symbol set. When the overall quality of a notation is supposed to be analyzed, these additional factors have to be taken into account. While symbols of visual modeling languages cannot be exchanged easily in practice: the usage of additional color can be a smooth mechanism to weaken the negative effects of deficiencies in perceptual discriminability. Such secondary notation (Green & Petre, 1996) is known to improve a

notation's understanding when used in a systematic way. Not only EPC modeling tools, but also BPMN modeling tools partially use different color schemes to improve model understanding. In this study, we aimed to control the effect of potential color usage. It would be desirable to investigate how color can be used systematically in notational design.

On the other hand, the development of (new) domain-specific modeling languages as another area of practice can benefit from our results. A domain-specific language is a tailor-made language for a specific problem domain (Strembeck & Zdun, 2009). The design of a domain-specific modeling language involves the definition of a suitable symbol set. Here, the recommendations of the "physics of notations" can be applied with a much higher level of design freedom as opposed to standardized notations which need to be consistent with prior versions of the standard.

## 7. Conclusion

Our experiment provides empirical evidence for the importance of symbol design in process model comprehension. It demonstrates that notational deficiencies in symbol sets may lead to heavier cognitive load for users, hampering model comprehension. We obtained support for our proposition from surface-level comprehension tasks, the subjective rating of cognitive load, and time taken. Altogether, we hope that these findings can guide future standardization efforts regarding process modeling notations. Our work makes a contribution to the literature on process model comprehension and research on cognitive load in conceptual modeling in general.

## References

- Aranda, J., Ernst, N., Horkoff, J., & Easterbrook, S. M. (2007). A framework for empirical evaluation of model comprehensibility. Paper presented at the International Workshop on Modeling in Software Engineering (MiSE-07).
- Bandara, W., Gable, G. G., & Rosemann, M. (2005). Factors and measures of business process modelling: model building through a multiple case study. *European Journal of Information Systems*, *14*, 347-360.
- Bodart, F., Patel, A., Sim, M., & Weber, R. (2001). Should optional properties be used in conceptual modelling? A theory and three empirical tests. *Information Systems Research*, *12*(4), 384-405.
- Bögl, A., Schrefl, M., Pomberger, G., & Weber, N. (2009). Automated construction of process goal trees from EPC-models to facilitate extraction of process patterns. In J. Filipe & J. Cordeiro (Eds.), *Enterprise information systems* (Vol. 24, pp. 427-442). Berlin: Springer Berlin Heidelberg.
- Burton-Jones, A., Wand, Y., & Weber, R. (2009). Guidelines for empirical evaluations of conceptual modeling grammars. *Journal of the Association for Information Systems*, *10*(6), 495-532.
- Canfora, G., Garc, F., Piattini, M., Ruiz, F., & Visaggio, C. A. (2005). A family of experiments to validate metrics for software process models. *Journal of Systems and Software*, *77*(2), 113-129.
- Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology*, *10*(2), 151-170.
- Charette, R. N. (2005). Why software fails. *IEEE Spectrum*, *42*(9), 42-49.
- Darke, P., & Shanks, G. (1997). User viewpoint modelling: Understanding and representing user viewpoints during requirements definition. *Information Systems Journal*, *7*(3), 213-239.
- Dollmann, T. J., Loos, P., Fellmann, M., Thomas, O., Hoheisel, A., Katranuschkov, P., Scherer, R. (2011). Design and usage of a process-centric collaboration methodology for virtual organizations in hybrid environments. *International Journal of Intelligent Information Technologies*, *7*(1), 45-64.
- Dumas, M., Aalst, W. M. P. v. d., & Hofstede, A. H. M. t. (2005). *Process aware information systems: Bridging people and software through process technology*. Hoboken, New Jersey: John Wiley & Sons.
- Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. United Kingdom: Cambridge University Press.
- Figl, K., Derntl, M., Rodriguez, M. C., & Botturi, L. (2010). Cognitive effectiveness of visual instructional design languages. [doi: DOI: 10.1016/j.jvlc.2010.08.009]. *Journal of Visual Languages & Computing*, *21*(6), 359-373.
- Figl, K., Mendling, J., Strembeck, M., & Recker, J. (2010). *On the cognitive effectiveness of routing symbols in process modeling languages*. Paper presented at the Business Information Systems (BIS), Berlin.
- Geiselman, R. E., Landee, B. M., & Christen, F. G. (1982). Perceptual discriminability as a basis for selecting graphic symbols. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *24*(3), 329-337.
- Gemino, A., & Wand, Y. (2004). A framework for empirical evaluation of conceptual modeling techniques. *Requirements Engineering*, *9*(4), 248-260.
- Gemino, A., & Wand, Y. (2005). Complexity and clarity in conceptual modeling: Comparison of mandatory and optional properties. *Data & Knowledge Engineering*, *55*(3), 301-326.
- Genon, N., Amyot, D., & Heymans, P. (2010). *Analysing the cognitive effectiveness of the UCM visual notation*. Paper presented at the System Analysis and Modeling Workshop.
- Genon, N., Heymans, P., & Amyot, D. (2010). *Analysing the cognitive effectiveness of the BPMN 2.0 visual syntax*. Paper presented at the Software Language Engineering.
- Gersch, M., Hewing, M., & Schöler, B. (2011). Business process blueprinting – an enhanced view on process performance. *Business Process Management Journal*, *17*(5), 732-747.
- Green, T. R. G., & Petre, M. (1996). Usability analysis of visual programming environments: A “cognitive dimensions” framework. *Journal of Visual Languages and Computing*, *7*(2), 131-174.
- Gruhn, V., & Laue, R. (2009). *Reducing the cognitive complexity of business process models*. Paper presented at the International Conference on Cognitive Informatics, Hong Kong.

- Hitchman, S. (2002). The details of conceptual modelling notations are important - a comparison of relationship normative language. *Communications of the Association for Information Systems*, 9(10), 167-179.
- Keller, G., Nüttgens, M., & Scheer, A.-W. (1992). *Semantische prozessmodellierung auf der grundlage "ereignisgesteuerter prozessketten (EPK)"*. Saarbrücken, Germany: Institut für Wirtschaftsinformatik.
- Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, 12(1), 1-10.
- Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, 17(2), 248-294.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11(1), 65-100.
- Lohmann, N., Verbeek, E., & Dijkman, R. (2009). Petri net transformations for business processes – a survey. In J. Kurt & M. A. Wil (Eds.), *Transactions on petri nets and other models of concurrency II* (pp. 46-63). Heidelberg: Springer-Verlag.
- Lowe, R. K. (1989). Search strategies and inference in the exploration of scientific diagrams. *Educational Psychology*, 9(1), 27-44.
- Maes, A., & Poels, G. (2007). Evaluating quality of conceptual modelling scripts based on user perceptions. *Data & Knowledge Engineering*, 63(3), 701-724.
- Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, 88(1), 49-63.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge, UK: Cambridge University Press.
- McDougall, S. J. P., Curry, M. B., & Bruijn, O. D. (1999). Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols. *Behavior Research Methods, Instruments, & Computers*, 31(3), 487-519.
- Melão, N., & Pidd, M. (2000). A conceptual framework for understanding business processes and business process modelling. *Information Systems Journal*, 10(2), 105-129.
- Mendling, J. (2008). *Metrics for process models: Empirical foundations of verification, error prediction, and guidelines for correctness* (Vol. 6). Germany: Springer.
- Mendling, J., Recker, J., & Reijers, H. A. (2010). On the usage of labels and icons in business process modeling. *International Journal of Information Systems Modeling and Design*, 1(2), 40-58.
- Mendling, J., & Strembeck, M. (2008). Influence factors of understanding business process models. In W. Abramowicz & D. Fensel (Eds.), *Business information systems* (Vol. 7, pp. 142-153). Germany: Springer.
- Moody, D. L. (1998). *Metrics for evaluating the quality of entity relationship models*. Paper presented at the Proceedings of the 17th International Conference on Conceptual Modeling.
- Moody, D. L. (2004). *Cognitive load effects on end user understanding of conceptual models: An experimental analysis*. Paper presented at the 8th East European Conference on Advances in Databases and Information Systems.
- Moody, D. L. (2009). The "physics" of notations: Towards a scientific basis for constructing visual notations in software engineering. *IEEE Transactions on Software Engineering*, 35(5), 756-779.
- Moody, D. L., Heymans, P., & Matulevicius, R. (2010). *Improving the effectiveness of visual representations in requirements engineering: An evaluation of i \* visual syntax*. Paper presented at the 17th IEEE International Requirements Engineering Conference.
- Moody, D. L., & Hillegersberg, J. (2008). *Evaluating the visual syntax of UML: An analysis of the cognitive effectiveness of the UML family of diagrams*. Paper presented at the Software Language Engineering.
- Nordbotten, J. C., & Crosby, M. E. (1999). The effect of graphic style on data model interpretation. *Information Systems Journal*, 9(2), 139-155.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Object Management Group. (2011a). Unified Modeling Language (UML), Version 2.4.1. Retrieved from <http://www.omg.org/spec/UML/2.4.1/>
- Object Management Group. (2011b). Business Process Model and Notation (BPMN), Version 2.0. Retrieved from <http://www.omg.org/spec/BPMN/2.0/>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255-287.

- Parsons, J., & Cole, L. (2005). What do the pictures mean? Guidelines for experimental evaluation of representation fidelity in diagrammatical conceptual modeling techniques. *Data and Knowledge Engineering*, 55(3), 327-342.
- Petre, M. (1995). Why looking isn't always seeing: Readership skills and graphical programming. *Communications of the ACM*, 38(6), 33-44.
- Recker, J., & Dreiling, A. (2007). *Does it matter which process modelling language we teach or use? An experimental study on understanding process modelling languages without formal education*. Paper presented at the 18th Australasian Conference on Information Systems.
- Recker, J., Rosemann, M., Indulska, M., & Green, P. (2009). Business process modeling – a comparative analysis. *Journal of the Association for Information Systems*, 10(4), 333-363.
- Reijers, H. A., & Mendling, J. (2011). A study into the factors that influence the understandability of business process models. *IEEE Transactions on Systems, Man, and Cybernetics - Part A*, 41(3), 449-462.
- 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.
- Rosemann, M. (2006). Potential pitfalls of process modeling: Part A. *Business Process Management Journal*, 12, 249-254.
- Rosemann, M., Recker, J., Indulska, M., & Green, P. (2006). A study of the evolution of the representational capabilities of process modeling grammars. In E. Dubois & K. Pohl (Eds.), *Advanced information systems engineering - CAiSE 2006* (Vol. 4001, pp. 447-461). Germany: Springer.
- Sarshar, K., & Loos, P. (2005, 2005). *Comparing the control-flow of EPC and petri net from the end-user perspective*. Paper presented at the 3rd International Conference on Business Process Management.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, 45(2), 185-213.
- Scheer, A. W. (2000). *ARIS - business process modeling* (3rd Ed.). Germany: Springer Verlag.
- Shanks, G., Tansley, E., Nuredini, J., Tobin, D., & Weber, R. (2008). Representing part-whole relations in conceptual modeling: An empirical evaluation. *MIS Quarterly*, 32(3), 553-573.
- Siau, K., & Tan, X. (2005). Improving the quality of conceptual modeling using cognitive mapping techniques. *Data & Knowledge Engineering*, 55(3), 343-365.
- Soffer, P., & Wand, Y. (2007). Goal-driven multi-process analysis. *Journal of the Association for Information Systems*, 8(3), 175-203.
- Stepanek, G. (2005). *Software project secrets: Why software projects fail*. New York: Apress.
- Strembeck, M., & Zdun, U. (2009). An approach for the systematic development of domain-specific languages. *Software: Practice and Experience*, 39(15), 1253-1292.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science: A Multidisciplinary Journal*, 12(2), 257-285.
- van der Aalst, W. M. P., & ter Hofstede, A. H. M. (2005). YAWL: Yet another workflow language. *Information Systems*, 30(4), 245-275.
- Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences*, 22(2), 219-240.
- Wand, Y., & Weber, R. (1993). On the ontological expressiveness of information systems analysis and design grammars. *Journal of Information Systems*, 3(4), 217-237.
- Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology*. London, UK: Routledge and Kegan Paul.
- Winn, W. (1993). An account of how readers search for information in diagrams. *Contemporary Educational Psychology*, 18(2), 162-185.



## Appendix: Experimental Material

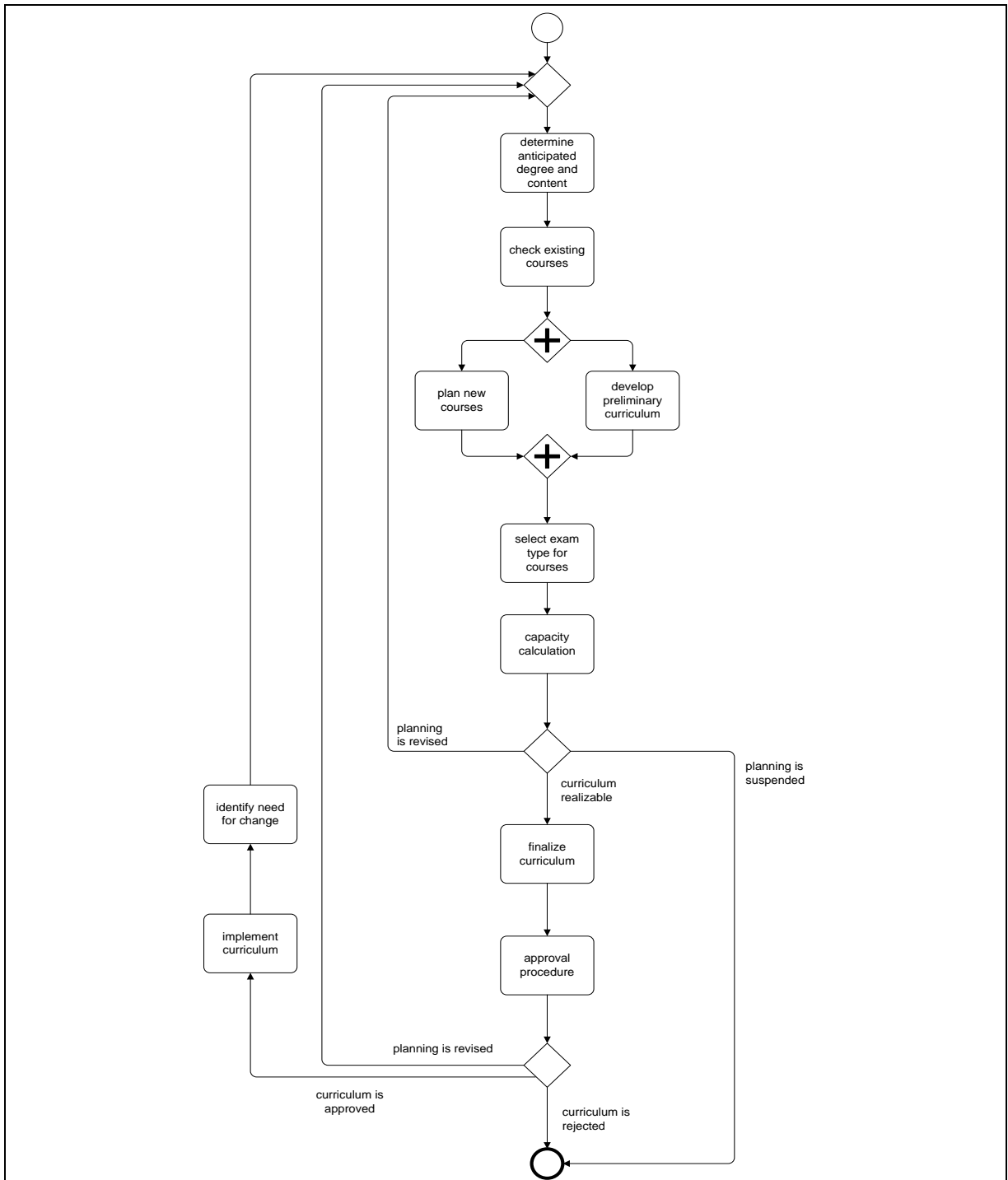


Figure A-1. Model on Curriculum Development ('Curriculum')

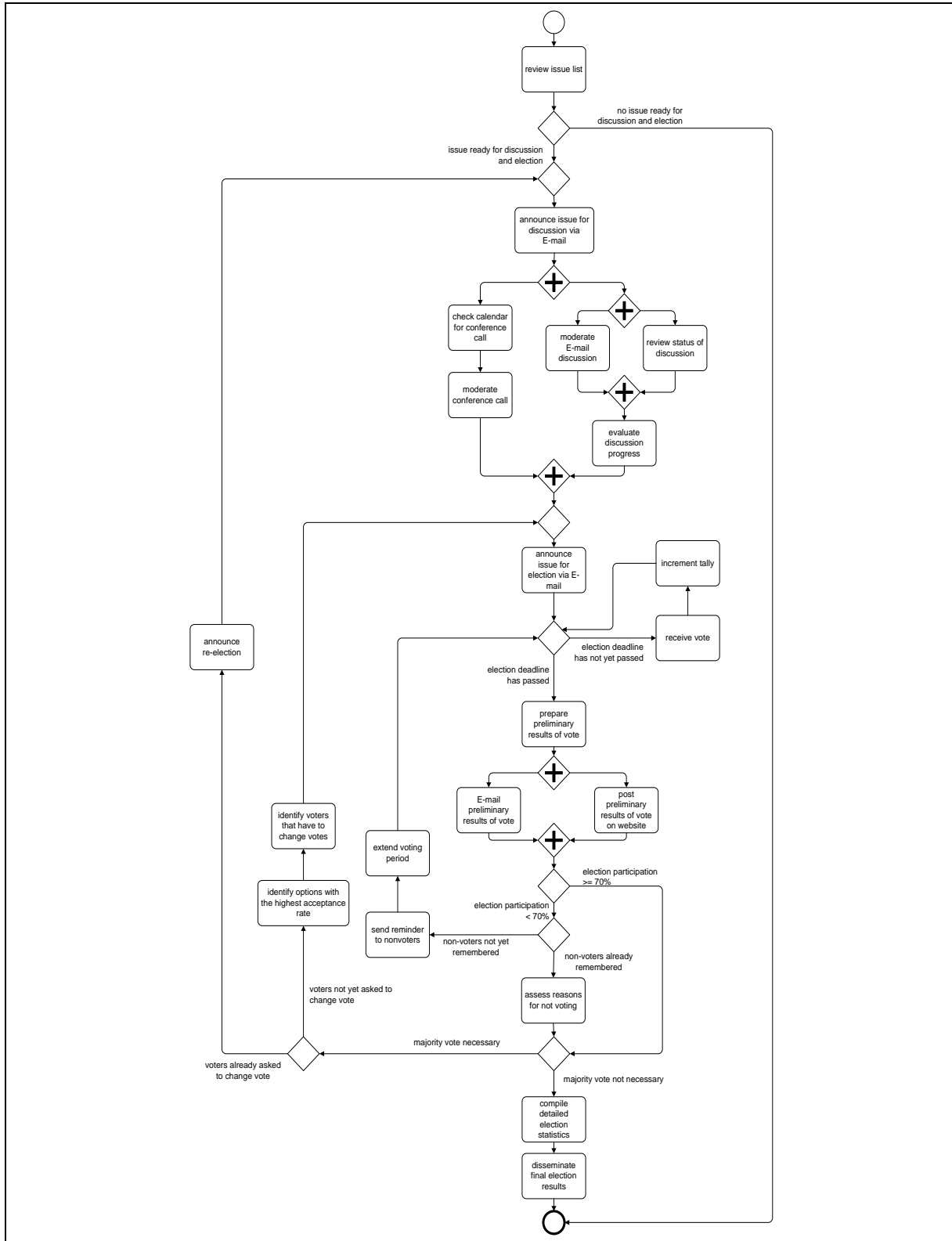


Figure A-2. Model on an E-mail Election Process ('Election')

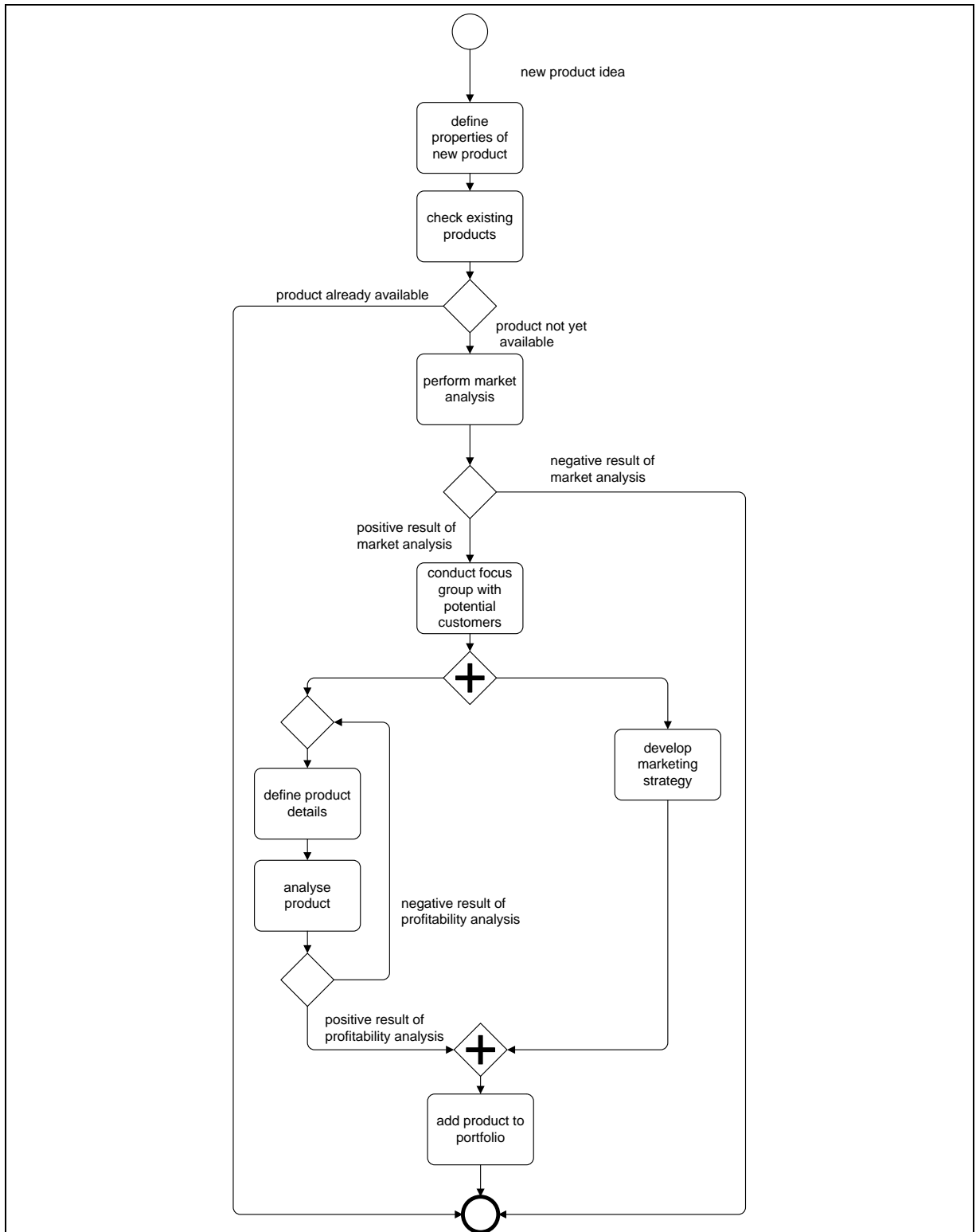


Figure A-3. Model on Product Planning ('Product')

## About the Authors

**Kathrin FIGL** is an Assistant Professor in the Institute for Information Systems and New Media at the Vienna University of Economics (WU). She received her Doctoral (awarded with the Dr. Maria Schaumayer Award) and two Master's in Information Systems and Psychology, both with honours, from the University of Vienna. Most of her research and teaching focuses on the intersection between information systems and psychology, including research on cognitive aspects of modeling, information systems education and human-computer-interaction. In 2010, she was awarded the excellent teaching award from the Vienna University of Economics for her lecture on information systems. She has published more than 50 research papers and articles, among others in the *AIS Transactions on Human-Computer Interaction*, *Decision Support Systems*, and the *Journal of Visual Languages and Computing*.

**Jan MENDLING** is a Professor with the Institute for Information Business at WU Vienna. His research areas include Business Process Management, Conceptual Modeling and Enterprise Systems. He received a PhD degree from WU Vienna (Austria), a diploma degree in Business Computer Science and in Business Administration both from the University of Trier, and held positions with QUT Brisbane (Australia) and HU Berlin (Germany). He has published more than 200 research papers and articles, a.o. in *ACM Transactions on Software Engineering and Methodology*, *IEEE Transaction on Software Engineering*, and *Decision Support Systems*. He is member of the editorial board of three international journals, one of the founders of the Berlin BPM Community of Practice ([www.bpmb.de](http://www.bpmb.de)), and board member of the Austrian Gesellschaft für Prozessmanagement.

**Mark STREMBECK** is an Associate Professor of Information Systems at the Vienna University of Economics and Business (WU Vienna), Austria. His research interests include business process management, model-driven software development, security engineering, and the modeling and management of dynamic (distributed) software systems. Among others, he has published in *ACM Transactions on Information and System Security*, *Decision Support Systems*, *IEEE Security & Privacy*, *Software: Practice & Experience*, and *Information & Software Technology*. He received his doctoral degree as well as his Habilitation degree (*venia docendi*) from WU Vienna. He is a key researcher at the Secure Business Austria Research Center (<http://www.sba-research.org/team/>), and the Vice Institute Head of the Institute for Information Systems at WU Vienna (<http://nm.wu.ac.at/>).