

The Structured Process Modeling Theory (SPMT) a cognitive view on why and how modelers benefit from structuring the process of process modeling

Jan Claes^{1,2}  · Irene Vanderfeesten² · Frederik Gailly¹ · Paul Grefen² · Geert Poels¹

Published online: 28 July 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract After observing various inexperienced modelers constructing a business process model based on the same textual case description, it was noted that great differences existed in the quality of the produced models. The impression arose that certain quality issues originated from cognitive failures during the modeling process. Therefore, we developed an explanatory theory that describes the cognitive mechanisms that affect effectiveness and efficiency of process model construction: the Structured Process Modeling Theory (SPMT). This theory states that modeling accuracy and speed are higher when the modeler adopts an (i) *individually fitting* (ii) *structured* (iii) *serialized* process modeling approach. The SPMT is evaluated against six theory quality criteria.

Keywords Business process modeling · Process of process modeling · Explanatory theory · Structured process modeling · Cognitive fit

1 Introduction

For the design and analysis of information systems for organizations, analysts typically deal with the complexity of the organization by using conceptual models. These models abstract from specific instances and represent the generic properties of the modeled system. The focus in this paper is on process models, which are considered to be a specific kind of conceptual models. A process model is a mostly graphical representation that documents the different steps that are or that have to be performed in the execution of a particular process under study, together with their execution constraints such as the allowed sequence or the potential responsible actors for these steps (Dumas et al. 2013; Weske 2007).

The recent developments in research about process models can be classified into three research streams. One stream studies the *application* of process models. For example, the construction of process models has shown to be a key success factor in process redesign (Kock et al. 2009; Xiao and Zheng 2012), software development (Krishnan et al. 1999), and communication (Abecker et al. 2000; Davies et al. 2006). Therefore it is important that the quality of process models is high.

A second research stream is thus investigating the *quality* of process models. Traditionally, it is believed that the quality of the model has to be evaluated relative to the purpose of the model (Juran and Gryna 1988; Lindland et al. 1994). An abundance of process model quality dimensions and metrics, targeted at various purposes, has thus been examined (Nelson et al. 2012; Vanderfeesten et al. 2007). For example, if the process model is created as a tool for communication about a particular process, the comprehensibility of the model by its intended readers can be regarded as an important quality dimension. In case the process model has to serve as input for a process-aware information system, syntactic correctness and

✉ Jan Claes
jan.claes@ugent.be
Irene Vanderfeesten
i.t.p.vanderfeesten@tue.nl
Frederik Gailly
frederik.gailly@ugent.be
Paul Grefen
p.w.p.j.grefen@tue.nl
Geert Poels
geert.poels@ugent.be

¹ Department of Business Informatics and Operations Management, Ghent University, Tweakerkenstraat 2, 9000 Gent, Belgium

² Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Paviljoen D13, Postbus 513, 5600 MB Eindhoven, The Netherlands

semantic completeness may be considered to be more crucial. An extensive overview of quality dimensions and related metrics is presented by (Sánchez-González et al. 2013) in their systematic literature review on process model quality research.

Recently, a third stream of process model research originated that shifts the focus from investigating *what* are characteristics of a good process model towards the study of *how* good process models are constructed. For instance, Brown et al. (2011) investigated how the use of virtual world technology increases modeler empowerment and consensual development during modeling in collaborative settings. Collaborative process modeling and how technology supports this activity is also the subject of (Recker et al. 2013). Further, Pinggera et al. (2013) identified three process modeling styles relating to variations in modeling speed and model reconciliation. Lastly, Claes et al. (2015) developed a visualization that represents how process models are created in terms of consecutive operations on the model in a modeling tool.

Similar trends of research shift existed already in the broad field of conceptual modeling (e.g., Hoppenbrouwers et al. 2005) and the even more general area of system analysis and development (e.g., Chakraborty et al. 2010; Nunamaker and Chen 1990). The underlying assumption in all of these studies is that the quality of the product depends on the quality of the process that creates the product, at least to some extent. Based on the observations described in this paper, we subscribe to this assumption and presume that certain quality concerns are caused during the modeling process. Therefore, we abstract from the different process model quality dimensions and study the cognitive mechanisms during the process of process modeling in which these quality issues originate.

We define the process of process modeling (PPM) as the sequence of steps a modeler performs in order to translate his mental image of the process into a formal, explicit and mostly graphical process specification: the process model. The modeler forms a mental representation of the process based on direct observation and/or various descriptions of the real or intended process such as interview transcripts, whiteboard notes and requirements documents (Chakraborty et al. 2010). It should also be noted that mental models are rarely stable, they keep evolving as more information is processed (Rogers and Rutherford 1992). Hence, the transformation of the (individual and dynamic) mental model into an explicit process model is a complex cognitive task. During this task the modeler iterates between shaping the mental model, evaluating the mental model, converting the mental model into a formal model, evaluating the formal model, adapting the mental model, etc.

Throughout this complex task the modeler is hindered by his cognitive limits, which results in cognitive ineffectiveness that can be manifested by a decrease of accuracy

and speed (Rockwell and Bajaj 2005; Sweller, 1988). Therefore, the end goal of our research is to help the modeler to reduce these negative effects by developing a method for process modeling that should warrant the optimal use of a modeler's cognitive functions. As advocated by (Avgerou 2000; Naumann 1986), a first, fundamental step for achieving this goal is the collection and description of the necessary knowledge that helps understand why, how and when cognitive failures occur during process modeling. Because such knowledge is not currently readily available, this paper handles entirely on its development. The presented contribution is the Structured Process Modeling Theory (SPMT), which explains why the modelers that implement an optimal structuring approach towards modeling may deal with the complexity of the task in a more cognitively effective and efficient way.

The SPMT being an explanatory theory will indeed serve as a foundation for the development of a method that prescribes how to create process models in a cognitively optimal way. Furthermore, the SPMT brings together cognitive theories about learning and problem solving in a fundamental new way and has the potential to “give back” to the cognition research field, because of the novel view on (the combined application) of these theories. This paper has practical significance by providing knowledge that can be used for process modeling training, for differentiated tool development, etc.

In Section 2, the methodology that was used to build and to test the SPMT is discussed. The theory was developed by adapting and combining cognitive theories to the context of process modeling in order to explain the varying success of different observed modeling approaches. The way these observations were collected is described in Section 3. Subsequently, Section 4 provides the theoretical background for the developed SPMT, which itself is presented in Section 5. Next, the SPMT is evaluated in Section 6. The context of the research is outlined in Section 7, which summarizes related work. Finally, Section 8 contains an extensive discussion and a brief conclusion is provided in Section 9.

2 Research methodology

Multiple research paradigms exist in Information Systems amongst which design science and behavioral science are prevalent paradigms (March and Smith 1995). Design science is centered on the development of research artifacts such as constructs, models or methods, which have to possess value or utility (Hevner et al. 2004). Behavioral science is concerned with developing knowledge about human behavior, represented by theories (Simon 1996). The selection of the research

methodology depends on the research question. Based on descriptions in the introduction, this question can be phrased as:

RQ. Why do people struggle with the complexity of constructing a process model?

The above research question asks for explanations of human behavior and thus an explanatory theory was developed that describes the cognitive leverages that play a role while constructing a process model. An explanatory theory “*provides explanations but does not aim to predict with any precision. There are no testable propositions.*” (Gregor 2006, p. 620). Other types of theories exist as well. A predictive theory for example does not provide explanations, but it does include testable propositions with predictable effects. Next to descriptive theories such as the explanatory or predictive theories, also prescriptive theories exist. Instead of only describing, explaining or predicting relations between constructs, they offer concrete prescriptions and relate the proposed actions to certain consequences (Gregor 2006).

2.1 Theory building

The input for theory development may include (objective) observations (Godfrey-Smith 2009; Nagel 1979), as well as (subjective) impressions (Popper 2005). New theory can then be developed by searching for explanations for the observations and impressions (Weick 1989). In order to collect observations and impressions about how modelers construct process models, exploratory modeling sessions were performed (see Section 3). An explanation for the observed relations between modeling approach and cognitive failures was searched for in cognitive literature. Section 4 explains how cognitive theories propose that the human brain is limited in handling complex tasks and if the brain gets overloaded, modelers tend to work slower and make more mistakes. These theories can explain the observed behavior and varying success of the modelers while constructing process models. We compiled and synthesized these theories into the central contribution of this paper: the Structure Process Modeling Theory (SPMT), presented in Section 5.

2.2 Theory testing

For most theories, the actual value can only be measured on the long term, by evaluating its actual use by others (Weick 1989). Nevertheless, in literature about theory in the information systems domain six assessable criteria for good (explanatory) theories were found: i.e., novelty, parsimony, consistency, plausibility, credibility, and transferability (Gregor 2006; Grover et al. 2008; Weber 2012; Weick

1989). Section 6 elaborates on the assessment of the SPMT against these criteria. For the evaluation of consistency, a second series of observational modeling sessions was examined in order to assess to what extent the described theory can be used to explain the additional observations.

3 Problem exploration

In order to explore how people construct process models, structured sessions were performed in which participants were asked to construct a business process model based on a given textual case description. These observational sessions supported the collection of the data that were studied in order to collect the observations and impressions that served as input for the development of the Structured Process Modeling Theory (SPMT).

3.1 Data collection method: Exploratory observational modeling sessions

During the exploratory modeling sessions it was observed how the modelers constructed a process model from a textual case description. The participants were instructed to aim for a high quality model. It was, however, not defined what was meant by ‘high quality model’.

Case The case to be modeled described the steps in the request handling of mortgages by a bank.¹ A textual description was handed over to the participants and comprised two A4 format sheets excluding instructions. The process models that were built by the participants contained on average 27 activities and construction took on average 276 recorded modeling operations in the tool (see further). This size indicates the complexity of the case and the modeling task according to (Mendling 2008), which will be further discussed in the following sections.

Participants In order to gain knowledge about how inexperienced modelers deal with the complexity of a case throughout a process-modeling endeavor, master students that attended a course in Business Process Management were selected as primary target group. The sessions were strategically planned after the lectures in which the students were introduced into process modeling, but before the training of specific modeling techniques or guidelines. This way a group was formed of participants that have enough maturity and knowledge about process modeling without possessing an abundance of modeling experience. The focus was on inexperienced modelers, because they

¹ Case description can be downloaded from <http://bpm.q-e.at/experiment/MortgageEindhoven>.

did not yet consciously learn any technique to cope with the complexity of a modeling task, which we expected to result in more variety in the observations and a more open search for potential interesting modeling approaches. The observational modeling sessions took place in December 2012 at Eindhoven University of Technology. The group of participants was composed of 118 master students in total, distributed over three different educational programs (i.e., Operations Management & Logistics, Innovation Management, and Business Information Systems). The mixture of educational profiles from technical-oriented to business-oriented students has the advantage of increasing the likelihood that a heterogeneous set of observations is obtained. Participation was voluntarily and the students could stop at any time without handing in a solution.

Modeling language A simplified modeling language was used for the modeling sessions. It contained constructs representing the main *concepts* of a control flow model²: start node, end node, activity, sequence flow, parallel branch (split and join), and optional branch (split and join). These constructs were chosen because they are found in the majority of currently used process modeling languages (e.g., BPMN, EPC, Petri-Net, UML Activity Diagrams, Workflow Net, YAWL, etc.) Moreover, they are considered the most used constructs for process modeling (Zur Muehlen and Recker 2008). The advantage of this approach is that the results can be transposed to existing or perhaps also future process model notations and the modeler could not be hindered by an abundance of model language constructs. The BPMN *symbols* for the constructs were used in order to be easily understood by the participants, who were familiar with the BPMN notation. This latter process model notation was used in a number of lectures of the BPM course in which the participants were enrolled.

Supporting tool The Cheetah Experimental Platform³ (Pinggera et al. 2010a) was used to support the data collection. This program was developed at the University of Innsbruck as an open source research platform to support experiments investigating the process of process modeling. The modeling sessions were entirely supported by this tool and consisted of three consecutive tasks. The *tool tutorial task* presented short videos together with a brief explanation to exemplify each feature of the modeling editor. To reassure that the tool features were sufficiently understood, the user had to

mimic the actions of the video in the modeling editor correctly before the next feature was presented. Next, in the *process-modeling task* the participants had to construct a process model for the given case description. Finally, the *survey task* had to be completed by answering a questionnaire.

Data collection The experimental tool recorded each modeling operation automatically in an event log. A list of the different types of operations that were recorded, is presented in Appendix A. Besides the name of the recorded operation, the event records contained additional information such as the time of its occurrence, position on the canvas, source and target activities of edges, etc. These data can be used for a step-by-step replay of the model construction process or to feed mining algorithms that support analyses of this process (such as the PPMChart visualization described in Section 3.2.1 below). Furthermore, the tool captured the constructed process models, which allows for inspecting different properties of the produced models. Finally, the questionnaire (see Appendix B) was used to collect data about the demographics of the respondents, as well as domain knowledge, modeling language and method knowledge and general tool and language issues.

3.2 Data analysis

The answers to the demographic questions about the participants revealed that they are students between 20 and 28 years old, mainly male (93 out of 118). The majority of participants were non-native English speaking, but only two of them indicated to have some difficulties in reading or understanding English text. Table 1 presents an overview of the demographic data. Further, the students indicated to have had an average of 5.7 workdays of formal training on process modeling and 8.7 workdays of self-education. The mental effort was rated between 2 and 7 out of 10 (4.4 on average). Participants indicated they had no problem understanding the case description or working with the tool. More details about the prior knowledge of participants are provided in Appendix C.

3.2.1 PPMChart visualization

The PPMChart visualization represents the operations of the construction process of one modeler that produced a single process model (Claes et al. 2015). An example is shown in Fig. 1. The chart consists of horizontal timelines, one for each model element that was present during modeling. The top-down ordering of these timelines is derived from the sequence flows in the process model. Each colored dot in the graph represents

² A control flow model is a process model that mainly represents the sequential order of process steps (i.e., the control flow).

³ More information about the tool can be found at <http://www.cheetahplatform.org>.

Table 1 Demographic information of participants

Gender	Age	Native language	Current profession	Education program
93 Male	1 age 20	99 Dutch	116 Student	86 Operations Management & Logistics (OML)
25 Female	7 age 21	3 Chinese	1 Part-time student	25 Business Information Systems (BIS)
	36 age 22	2 English	1 PhD student	4 Innovation Management (IM)
	38 age 23	2 Greek		3 Human-Technology Interaction (HTI)
	22 age 24	2 Russian		1 Professional Doctorate
	11 age 25	1 Danish, French, German, Indonesian,		1 Doctorate (PhD)
	2 age 26	Macedonian, Persian, Polish,		
	1 age 28	Portuguese, Romanian		

one operation on one model element on the modeling canvas with the following characteristics:

- The *line* of the dot represents the model element on which the operation was performed (the identifier of the model element is displayed at the beginning of the line).
- The *position* of the dot on the line represents the time when the operation occurred (the default width of a PPMChart is 1 h).
- The *color* of the dot represents the type of operation (i.e., green for creation, blue for movement, red for deletion, orange for (re)naming, and grey for reconnection of edges).
- The *shape* of the dot represents the type of model element of the operation (i.e., circle for events, square for activity, diamond for gateways, triangle for edges).

For example, in the annotated highlight of Fig. 1 it can be observed that the first created element (i.e., the far most left green dot) was the start event (i.e., a circular dot on the first line). Next, an activity was put on the canvas somewhat later (i.e., a green square dot on another line, slightly more to the right). After the creation of some elements (i.e., left vertical zone of green dots), an almost simultaneous movement of all existing elements can be observed (i.e., vertical blue line of

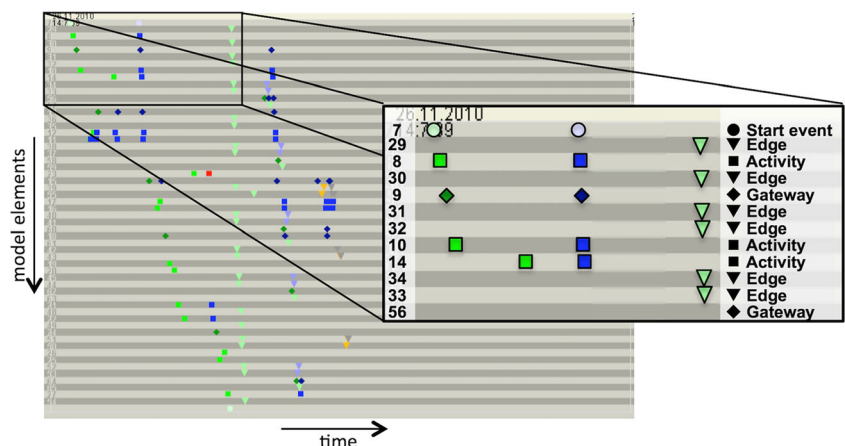
dots). Only much later, the edges that connect these elements were created (i.e., line of green triangular dots at the right).

3.2.2 Process model quality

Lindland et al. (1994) define three main quality dimensions of conceptual models: (i) *syntactic quality* indicates to which degree the symbols of the modeling language were used according to the rules of the language, (ii) *semantic quality* indicates how adequate the model represents the modeled phenomenon in terms of correctness and completeness, and (iii) *pragmatic quality* indicates the extent to which the users of the model understand the model as intended by the modeler.

In the study of the observational modeling sessions, only syntactic quality was evaluated to form impressions of the quality of the produced models. Rather arbitrarily, this dimension was selected because a measurement can be determined easily and objectively on the basis of the modeling language specification. It was assumed that the syntactic quality provides a sufficient insight at this stage of the research. Furthermore, a distinction was made between errors that originate in a lack of knowledge of the modeling language and errors that originate in cognitive failure. Especially the latter type of error is interesting for investigating our research question. In the remainder of the paper, the term ‘*mistake*’ is used to identify those syntactic errors in the process models, which

Fig. 1 PPMChart visualization representing one process-modeling instance



did not clearly arise from a lack of knowledge of the process modeling language. A list of the observed syntax errors is included in Appendix D, together with their classification in ‘mistakes’ and other syntactic errors.

3.2.3 Observations and impressions about the modeling process

PPMCharts allow for zooming in on specific operations (i.e., on individual dots in the charts), as well as on aggregated modeling phases and patterns (i.e., combinations of dots in the charts). Different PPMCharts were compared to extract patterns that reflect identifiable modeling approaches. This section presents a selection of such observations together with our impressions about the relation between these approaches and the properties of the resulting process models.

Serializing the modeling process When tasks are complex, people tend to deal with task complexity by splitting up the task in implicit subtasks that are executed sequentially (De Jong 2010). This complexity management technique is called serialization. Because this technique needs some cognitive administration, it results in frequent pauses during the modeling process in which no visible activities occur. From the 118 recorded modeling sessions, it was observed that in all but one of the sessions pauses were observed in the modeling replays as a timespan in which no operations occurred. These pauses were also evidenced in the PPMCharts as a vertical zone in which no dots occur.

Observation 1: All but one of the modelers paused frequently during the modeling process.

Of course, different events can have caused these pauses. Potentially the modeler was distracted, the modeling tool was lagging, the modeler was reading the case description, the modeler was thinking about the previous or the next steps, etc. Because often a high concentration of dots was observed right after a pause, the pauses seemed to us to be deliberate interruptions of the modeling pace in which the previous and/or future modeling operations were considered. It gave us the impression that the modelers needed to serialize the modeling process.

Impression 1: Modelers are in need of serializing the modeling process to deal with its complexity.

Structuring the modeling process Whereas serialization is defined as splitting up a task in sequentially executed subtasks, structuring can be defined as the extent to which a consistent strategy is applied for defining those

subtasks. The way of (not) structuring the modeling process can be recognized in the PPMChart by the patterns that can (not) be clearly discovered in the arrangement of the dots in the chart.

The analysis of the PPMCharts revealed that 33 of the 118 modelers (28 %) have built the process model in a flow-oriented way, meaning that they constructed parts of the process model according to the control flow structure of the process. Once a part of the model was considered complete, these modelers did not change that part of the model anymore. In the PPMChart, because of the sorting of timelines according to the sequence flows in the model, this pattern was observed as a diagonal zone of operations (see Fig. 2).

Observation 2: A large group of the sessions can be categorized as “flow-oriented process modeling”.

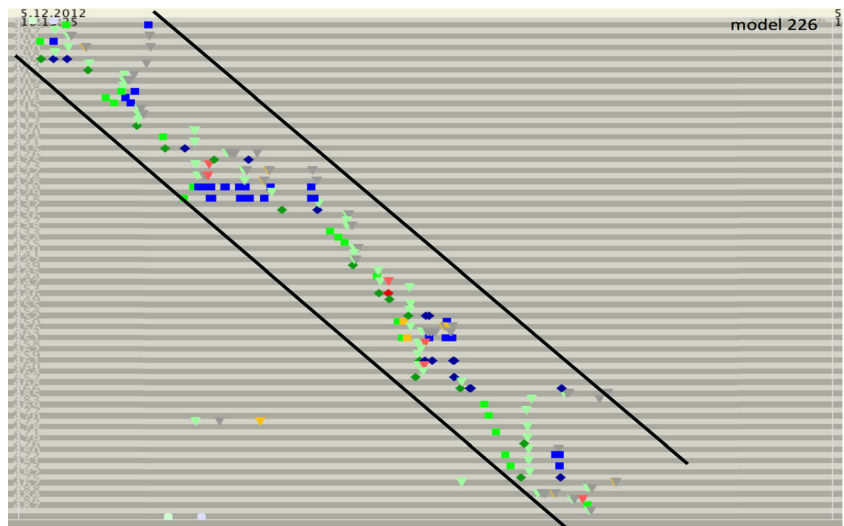
Conversely, ten other modelers (8 %) organized the modeling process in an aspect-oriented way. Aspect-oriented process modeling is observed when the modeler consecutively directs attention to different aspects of modeling.⁴ They may for example first focus on the content of the model (i.e., placing every activity and gateway on the canvas), then on the sequence flow of the activities (i.e., connecting the elements with sequence flow arrows), and finally on the layout of the model (i.e., moving and aligning elements). In Fig. 3 aspect-oriented process modeling can be observed as several non-overlapping zones each enclosing similar operations.

Observation 3: A smaller group of the sessions can be categorized as “aspect-oriented process modeling”.

Whereas many situations were discovered in which a flow-oriented or aspect-oriented organization of the modeling process was used consistently, it should be noted that also combinations were observed in 33 of the 118 cases (28 %). Figure 4 shows an example where process model elements are *created* in a flow-oriented manner, but overall the modeler alternated between dedicated phases of working on different modeling aspects. This can be observed by a diagonal zone of

⁴ The term aspect-oriented process modeling must not be confused with aspect-oriented modeling. The former is our description of splitting up the modeling *process* according to the various aspects that are targeted sequentially. The latter is a way of splitting up the model itself in submodels that each represent another aspect of the system to be modeled. Both terms are derived from aspect-oriented programming, a technique for splitting up the programming process as well as the program code according to the different aspects to be programmed.

Fig. 2 Example of flow-oriented process modeling



green dots followed by a number of zones of limited height that each consist of similar dots.

Observation 4: Another large group of the sessions used a combination of “flow-oriented process modeling” and “aspect-oriented process modeling”.

It should also be noted that not every modeler seemed to implement a particular way of organizing the modeling process, as could be concluded from Fig. 5. No clear pattern of dots was discovered in the charts. A subset of 12 of the 118 instances (10 %) was labeled “undirected process modeling”. The term “undirected” is preferred over “unstructured”, because it is physically impossible for most people to perform actions without any form of structured approach.

Observation 5: Another small group of the sessions can be categorized as “undirected process modeling”.

So far, three structuring strategies for serialization were observed: flow-oriented process modeling, aspect-oriented process modeling, and a combination of both approaches. We also observed undirected process modeling.

The remaining 30 sessions (25 %) could not clearly be categorized as structured or undirected. They were labeled “uncategorized” and were left out of scope for further analysis.

In order for a process model to be syntactically correct, no syntax errors may exist in the model. The number of ‘mistakes’ in each model was assessed by the authors. In their assessment, for certain syntactical errors they had to make a subjective decision whether the error should be classified as a

Fig. 3 Example of aspect-oriented process modeling

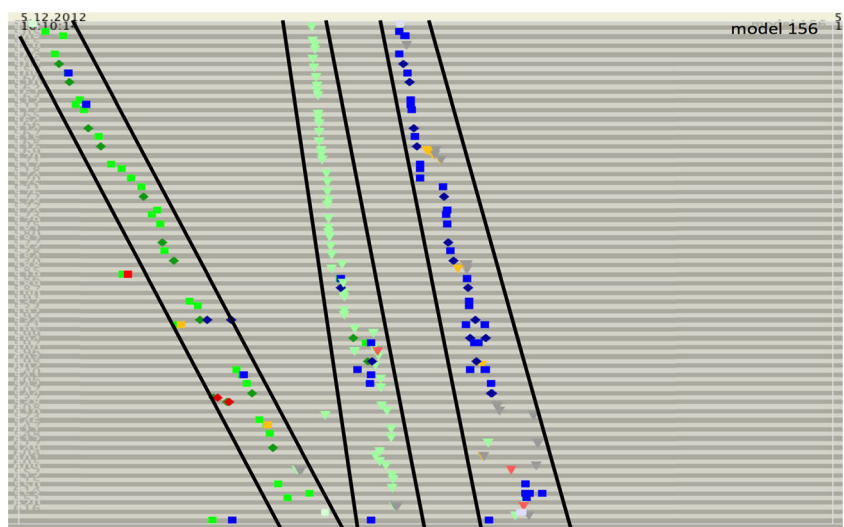
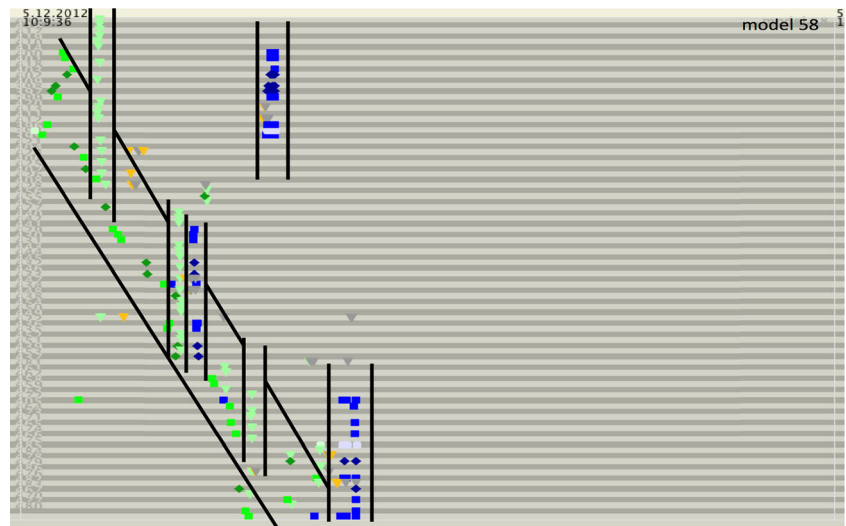


Fig. 4 Example of a combination of flow-oriented and aspect-oriented process modeling



‘mistake’ or not (see classification of errors in Appendix D). From the models constructed according to a structured serialization strategy (i.e., flow-oriented, aspect-oriented, and combination of both approaches), a bigger proportion seemed to contain no ‘mistakes’ than models originated from the undirected modeling approach. The impression arose that serializing the modeling process in a structured way helps avoiding these errors caused by cognitive failure.

Impression 2: Structured serializing of the modeling process helps avoiding ‘mistakes’.

However, not every model that was created in a structured way ended up containing no ‘mistakes’. This could for example be explained if factors exist that counter the effect of the structured approach. Nevertheless, no factors hindering the modeler were observed and for that reason we got the

impression that the structuring did not help every modeler in the same way.

Impression 3: Structured serializing does not support every modeler to avoid ‘mistakes’ to the same extent.

Speed of the modeling process Finally, it was observed how in some PPMcharts the zone that contains dots is narrower than in other charts (e.g., compare Fig. 4 with Fig. 5). This means that some modelers took less time to construct the process model. A comparison of time distribution of the four defined serialization strategies revealed that the modeling sessions of the category “undirected” lasted clearly longer than the three other categories (see Fig. 6). Independent t-tests indicated that the mean modeling time of the three structured approaches was significantly different from that of the

Fig. 5 Example of undirected process modeling

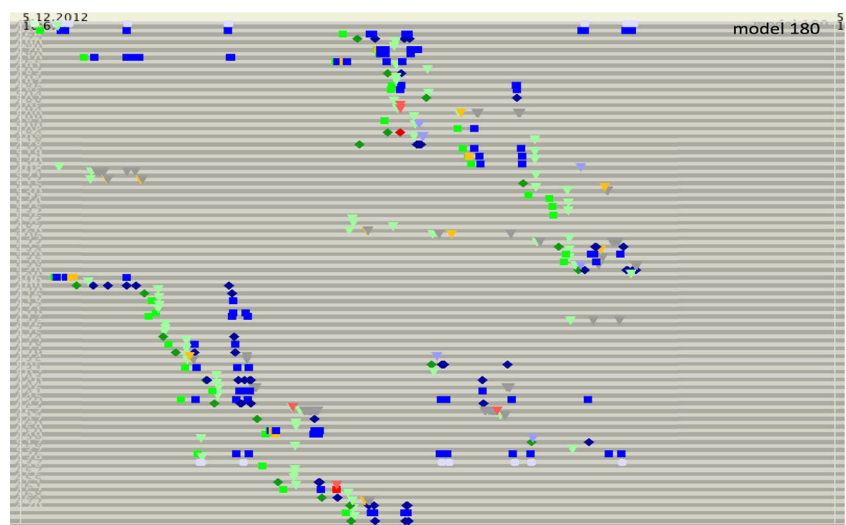
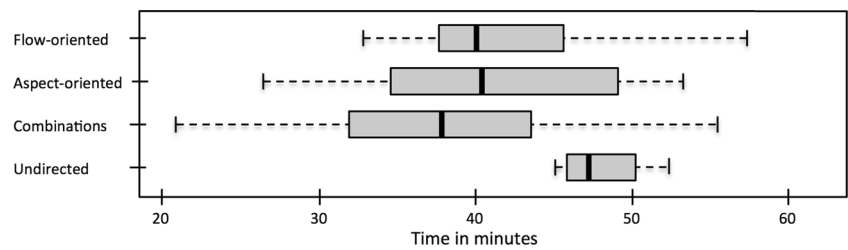


Fig. 6 Boxplots of modeling time in minutes for each of the observed serialization styles



undirected approaches ($p_{FO-UD} = 0023$, $p_{AO-UD} = 0012$, $p_{C-UD} = 0000$). The structured approach seems not only to help reducing the number of ‘mistakes’; it may also speed up the modeling process.

Observation 6: The sessions labeled “undirected process modeling” lasted longer than the other approaches.

Overview An overview of the data regarding Observations 1 to 6 is presented in Table 2. The outcomes of the exploratory study are the six observations and the three impressions, summarized in Table 3. As proposed by (Godfrey-Smith 2009; Nagel 1979; Popper 2005) both observations and related impressions can then be used as input to build a theory, which is described extensively in Sections 4 and 5.

4 Theoretical background

Different cognitive theories can be combined to provide explanations for the observed modeling approaches and their relation with modeling accuracy and speed. In the next section, the theoretical background presented here is used to formulate a theory for explaining modelers’ cognitive strategies for dealing with complexity.

4.1 Kinds of human memory

The literature on cognition describes three main kinds of human memory. *Sensory memory* is very fast memory where the

stimuli of our senses are stored for a short period (Sperling 1963). During this instant the information that is unconsciously considered relevant is handed over to working memory (Sperling 1963). Next, the information in working memory is complemented with existing knowledge that is retrieved from *long-term memory* (Sweller et al. 1998). This latter kind of memory is slow but virtually unlimited (Sweller et al. 1998). Information is stored in long-term memory as cognitive schemas composed of patterns of connected elementary facts (Sweller et al. 1998). Relevant information for process modeling that is retrieved from long-term memory includes domain knowledge, and modeling language and modeling method knowledge. In *working memory* the information is organized and processed in order to initiate certain performances (e.g., to put an activity on the modeling canvas with a mouse click) or to complement the knowledge in long-term memory (e.g., to complement the mental model of the case with new insights from a line of text that was read) (Atkinson and Shiffrin 1968). Because working memory has a limited capacity (Cowan 2010; Miller 1956) and information can only be stored in this memory for a short period (Van Merriënboer and Sweller 2005), it is important to use it effectively when dealing with highly complex tasks, such as process modeling.

4.2 Types of cognitive load

Process modeling requires input information to be absorbed and complemented with knowledge from long-term memory such as domain knowledge, in order to be processed in working memory leading to the actions of constructing the process model. The center of this complex task are the operations in working memory (Atkinson and Shiffrin 1968; Sweller et al.

Table 2 Observed serialization strategies and their measured properties

Serialization strategy	Number of cases	Number of serialized approaches	Mean modeling time
Flow-oriented process modeling	33/118 (28 %)	33/33 (100 %)	42,80 ± 9,56 min.
Aspect-oriented process modeling	10/118 (8 %)	10/10 (100 %)	40,32 ± 9,44 min.
Combined process modeling	33/118 (28 %)	33/33 (100 %)	37,37 ± 7,92 min.
Undirected process modeling	12/118 (10 %)	11/12 (99 %)	49,87 ± 6,63 min.
Uncategorized cases	30/118 (25 %)	30/30 (100 %)	–

Table 3 Observations and impressions about how people deal with complexity during process modeling

Observations	
Observation 1	All but one of the modelers paused frequently during the modeling process.
Observation 2	A large group of the modeling sessions can be categorized as “flow-oriented process modeling”.
Observation 3	A smaller group of the sessions can be categorized as “aspect-oriented process modeling”.
Observation 4	Another large group of the sessions used a combination of “flow-oriented process modeling” and “aspect-oriented process modeling”.
Observation 5	Another small group of the sessions can be categorized as “undirected process modeling”.
Observation 6	The sessions labeled “undirected process modeling” lasted longer than the other approaches.
Impressions	
Impression 1	Modelers are in need of serializing the modeling process to deal with its complexity.
Impression 2	Structured serializing of the modeling process helps avoiding ‘mistakes’.
Impression 3	Structured serializing does not support every modeler to avoid ‘mistakes’ to the same extent.

1998). The necessary information fills up working memory and is subdivided in three types of cognitive load (Sweller and Chandler 1994). *Intrinsic cognitive load* is the amount of information that needs to be loaded in working memory for deciding how to conduct a particular task. It mainly depends on the properties of the task and the amount of relevant prior knowledge of the performer of the task (i.e., knowledge about the domain, about the modeling language and about the modeling method). *Extraneous cognitive load* is the load that is raised for processing and interpreting the input material of the task such as descriptions or direct observations of the process to be modeled. This type of cognitive load depends on the representation of the input material as well as the fit of this representation with the task it has to support and with the characteristics of the interpreter of the material (Vessey and Galletta 1991) (see also Section 4.4). Finally, during the execution of a task humans usually are able to reserve some load in working memory for building, restructuring and completing cognitive schemas to be stored in long-term memory. This will help reducing cognitive load for performing similar tasks in the future. This activity is called learning and the associated load is the *germane cognitive load*. Furthermore, a distinction can be made between the *overall* cognitive load (i.e., the total amount of information sequentially loaded in working memory for performing a specific task) and *instantaneous* cognitive load (i.e., the amount of information that is loaded in working memory at a certain point in time) (Paas et al. 2003b).

4.3 Cognitive load theory

The capacity of *working memory is limited*. In the past, researchers have tried to define how much information can be loaded at the same time in this kind of memory. Miller estimated the amount of information that can be remembered in short term memory at about 7 units (Miller 1956). More recent research concludes that only 3 to 4 units of information can be activated and processed in working memory at the same time (Sweller et al. 1998; Van Merriënboer and Sweller 2005). Although there appears to be a limit on the amount of units that can be loaded simultaneously in working memory, there seems to be no constraint on the size and complexity of these units of information (Sweller et al. 1998). More specifically, it is believed that one unit of information loaded in working memory (often referred to as ‘information chunk’) corresponds with one cognitive schema in long-term memory (Sweller et al. 1998). This can explain why a person seems to be able to store more information in working memory for tasks in which he is experienced, because for such tasks he was able to build up larger and stronger cognitive schemas in the past.

Therefore, for complex tasks or tasks in which a person is not adequately experienced, it is imaginable that the limited capacity of the working memory is not sufficient for the (maximum instantaneous) load that is needed to accomplish the task. This is called *cognitive overload* (Sweller 1988). The Cognitive Load Theory states that when working memory is overloaded, there is *no room for learning* (i.e., schema building) and *accuracy and speed* of information processing *decrease* (Rockwell and Bajaj 2005; Sweller 1988). In other words, cognitive overload has a negative impact on the effectiveness and efficiency of the modeling performance.

4.4 Cognitive fit theory

The *Cognitive Fit Theory* states that humans are able to solve problems more effectively and efficiently if the *representation of the input material of a certain task ‘fits’ with the task itself* (Vessey 1991). For example, when a task involves exploring relationships between data a visual representation such as a diagram is preferred. For more statistical purposes such as determining the average of a series of numbers a textual representation in the form of a list or table is more cognitive efficient (Vessey 1991). Whereas the focus of Cognitive Fit Theory is on the match between problem representation and task, a secondary effect is described as the *match between the task and its performer*. For example, most people excel in either graphical or logical tasks (Pithers 2002). The former type of people probably needs less effort to work on the layout of the model, whereas the latter may find it easy to warrant the semantic correctness of the model. For the development of our theory, we focused mainly on this secondary relation between task and performer. Since the initial publication of the theory

in 1991, the work is refined and concepts of domain knowledge, method knowledge and problem solving tools are taken into account as well (Khatri et al. 2006a, b; Shaft and Vessey 2006; Sinha and Vessey 1992; Vessey 1991).

4.5 Overview

Figure 7 provides an overview of the reviewed cognitive theories and integrates them into a conceptual framework displaying the causal relations that might explain the phenomenon of cognitive overload in working memory during task performance. The task under consideration is the construction of a process model. The central construct of the constructed theoretical framework is cognitive overload, which depends on the modeler’s working memory capacity and the cognitive load that the task requires. This cognitive load is composed of extraneous, intrinsic and germane cognitive load.

Extraneous cognitive load mainly depends on the input material representation fit with the task and the modeler. A higher fit requires a lower cognitive load. For process modeling the input material includes any descriptive process information such as verbal or oral transcripts of interviews with process managers/workers and existing documents describing the process.

The intrinsic cognitive load increases for more complex tasks and decreases in case the modeler possesses more relevant prior knowledge. Differences between various process modeling tasks are mainly related to the complexity of the case to be modeled. Prior knowledge incorporates domain knowledge, and modeling language and method knowledge.

Germane cognitive load is caused by loading information in working memory for the construction of cognitive schemas, which is not a prerequisite for the task, but rather the result of learning. This can only occur if during previous processing of information the working memory was not overloaded.

If the sum of these three types of cognitive load at a certain point in time transcends working memory capacity, cognitive overload occurs. This has a negative effect on process model quality (i.e., more ‘mistakes’ are made), speed of modeling, and learning. Note that learning means that the set of cognitive schemas of the modeler is broadened and strengthened, which gradually improves the useful knowledge of the modeler for future similar tasks.

5 The structured process modeling theory (SPMT)

In order to explain the observations and impressions presented in Section 3, the cognitive theories listed in Section 4 were integrated and transformed into the newly developed Structured Process Modeling Theory (SPMT). Three key concepts were extracted from the observations and impressions: serialization, structuring and individual differences. Therefore the SPMT consist of three parts, each targeting one of these concepts.

5.1 Part 1: Serialization of the process modeling task can reduce cognitive overload

We observed that inexperienced modelers use a serialization approach to construct the process model (i.e., Observation 1). Our impression was that the serialization appears to help these modelers dealing with the complexity of the modeling task (i.e., Impression 1). Cognitive theories also recognize the concept of cognitive serialization to deal with cognitive overload. If a task requires too much information to be stored in working memory simultaneously, then it is advised to load the information sequentially (Bannert 2002; De Jong 2010; Gerjets et al. 2004; Paas et al. 2003a; Pithers 2002; Pollock et al. 2002; Van Merriënboer et al. 2003). This means that intrinsic cognitive load can be spread out over a longer period, which

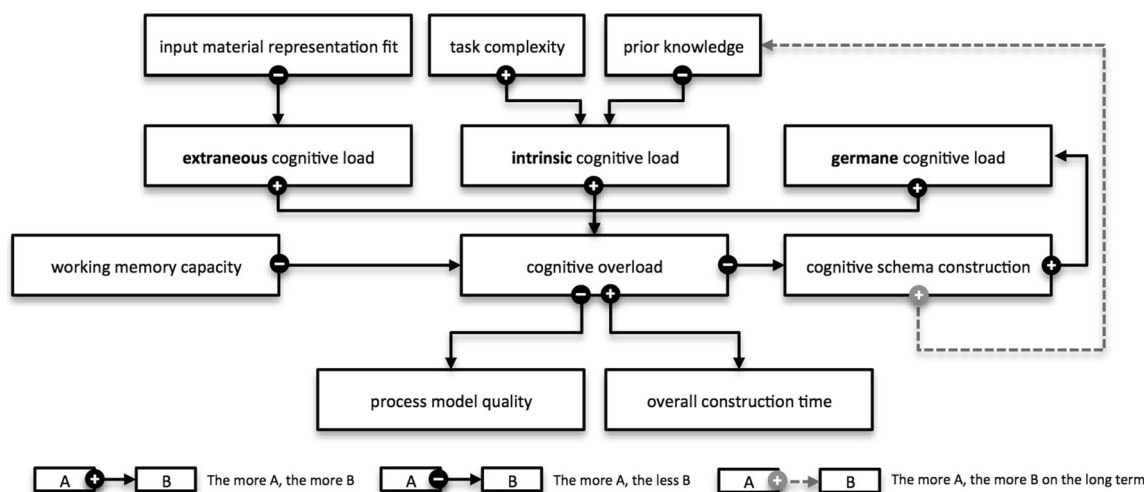


Fig. 7 Causal model centered on cognitive overload in working memory

reduces the probability of instantaneous cognitive overload (De Jong 2010).

On the other hand, serialization causes more intrinsic cognitive load for integration and for administration of the sequentially processed and produced information (Gerjets et al. 2004). In other words, extra load is created to aggregate the information of the separate parts of a solution and for building the modeling strategy (Gerjets et al. 2004; Van Merriënboer et al. 2003). The modeling strategy determines how to divide the modeling task in subtasks, in which order to proceed, how to execute each subtask, how to aggregate the different partial results, etc. The extra load for aggregation and strategy building results in a total *overall* intrinsic cognitive load that can be higher in case of serialization. But if the intrinsic load for aggregation and strategy building can be kept low, the maximum *instantaneous* load decreases together with the probability of cognitive overload.

Figure 8 shows these relations graphically. The adopted serialization style symbolizes how the model construction was serialized (e.g., flow-oriented, undirected, etc.). The *degree* of serialization indicates how much the modeling was subdivided. The *structuredness* of the serialization indicates the consistency of the implemented serialization strategy. According to the aforementioned phases of model parts creation, information aggregation and modeling strategy building, the intrinsic cognitive load is artificially subdivided into three subtypes of intrinsic load, but it is practically impossible to distinguish between those three types (Gerjets et al. 2004). Whereas the degree of serialization mainly impacts the intrinsic cognitive load for process modeling and aggregation, the structuredness of the

serialization determines further how much load the serialization poses on aggregation and strategy building. Because the effect of structuring is explained in Part 2 of the SPMT, the effect of serialization on strategy building is not included in Part 1 of the SPMT which centers only on the degree of serialization.

We conclude that serialization of the process of process modeling helps reducing the probability of instantaneous cognitive overload if the important condition is met that aggregation of the partial solutions (and modeling strategy building) do not consume the freed resources in working memory.

5.2 Part 2: Structured process modeling reduces cognitive overload

As discussed in the previous subsection, the benefits of serializing complex cognitive tasks can only be realized if the accompanying additional cognitive effort for strategy building and aggregation does not surpass the gain of serializing. Observations 2–5 state that different serialization approaches exist. Observation 6 reports that the observed structured approaches (i.e., flow-oriented, aspect-oriented or a combination of these) were faster than the undirected approach. In Impression 2 our perception is expressed that the structured approaches also help to reduce the occurrence of ‘mistakes’. Structuring the process modeling approach seems to increase the effectiveness and efficiency of the construction of process models, which can be explained if these techniques ensure that the cognitive load for strategy building and aggregation

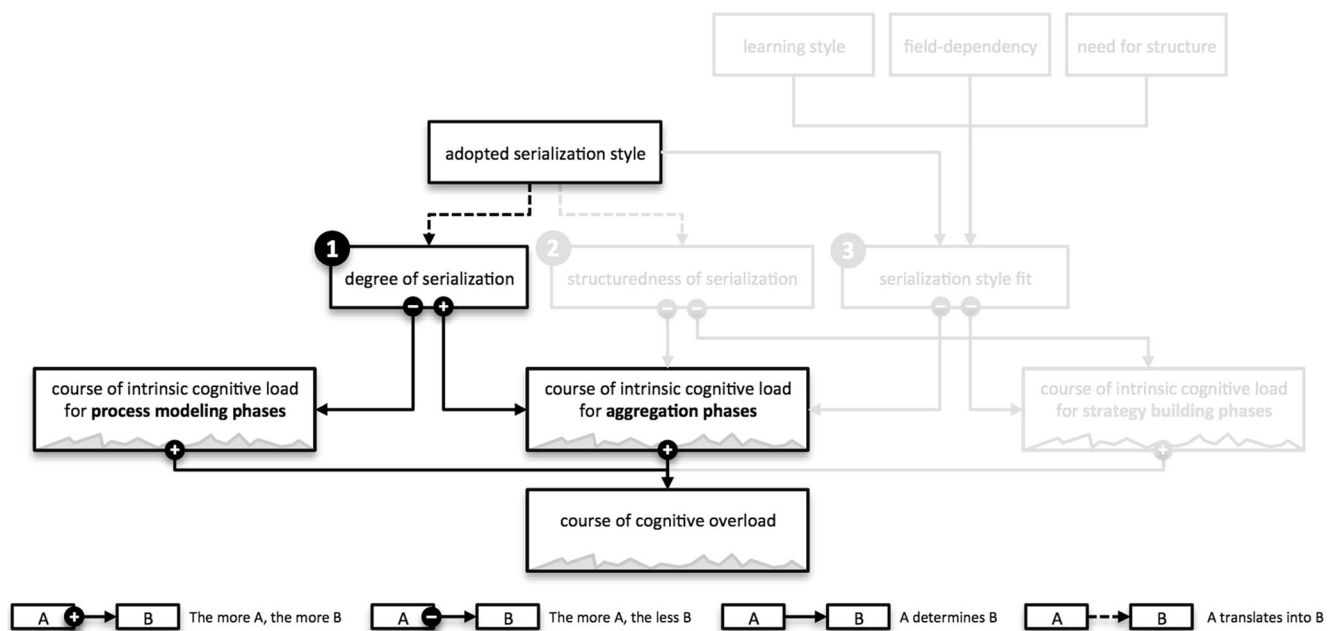


Fig. 8 The effect of serialization on the course of cognitive overload

is limited. Hence, Part 2 of the SPMT provides the theoretical support for this conclusion (see Fig. 9).

A more structured serialization approach towards process modeling makes it easier to keep track of the progress of the modeling endeavor (Van Merriënboer et al. 2003). This in turn lowers the effort to evaluate and adjust the modeling strategy (Van Merriënboer 1997). By structuring the serialization process, also the outcome of this process (i.e., the process model) will probably be more structured, which facilitates the aggregation of the separately developed parts of the process model (Kim et al. 2000).

Therefore, part 2 of the SPMT states that structuring the (serialized) approach towards process modeling lowers the intrinsic cognitive load for aggregation of the partial solutions and modeling strategy building and thus reduces the probability of instantaneous cognitive overload.

5.3 Part 3: Serialization style fit is a prerequisite for cognitive overload reduction

Nevertheless, based on Impression 3, it is proposed that a third factor has to be considered. Besides the degree and structuredness of the serialization, also the fit of the adopted serialization style with the characteristics of the problem solver plays an important role in the cognitive load that the problem imposes on the modeler (Vessey 1991). This is represented in Fig. 10.

Cognitive literature suggests that each human being has a specific intrinsic learning style (Felder and Silverman 1988). One of the defined dimensions of learning style is called ‘global/sequential understanding’ (Felder and

Silverman 1988). It specifies to what extent a learner needs the material to be processed sequentially. For example, we hypothesize that the flow-oriented modeling method is better suited to a sequential learner, because it builds up the model in a sequential manner. The aspect-oriented approach starts from a global view of the content of the model and drills down into the different details (aspects) of the process model to be constructed. Therefore, it is matched with the global learning style.

Similarly, the Field Dependence-Field Independence Theory (Pithers 2002; Witkin and Goodenough 1981) states that some people are better in abstract reasoning than others (i.e., they do not need to load a lot of contextual information in memory). Field dependent modelers find it harder to break up the model in smaller parts that they will construct separately and without considering its context (Pithers 2002), which means they may prefer the aspect-oriented style for structuring the modeling process, because for each aspect that is targeted sequentially the whole process model is considered before turning to the next aspect.

The Need for Structure scale defines to what extent the performance of a person depends on the structuredness of the adopted solution method (Neuberg and Newsom 1993). Therefore it is hypothesized that modelers with a high need for structure will benefit most of structuring the modeling process according to the one or the other structuring style.

In summary, based on cognitive theories, we suggest that the load for aggregation of the partial solutions and for modeling strategy building can be kept low if the serialization of the process-modeling task is conducted in a structured way that fits with the characteristics of the modeler.

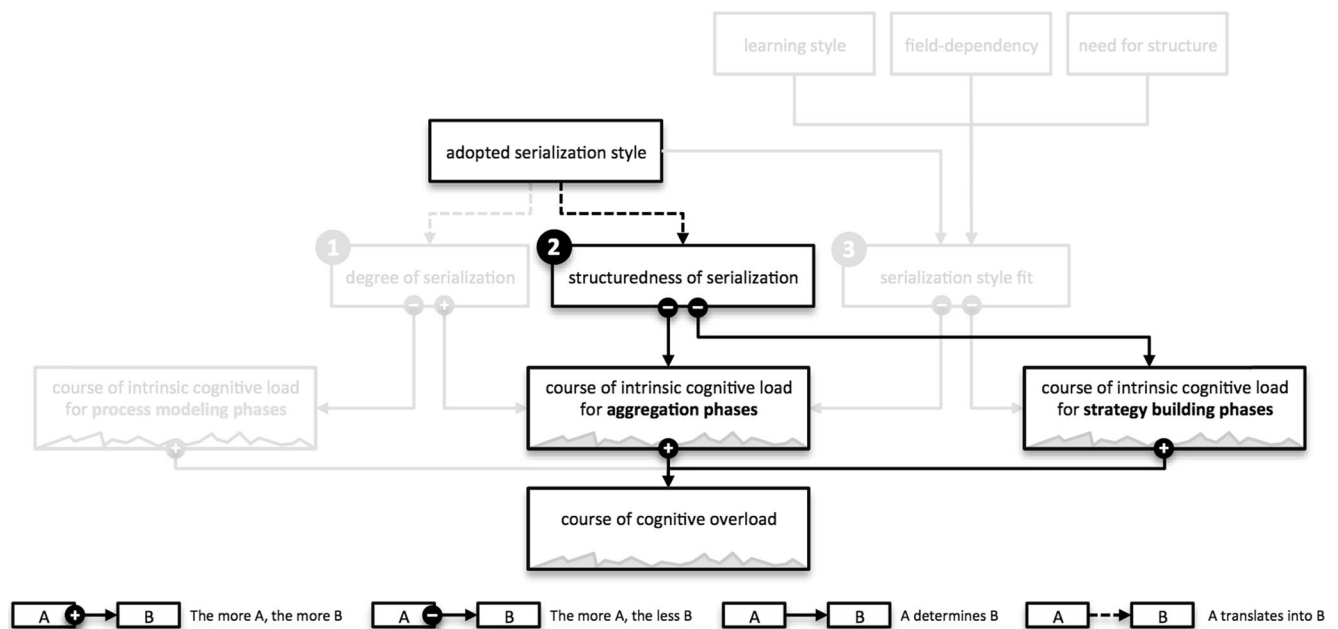


Fig. 9 The effect of structuredness of the serialization on the course of cognitive overload

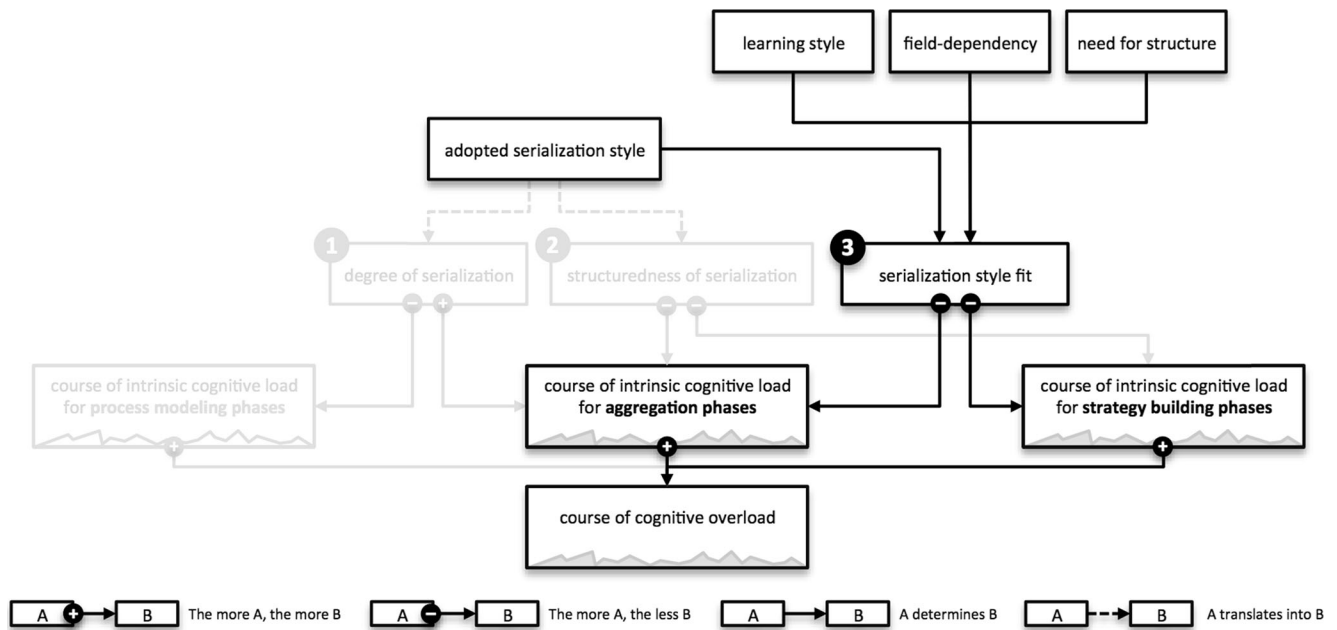


Fig. 10 The effect of serialization style fit on the course of cognitive overload

5.4 Summary of the SPMT

The SPMT is summarized below in (i) a theoretical model, which graphically represents the included constructs and their relation; (ii) the propositions, which describe the produced knowledge in a textual format; and (iii) a brief description of the boundaries of the theory.

5.4.1 Theoretical model

As stated before, the intrinsic cognitive load together with the extraneous and germane cognitive load that is needed to solve a certain problem can exceed working memory capacity in which case cognitive overload occurs. When this happens, a negative effect on modeling accuracy and modeling speed results in a decrease of effectiveness and efficiency of the overall modeling endeavor. The SPMT explains how the technique of *individually fitting structured serialized process modeling* can lower the course of intrinsic cognitive load (and thus also the chance of cognitive overload) for a given case complexity and prior knowledge (see Fig. 11).

5.4.2 Propositions

Based on the three parts of the research model, the research model of the SPMT can be complemented with three propositions.

Proposition 1: Serialization.

When the construction process of process models is serialized, the instantaneous intrinsic cognitive load for modeling

can be kept lower. If this reduction is greater than the accompanying increase of instantaneous intrinsic cognitive load for aggregating and for strategy building, the total cognitive load is decreased.

Proposition 2: Structured serialization.

When the serialized process modeling approach occurs in a structured fashion, the increase in instantaneous intrinsic cognitive load for aggregating and for strategy building can be reduced.

Proposition 3: Individually fitting structured serialization.

If the structured serialization approach (e.g., aspect-oriented or flow-oriented process modeling) fits with the characteristics of the modeler (i.e., learning style, need for structure and field-dependency) the increase in instantaneous intrinsic cognitive load for aggregating and for strategy building can be further reduced.

5.4.3 Boundaries

The SPMT was based on observations and impressions in a specific setting. The observed subjects were master students. They served as a proxy for inexperienced modelers. The observed task was the construction of a control flow model in a simplified modeling language. Therefore, the SPMT applies at least for control flow modeling by inexperienced modelers. However, the SPMT is composed of constructs and relations that were found in literature. The only boundary of these existing theories is that they describe cognitive properties,

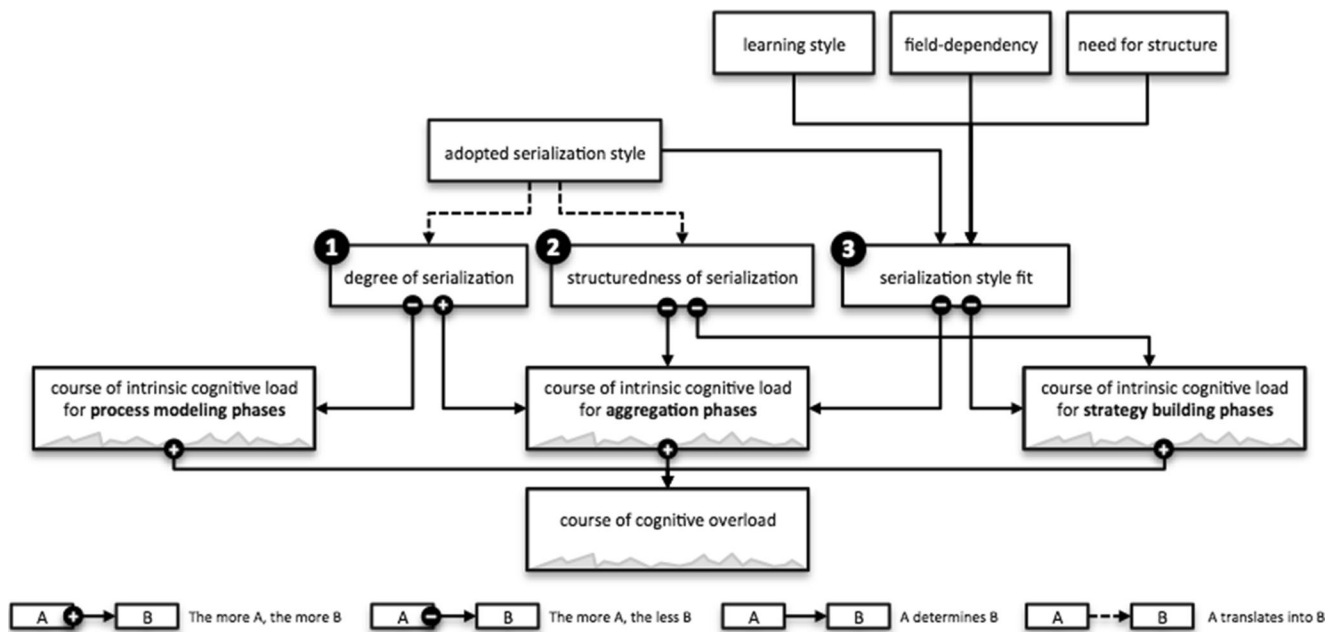


Fig. 11 Theoretical model of the Structured Process Modeling Theory (SPMT)

processes or relations of human beings. The SPMT may thus apply to more generic situations.

6 Evaluation of the structured process modeling theory (SPMT)

In this section the six criteria for evaluating an explanatory theory mentioned in Section 2.2 are applied to the SPMT: *novelty*, *parsimony*, *consistency*, *plausibility*, *credibility*, and *transferability*. These criteria were found in various academic articles about theory testing (Gregor 2006; Grover et al. 2008; Weber 2012; Weick 1989). Nevertheless, we found no concrete guidelines on how to assess the SPMT against these criteria. In this paper, where the emphasis is on the theory building, logical arguments rather than empirical data are used to evaluate the criteria (Whetten 1989). The section concludes with a brief discussion of two other important theory testing criteria that we consider currently not feasible to evaluate: *falsifiability* and *utility* (Bacharach 1989).

6.1 Novelty

There are different ways in which a theory can be novel: (i) it describes constructs or associations that were not established before, (ii) it describes well-known constructs or associations in a fundamental new way, (iii) it makes important changes to existing theory (Weber 2012). The SPMT is novel because it combines several existing cognitive theories in a fundamental new way. The consideration of the first part of the SPMT - describing how serialization of the modeling effort helps reducing intrinsic cognitive load - has been touched before

(Rockwell and Bajaj 2005; Soffer et al. 2012). Yet, the idea of structuring the construction process of the process model (i.e., the second part of the SPMT) seems more original, although structuredness of the outcome of such a construction process is well studied (Laue and Mendling 2010; Zugal et al. 2013). Also in software engineering for example, there are many studies about the structuredness of program code (e.g., procedural versus object-oriented code (Wiedenbeck and Ramalingam 1999)).

The real novelty of the SPMT lies in the third part. The technique of serialization is described in cognitive literature as “cognitive sequencing” (De Jong 2010). Different structured sequencing strategies are defined: e.g., simple-to-complex sequencing, part-whole sequencing (similar to flow-oriented modeling), simplified whole tasks or whole-task sequencing (similar to aspect-oriented modeling), and modular presentation (Gerjets et al. 2004; Van Merriënboer et al. 2003). However, while the notions of cognitive fit where already published in 1986 (Vessey and Weber 1986), the principle of cognitive fit is not considered in literature when advising which of these sequencing strategies to use. For example whole-task sequencing is considered to always outperform part-whole sequencing (Van Merriënboer et al. 2003) and modular presentation in turn was presented as an improvement of whole-task sequencing (Gerjets et al. 2004). Nevertheless, we propose that cognitive fit should be considered for selecting the appropriate sequencing technique, as is stated in part 3 of the SPMT.

6.2 Parsimony

A theory is considered parsimonious if it uses only a small number of constructs and associations to accurately describe

their focal phenomena (Weber 2012). Still, a high amount of relevant constructs and associations are presented throughout this paper. Most of them however are used to describe existing knowledge that constitutes the context of the SPMT. When the three parts of the SPMT themselves are considered, only a small number of constructs and associations is used. The number of constructs and associations of the separate parts and the whole of the SPMT are summarized in Table 4.

The artificial distinction between the three types of intrinsic cognitive load (i.e., load for process modeling, for aggregating and for strategy building) and between the two attributes of selected serialization style (i.e., degree and structuredness) could have been omitted. This would reduce the total amount of constructs to 7 and the amount of associations to 8 (i.e., still distinguishing between a positive and a negative effect of adopted serialization style on intrinsic cognitive load). Nevertheless, this would - in our opinion - also significantly diminish the explanatory power and the understandability of the theory.

6.3 Consistency

A theory is consistent if various observations can be explained with the same theory. Therefore, other available datasets with recorded data about the process of process modeling were examined for supplementary observations about complexity handling during the modeling activities. Another set of observational modeling sessions in 2013 contained such additional observations. Participants were master students of Business Engineering at Ghent University. They have a similar background and are enrolled in a similar educational program as the students from the exploratory modeling sessions in Eindhoven. 143 additional modeling sessions were recorded. The case to be modeled described a process about collecting fines.⁵

A new way of structuring the modeling process was observed in this additional dataset. Twelve of the 143 modelers (8 %) used a way of structuring that was labeled “happy path first modeling” (see Fig. 12). The modelers seemed to have first modeled the main process behavior (i.e., the happy path), and afterwards they modeled an exceptional route. They all ended up with a process model without ‘mistakes’ that took far less time to construct than the 11 undirected ones in the dataset.

The SPMT was developed to describe and explain the observations of Section 3.2.3, but it was also intended to be applicable in a broader sense. As a consequence, it can also be used to explain this additional observation about a previously not discovered way of structuring the modeling process. The modelers clearly serialized the modeling process and used

Table 4 Number of constructs and associations in the SPMT

	Part 1	Part 2	Part 3	SPMT
Constructs	5	5	8	11
Associations	5	5	8	15

a structured approach (i.e., the happy path first modeling). Potentially, the use of this particular structuring style fitted more to the modeler. For example, a sequential learner that can be classified as field independent, would prefer the flow-oriented approach towards modeling the happy path, but may like to abstract from exceptional behavior at first. According to the SPMT, this can explain why they appeared to have made fewer ‘mistakes’ and were faster than the modelers from the “undirected process modeling” subset. Structuring their approach to modeling in a way that fitted with their characteristics has helped them avoid cognitive overload, which has increased their modeling accuracy and speed.

In other words, the SPMT can be used in a consistent way to explain this observation. There was no need to adapt or complement the SPMT in order to be used to explain why the happy path first modeling approach helped these particular modelers. Moreover, a retrospective examination of the modeling sessions in the dataset described in Section 3, showed that 15 of the 118 sessions (13 %) could have been labeled “happy path first modeling”. It was also noticed that all but one of these instances were currently labeled ‘uncategorized’.

6.4 Plausibility & credibility

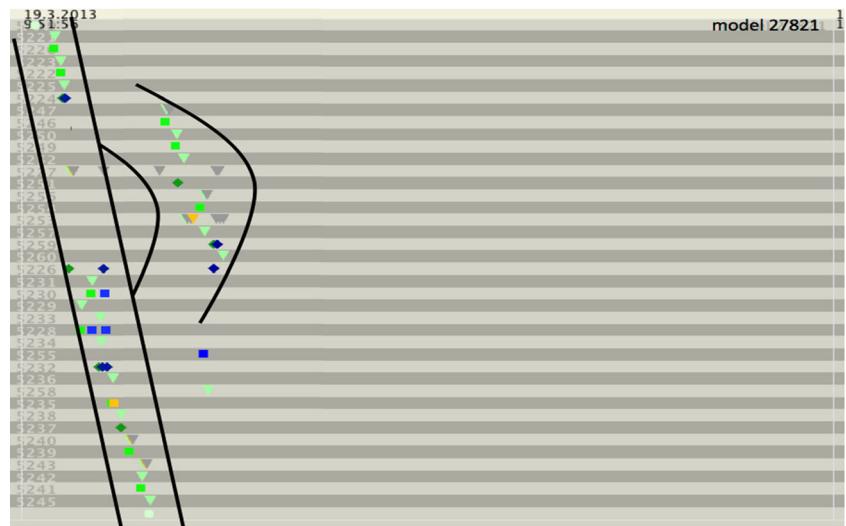
The real observed behavior, pronounced in the observations and impressions in Section 3.2.3, was explained based on established theory. Existing cognitive theories were used to provide all the constructs and associations that make up the theory. This theory building methodology warrants both plausibility and credibility. The SPMT is plausible, because it explains accurately and profoundly the effects that were observed in reality. It is also credible, because it uses only constructs and associations from established existing theories to explain those effects.

6.5 Transferability

A good theory is transferable to other research contexts. The SPMT was developed as a mid-range theory (Weber 2012) with the observations and impressions in the context of process modeling in mind. Nevertheless, the constructs and associations that constitute the theory were taken from general cognitive literature. Therefore, the SPMT has the potential to be transferred beyond the process-modeling domain. It can apply also in other domains such as conceptual modeling in

⁵ The case description can be downloaded at <http://www.janclaus.info/papers/PPMISF>.

Fig. 12 Example of happy path first process modeling



general, programming, text writing, etc., which would make it a macro-level theory (Weber 2012). In order to establish the real rather than the potential theory level and transferability, the theory needs to be applied and tested in various domains, which is addressed as future work in Section 8.3.

6.6 Falsifiability and utility

According to (Bacharach 1989) a theory should be evaluated against two other primary criteria: falsifiability and utility. We acknowledge this point of view, but evaluating our theory against these criteria is considered infeasible at this point and therefore out of the scope of this paper. The evaluation of falsifiability of the SPMT is explicitly addressed as future research in Section 8.3, because this requires the propositions of the theory to be operationalized into testable hypotheses. The best way of evaluating the utility of a theory is to measure how much it is actually used for practical and academic purposes, which is off course only possible on a longer term.

7 Related work

Although the constructs of serialization, structuredness and cognitive fit, the three parts of the Structured Process Modeling Theory (SPMT), were not considered together before, they were studied separately in various contexts. In this section, related work is presented that takes a cognitive view on general conceptual modeling or process modeling in particular with a focus on serialization, structuredness or cognitive fit.

Serialization Rockwell and Bajaj (2005) propose the COGEVAL framework that consists of a collection of 8

propositions about modeling complexity and model readability based on cognitive theories. One of the propositions presents chunking as a technique in conceptual modeling to improve modeling effectiveness and efficiency. It is not clear if the term ‘chunking’ refers to splitting up the model in smaller subparts, or splitting up the modeling process in smaller subparts.⁶ If the latter applies, this is similar to part 1 of the SPMT, but without considering the increase in cognitive load for aggregation and strategy building. Next, the process of constructing process models is described by (Soffer et al. 2012) as a sequence of two phases. A modeler first builds a mental model of the process to be represented in the diagram and then the mental model is mapped onto the constructs of a formal process modeling language in order to build the process model. The focus of the paper is on optimizing the formation of the mental model as a prerequisite to increase semantic quality of the process model. It is advised to lower cognitive load by building this mental model chunk by chunk. Furthermore, the paper suggests to examine the impact of model structuredness on domain understanding. It does not consider however a structured approach towards the chunking.

Structuredness Most cognition inspired literature on structuredness in conceptual modeling describes a relation between structuredness *of the model* and some other characteristic of the model. A model is considered well-structured if every branch of a split of a certain type is joined in a single join construct of the same type. For example, well-structuredness is proposed to have an impact on correctness because it makes it easier for the modeler to navigate through

⁶ In literature the term ‘chunk’ has different meanings: a part of a process, a part of an artifact, a collection of information in memory. Therefore, the term was used sparsely in this paper. Splitting up a process in parts was named ‘serialization’ and a collection of information is stored in memory as a ‘cognitive schema’.

the model that was build so far, which reduces the chance on introducing errors (Laue and Mendling 2010). Further, also nesting depth of split and join constructs is an aspect of structuredness and a greater nesting depth is proposed to imply greater model complexity (Gruhn and Laue 2006). Finally, (Zugal et al. 2013) describe the effect of hierarchical structuring (i.e., decomposing the model in sub-models) on expressiveness and understandability. It is proposed that hierarchical models suffer from two opposing effects: (i) abstraction decreases mental effort⁷ by hiding information and supporting pattern recognition, but (ii) fragmentation increases mental effort because of attention switch and integration effort. The opposing effects of abstraction and fragmentation are described in part 1 of the SPMT. Serialization of the modeling process allows focusing on one part of the model at a time (abstracting from the other parts), but there is a cost of aggregating the different parts (integration effort).

Except for structuredness of the model, there is also literature about structuredness of the input (e.g., a textual case description). Pinggera et al. (2010b) propose that a breadth-first ordering of text was best suited to yield good results. Breadth-first ordering was defined as “*begins with the start activity and then explains the entire process by taking all branches into account*” (p. 448). It corresponds with the flow-oriented approach to modeling described in this paper (whereas depth-first can be matched with the happy path first modeling style). Cognitive fit however, was not considered in their work.

Cognitive fit In their summarizing framework of cognition variables for conceptual modeling, it is proposed in (Stark and Esswein 2012) that problem-solving skills of the modeler have to match with the task of modeling and that this (mis)match can cause effects on the resulting conceptual model. Regrettably, this was not further investigated or tested. Further, (Agarwal et al. 1996a, b, 2000) propose that an object-oriented representation is not universally more usable or less usable than other representations. Cognitive fit and prior method knowledge should be considered to evaluate the usability of object-oriented representations. This is fully in line with part 3 of the SPMT, but the focus is not on object-oriented modeling as a process (which would be similar to structured modeling), though it is on object-oriented representations. Therefore the research centers on extraneous load, rather than intrinsic load (as is the case for the SPMT). Lastly, the understandability of a process model is proposed to be more impacted by personal factors, than by model factors (Reijers and Mendling 2011). This work also recognizes

the need for studying cognitive fit, albeit in the context of model reading.

Guidelines for modeling Most of the work mentioned above describes causal effects between various variables. The emphasis is on predicting, rather than explaining. (Gregor 2006) states that both theories for explaining and theories for predicting can be used as input for a theory for design and action. The ambition of the SPMT is also to describe the necessary knowledge in order to build a prescriptive theory for process modeling. Two of such prescriptive theories were found already in literature. (Mendling et al. 2010) propose seven process-modeling guidelines (7PMG) that are based on strong empirical evidence and are simple enough to be used by practitioners. Guideline 4 proposes to model as structured as possible. The guidelines of modeling (GOM) presented in (Becker et al. 2000) are less concrete guidelines that claim to assure the quality of process models beyond syntactical aspects. Both prescriptive theories, however, provide recommendations about desired process model properties that can be guarded during modeling without considering the cognitive fit of the recommendation with the characteristics of the modeler.

8 Discussion

The research described in this paper is limited in several ways. Nevertheless, the SPMT can be valuable in practice and for research. The limitations and implications of the presented research are discussed below. In order to work on the limitations and to increase its usefulness, future research is described in this section as well.

8.1 Limitations

8.1.1 Limited ecological validity

The observations and impressions that were used as input for building the Structured Process Modeling Theory (SPMT) stem from modeling sessions with master students. Furthermore, they were given an artificial case description. In real life modeling sessions the modelers seldom start from a structured case description such as the one that was used for the observations. They rather use direct observation, interview transcripts, notes and pictures from whiteboard sessions, etc. Finally, only syntactic quality was considered when evaluating the produced models. Because the Structure Process Modeling Theory (SPMT) was only inspired by these observations and impressions, but it was compiled from existing cognitive theories that apply widely, there is no reason to suspect that the SPMT does not apply in a more realistic setting. However, the limited ecological validity of the

⁷ Whereas mental load is defined as the amount of information *needed to store* in working memory at a certain time to perform a task, mental effort can be regarded as the amount of information that is *actually stored* in working memory during the execution of a task.

observations and impressions may have hindered the disclosure of all relevant effects of serialization on cognitive load.

8.1.2 Limited content and construct validity

The SPMT and its constructs and associations may have limited content validity. First, only (structured) serialization was investigated (in accordance with the observations), no other general problem solving techniques were considered. Second, the assessment of syntactic quality that formed the base of Impressions 2 and 3, was partly subjective. It is possible that the impressions are not entirely accurate, which may have hindered the disclosure of certain relevant effects of serialization on cognitive load. The credibility of the theory however, is guaranteed by the deductive approach, which builds on existing, established theories. Additional observations in several different settings can help to assess the content validity of the SPMT in the future. Third, although the constructs are clearly described in the SPMT, some of them may be hard to transform into a variable that can be measured properly (i.e., with high construct validity). For example, to date there are no known metrics that measure intrinsic cognitive load separately from extraneous or germane cognitive load, not to mention metrics for the artificially separated constructs of intrinsic cognitive load for modeling, for aggregating and for strategy building of the SPMT.

8.2 Implications

8.2.1 Implications for practice

We have experienced that in practice a lot of modelers (experienced and inexperienced) often struggle with the complexity of the case at hand. Although it was observed how some inexperienced modelers automatically turned to a structuring approach and although the structuring techniques are not particularly hard to apply, other modelers do not seem to structure their modeling processes. A slower constructed and lower quality process model was observed, that – according to the SPMT – can be a consequence of applying an undirected process modeling strategy. The SPMT will help building the knowledge that is necessary to (i) be aware of suboptimal modeling conditions (e.g., when modelers apply a structuring technique that does not fit with the task and with their characteristics as a problem solver), (ii) train the modelers to use an *individually fitting structured serialization technique* for process modeling in order to raise effectiveness and efficiency, (iii) provide the means to better support the modelers in handling complexity (e.g., differentiated or adaptive tools that support structured process modeling in accordance to different modeling approaches or with changing features for consecutively modeling phases).

8.2.2 Implications for research

The SPMT is novel in its recognition of cognitive fit between modeling task and modeler characteristics of the proposed modeling structuring technique for optimal effectiveness and efficiency. This fundamental focus point can inspire researchers in other research domains to develop adaptive techniques as well. The SPMT can be applied in a broader context and can add to the existing cognitive theories about serialization as a generic problem solving technique. Furthermore, within the domain of process modeling, the (descriptive) SPMT is considered as a first, necessary step towards the development of a prescriptive theory that will further extend our knowledge about the effect and applicability of structuring and individual fit during process modeling.

8.3 Future work

8.3.1 More extensive evaluation of the SPMT

The SPMT needs to be tested more profoundly. The propositions will be converted into empirically testable *hypotheses*, accurate *metrics* need to be developed for each of the involved variables of these hypotheses and new series of *observational modeling sessions* will be performed in which these variables are measured and correlations are calculated.

Because of limited ecological validity of observations and impressions that were used as input for the development of the SPMT, the *external validity* of the SPMT itself needs to be examined further. Current observations were made on master student behavior where the prior knowledge of existing modeling techniques is assumed to be very low. Therefore, one of the factors to examine is how much this prior knowledge of experienced modelers influences the observed effects. Cognitive theories suggest, that retraining an experienced modeler to use a different technique than the ones he is used to, consumes a lot of germane load, which is expressed in an initial decrease of performance (this is called the Expert Reversal Effect, Kalyuga et al. 2003).

8.3.2 Development of prescriptive theory and a method for cognitive effective and efficient process modeling

Furthermore, in order to convert the SPMT, which is a (descriptive) theory for explaining, towards a (prescriptive) theory for design and action, next actions still need to be undertaken.

First, it should be examined if modelers can be trained to apply the three aspects of the SPMT. This requires the development of a *method* (i.e., prescribing how to construct the process model according to the *individually fitting structured serialized process modeling* principle of the SPMT) and a *treatment* (i.e., describing how to train modelers to use that

method). Subsequently, the degree of *treatment adoption* in an experimental context can be measured.

Second, it should be examined if the positive effect on load, overload and by consequence accuracy and speed manifests itself indeed when modelers are trained to apply the developed method based on the three aspects of the SPMT (i.e., *testing causality*). This requires reformulating the hypotheses into *causal* relations between the variables and the set-up of a *controlled comparative experiment* to isolate the effect of the treatment in the measurements of these causal relations.

9 Conclusion

In experimental modeling sessions with master students that were instructed to construct a process model based on the same textual description, we noted various differences in the produced process models. For example different syntactical errors were found in the models. Some errors were made consistently and can be caused by a lack of knowledge, whereas other errors seem to be a result of cognitive failures during process modeling. For the development of tools to help modelers to reduce the latter type of errors, knowledge is needed about why, how and when these failures occur and impact the accuracy and speed of the modeling process. This knowledge was not readily available and therefore it is provided in this article in the form of an explanatory theory.

The developed theory is called the Structured Process Modeling Theory (SPMT) and consists of three parts. Based on observations and impressions, and on explanations from cognitive literature, it describes how the probability of cognitive overload can be reduced by (i) *serializing* the modeling process, and (ii) *structuring* that serialization (iii) *in a way that fits* with the characteristics of the modeler. The research methodology of theory building based on observations and impressions and using components of existing theories should warrant the utility of the newly developed theory. However, a brief evaluation of the theory and a discussion on the limitations are described in Sections 6 and 8.

This work is important on three levels. Firstly, it provides new *knowledge* on the relation between serializing, structuring and fit of the process modeling approach on the one hand and cognitive effectiveness and efficiency on the other hand. It explains why some modelers struggle (more than others) with the complexity of constructing a process model. This knowledge in itself is useful because it facilitates the selection of suitable modelers or modeling approaches for concrete projects.

Secondly, it is a step towards the development of a *method* that aims at supporting modelers to select and implement an optimized process modeling strategy that fits with the task at hand and with the characteristics of the modeler. If the theory is true, if a modeler can be trained to modify his modeling

technique and if this change of approach preserves the described effects, the SPMT has the potential to significantly and positively impact the quality of future process modeling projects.

Lastly, the knowledge and the method can be used to develop *tool support* for process modeling that is differentiated (i.e., the features of the tool can differ according to the use(r) of the tool) or adaptive (i.e., the features of the tool change during the modeling process, for example to support consecutive phases of modeling). Tools can ease a modeler's transition to an improved process-modeling technique and can aid the application of such a technique.

Acknowledgments We wish to express our gratitude to the developers of Cheetah Experimental Platform (CEP), to all the students and administrators involved in the observational modeling sessions, and to Martin Valcke who pointed out some relevant cognitive theories that apply for the observed behavior.

Appendix A. Recorded modeling operations in CEP

Cheetah Experimental Platform records every interaction with the tool while modeling. Figure 13 displays a list that contains the subset of the recorded operations that are represented in the PPMChart visualization, together with the used symbols in the visualization.

Appendix B. Questionnaire used in the observational modeling sessions

1. What is your gender? (M/F)
2. How old are you?
3. What is your native language?
4. What is your current profession? (Student/Other)
If you have selected 'Other' please specify which profession.
5. How many years ago did you start process modeling?
6. How many process models have you analyzed or read within the last 12 months? (A year has about 250

■ CREATE_ACTIVITY	■ DELETE_ACTIVITY
● CREATE_START_EVENT	● DELETE_START_EVENT
● CREATE_END_EVENT	● DELETE-END_EVENT
◆ CREATE_AND	◆ DELETE_AND
◆ CREATE_XOR	◆ DELETE_XOR
▶ CREATE_EDGE	▶ DELETE_EDGE
■ MOVE_ACTIVITY	■ NAME_ACTIVITY
● MOVE_START_EVENT	■ RENAME_ACTIVITY
● MOVE_END_EVENT	▼ NAME_EDGE
◆ MOVE_AND	▼ RENAME_EDGE
◆ MOVE_XOR	▼ RECONNECT_EDGE

Fig. 13 Model interactions recorded in CEP and visualized in the PPMChart

- workdays. In case you read one model per day, this would sum up to 250 models per year)
7. How many process models have you created or edited within the last 12 months?
 8. How many activities did all these models have on average?
 9. How many workdays of formal training on process modeling have you received within the last 12 months? (This includes e.g. university lectures, certification courses, training courses. 15 weeks of a 90 min university lecture is roughly 3 work days)
 10. How many workdays of self-education have you made within the last 12 months? (This includes e.g. learning-by-doing, learning-on-the-fly, self-study of textbooks or specifications)
 11. Which education program are you following?_(OML/BIS/IM/CSE/Other)
If you have selected ‘Other’ please specify which education program.
 12. Which process modeling languages have you used before?
(Aris (express)/BPEL/BPMN/BPM|one/Petri Nets-Colored Petri Nets-CPN Tools/Tibco-COSA/Workflow Nets-WoPeD/Other).
If you have selected ‘Other’ please specify which modeling languages.
 13. I consider myself being a process-modeling expert. (*)
 14. I have troubles reading English texts. (*)
 15. I have troubles understanding English text. (*)
 16. I am familiar with processes in the financial domain. (*)
 17. I am familiar with mortgage approval processes (*)
 18. I have created process models in the financial domain before. (*)
 19. I have created process models for a mortgage process before. (*)
 20. Overall, I am familiar with the BPMN modeling language. (*)
 21. I feel confident in understanding process models created with the BPMN modeling language. (*)
 22. I feel competent in using the BPMN modeling language for process modeling. (*)
 23. How many months ago did you start using BPMN? (The first version of BPMN stems from May 2004, i.e. 60 months until May 2009)
 24. Have you completed the assignment? (yes/no)
 25. How would you assess the mental effort for completing the modeling task? (1–10)
 26. I have experienced language difficulties when reading and understanding the case description. (*)
 27. I found the case description easy to understand. (*)
 28. I found the case description complex and difficult to follow. (*)
 29. The case description was ambiguous. (*)
 30. The case was clearly described. (*)
 31. Have you experienced any disturbances during the execution of the modeling task? (yes/no)
If yes, which (kind of) disturbances have you experienced?
 32. Learning to use the BPMN modeling language is easy for me. (*)
 33. I find it easy to get the BPMN modeling language to do what I want it to do. (*)
 34. My interaction with the BPMN modeling language is clear and understandable. (*)
 35. I find the BPMN modeling language to be flexible to interact with. (*)
 36. It would be easy for me to become skillful at using the BPMN modeling language. (*)
 37. I find the BPMN modeling language easy to use. (*)
 38. Learning to operate the modeling tool is easy for me. (*)
 39. I find it easy to get the modeling tool to do what I want it to do. (*)
 40. My interaction with the modeling tool is clear and understandable. (*)
 41. I find the modeling tool to be flexible to interact with. (*)
 42. It would be easy for me to become skillful at using the modeling tool. (*)
 43. I find the modeling tool easy to use. (*)
- (*) This item had to be scored on a 7-point Likert scale (strongly agree – agree – somewhat agree – neutral – somewhat disagree – disagree – strongly disagree)

Fig. 14 Indicated prior domain knowledge (PDK) and prior modeling knowledge (PMK)

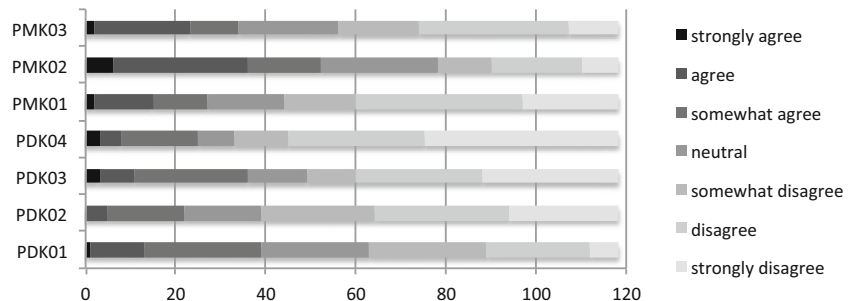
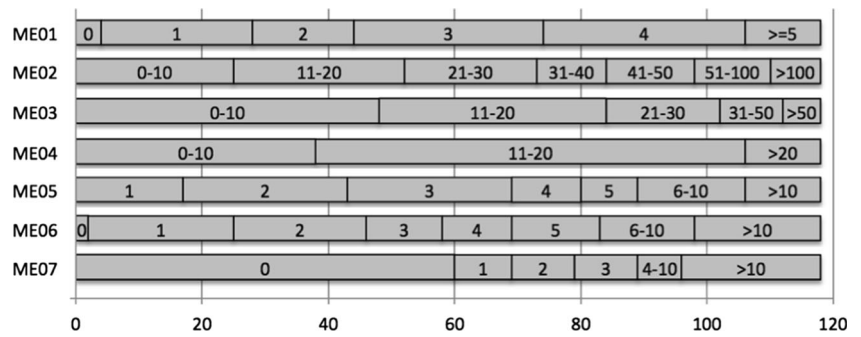


Fig. 15 Indicated modeling experience (ME)



Appendix C. Prior knowledge and modeling experience of the participants

Figures 14 and 15 show information about the prior knowledge and modeling experience of the participants of the observational modeling sessions described in Section 3.

Prior domain knowledge (PDK)

- I am familiar with processes in the financial domain. (PDK01)
- I am familiar with mortgage approval processes (PDK02)
- I have created process models in the financial domain before. (PDK03)
- I have created process models for a mortgage process before. (PDK04)

Prior modeling knowledge (PMK)

- Overall, I am familiar with the BPMN modeling language. (PMK01)
- I feel confident in understanding process models created with the BPMN modeling language. (PMK02)
- I feel competent in using the BPMN modeling language for process modeling. (PMK03)

Modeling experience (ME)

- How many years ago did you start process modeling? (ME01)
- How many process models have you analyzed or read within the last 12 months? (A year has about 250 workdays. In case you read one model per day, this would sum up to 250 models per year) (ME02)
- How many process models have you created or edited within the last 12 months? (ME03)
- How many activities did all these models have on average? (ME04)
- How many workdays of formal training on process modeling have you received within the last 12 months? (This includes e.g. university lectures, certification courses,

training courses. 15 weeks of a 90 min university lecture is roughly 3 work days) (ME05)

- How many workdays of self-education have you made within the last 12 months? (This includes e.g. learning-by-doing, learning-on-the-fly, self-study of textbooks or specifications) (ME06)
- How many months ago did you start using BPMN? (ME07)

Appendix D. Observed syntactic errors in the observational modeling sessions

Table 5 lists the observed syntactic errors in the produced models of the observational modeling sessions. Most errors were considered to potentially be caused by cognitive failure (marked ‘cognitive’ in Table 5), whereas the consistent absence of (a certain kind of) gateways was assumed to originate

Table 5 Observed syntactic errors in the observational modeling sessions

Syntactic error	Type
Contains no end event (but does contain a start event)	Cognitive
Contains an end event in the middle (outgoing edge on end event)	Cognitive/knowledge
Some, but not all of the paths are not closed (missing end event?)	Cognitive
Contains no split gateways at all	Knowledge
Contains either no XOR or no AND split gateways at all	Knowledge
Forgot some, but not all split gateways	Cognitive
Contains no join gateways at all	Knowledge
Contains either no XOR or no AND join gateways at all	Knowledge
Forgot some, but not all join gateways	Cognitive
One gateway combines a join and split feature	Cognitive/knowledge
Wrong type of join combined with a certain split	Cognitive
Gateway with only one ingoing and one outgoing edge	Cognitive
Wrong nesting of gateways	Cognitive
AND and XOR are joined together in one join gateway	Cognitive

in a lack of knowledge of the modeling language rules (marked ‘knowledge’ in Table 5).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abecker, A., Bernardi, A., & Hinkelmann, K. (2000). Context-aware, proactive delivery of task-specific information: the KnowMore project. *Information Systems Frontiers*, 2(3–4), 253–276.
- Agarwal, R., Sinha, A. P., & Tanniru, M. (1996a). Cognitive fit in requirements modeling: a study of object and process methodologies. *Journal of Management Information Systems*, 13(2), 137–162.
- Agarwal, R., Sinha, A. P., & Tanniru, M. (1996b). The role of prior experience and task characteristics in object-oriented modeling: an empirical study. *International Journal of Human-Computer Studies*, 45(6), 639–667.
- Agarwal, R., De, P., Sinha, A. P., & Tanniru, M. (2000). On the usability of OO representations. *Communications of the ACM*, 43(10), 83–89.
- Atkinson, R., & Shiffrin, R. (1968). Human memory: a proposed system and its control processes. *Psychology of Learning and Motivation*, 2, 89–195.
- Avgerou, C. (2000). Information systems: what sort of science is it? *Omega*, 28(5), 567–579.
- Bacharach, S. (1989). Organizational theories: some criteria for evaluation. *Academy of Management Review*, 14(4), 496–515.
- Bannert, M. (2002). Managing cognitive load—recent trends in cognitive load theory. *Learning and Instruction*, 12(1), 139–146.
- Becker, J., Rosemann, M., & Von Uthmann, C. (2000). Guidelines of business process modeling. In W. Van der Aalst, J. Desel, & A. Oberweis (Eds.), *Business process management. Models, techniques, and empirical studies. Part I*. (Vol. LNCS 1806, pp. 30–49). Springer Berlin Heidelberg.
- Brown, R. A., Recker, J. C., & West, S. (2011). Using virtual worlds for collaborative business process modeling. *Business Process Management Journal*, 17(3), 546–564.
- Chakraborty, S., Sarker, S., & Sarker, S. (2010). An exploration into the process of requirements elicitation: a grounded approach. *Journal of the Association for Information Systems*, 11(4), 212–249.
- Claes, J., Vanderfeesten, I., Pinggera, J., Reijers, H. A., Weber, B., & Poels, G. (2015). A visual analysis of the process of process modeling. *Information Systems and e-Business Management*, 13(1), 147–190.
- Cowan, N. (2010). The magical mystery four: how is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1), 51–57.
- Davies, I., Green, P., Rosemann, M., Indulska, M., & Gallo, S. (2006). How do practitioners use conceptual modeling in practice? *Data & Knowledge Engineering*, 58(3), 358–380.
- De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, 38(2), 105–134.
- Dumas, M., Rosa, M. La, Mendling, J., Reijers, H. A., & La Rosa, M. (2013). *Fundamentals of business process management*. Springer Berlin Heidelberg.
- Felder, R., & Silverman, L. (1988). Learning and teaching styles in engineering education. *Engineering Education*, 78(June), 674–681.
- Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: molar versus modular presentation of solution procedures. *Instructional Science*, 32(1–2), 33–58.
- Godfrey-Smith, P. (2009). *Theory and reality: An introduction to the philosophy of science*. Chicago:University of Chicago Press.
- Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, 30(3), 611–642.
- Grover, V., Lyytinen, K., Srinivasan, A., & Tan, B. C. Y. (2008). Contributing to rigorous and forward thinking explanatory theory. *Journal of the Association for Information Systems*, 9(2), 40–47.
- Gruhn, V., & Laue, R. (2006). Complexity Metrics for Business Process Models. In W. Abramowicz & H. C. Mayr (Eds.), *9th international conference on business information systems (BIS 2006)* (Vol. LNI85, pp. 1–12). Klagenfurt: Gesellschaft für Informatik (GI).
- Hevner, A. R. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105.
- Hoppenbrouwers, S. J. B. A., Proper, H. A., & Van der Weide, T. P. (2005). A fundamental view on the process of conceptual modeling. In L. Delcambre, C. Kop, H. C. Mayr, J. Mylopoulos, & O. Pastor (Eds.), *Conceptual modeling - ER 2005. 24th International Conference on Conceptual Modeling, Klagenfurt, Austria, October 24–28, 2005. Proceedings*. (Vol. LNCS 3716, pp. 128–143). Springer Berlin Heidelberg.
- Juran, J. M., & Gryna, F. M. (1988). In J. M. Juran, & F. M. Gryna (Eds.), *Juran's quality control handbook*. (4th ed.). McGraw-Hill.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23–31.
- Khatri, V., Vessey, I., Ram, S., & Ramesh, V. (2006a). Cognitive fit between conceptual schemas and internal problem representations: the case of geospatial-temporal conceptual schema comprehension. *IEEE Transactions on Professional Communication*, 49(2), 109–127.
- Khatri, V., Vessey, I., Ramesh, V., Clay, P., & Park, S.-J. (2006b). Understanding conceptual schemas: exploring the role of application and IS domain knowledge. *Information Systems Research*, 17(1), 81–99.
- Kim, J., Hahn, J., & Hahn, H. (2000). How do we understand a system with (so) many diagrams? Cognitive integration processes in diagrammatic reasoning. *Information Systems Research*, 11(3), 284–303.
- Kock, N., Verville, J., Danesh-Pajou, A., & DeLuca, D. (2009). Communication flow orientation in business process modeling and its effect on redesign success: results from a field study. *Decision Support Systems*, 46(2), 562–575.
- Krishnan, M., Mukhopadhyay, T., & Zubrow, D. (1999). Software process models and project performance. *Information Systems Frontiers*, 1(3), 267–277.
- Laue, R., & Mendling, J. (2010). Structuredness and its significance for correctness of process models. *Information Systems and e-Business Management*, 8(3), 287–307.
- Lindland, O. I., Sindre, G., & Solvberg, A. (1994). Understanding quality in conceptual modeling. *Software, IEEE*, 11(2), 42–49.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251–266.
- Mendling, J. (2008). *Metrics for process models: Empirical foundations of verification, error prediction and guidelines for correctness. LNBIP 6*. (J. Mylopoulos & N. M. Sadeh, Eds.). Berlin Heidelberg: Springer Science & Business Media.
- Mendling, J., Reijers, H. A., & Van der Aalst, W. M. P. (2010). Seven process modeling guidelines (7PMG). *Information and Software Technology*, 52(2), 127–136.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.
- Nagel, E. (1979). *The structure of science*. Indianapolis:Hackett Publishing Company.
- Naumann, J. D. (1986). The role of frameworks in MIS research. In *Proceeding of the 1986 DSI National Meeting, Honolulu, Hawaii* (pp. 569–571).

- Nelson, H. J., Poels, G., Genero, M., & Piattini, M. (2012). A conceptual modeling quality framework. *Software Quality Journal*, 20(1), 201–228.
- Neuberg, S., & Newsom, J. (1993). Personal need for structure: individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, 65(1), 113–131.
- Numanaker, J. F., & Chen, M. (1990). Systems development in information systems research. In *System Sciences, 1990., Proceedings of the Twenty-Third Annual Hawaii International Conference on* (Vol. 3, pp. 631–640). Kailua-Kona, HI: IEEE. doi:10.1109/HICSS.1990.205401.
- Paas, F. G. W. C., Renkl, A., & Sweller, J. (2003a). Cognitive load theory and instructional design: recent developments. *Educational Psychologist*, 38(1), 1–4.
- Paas, F. G. W. C., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003b). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63–71.
- Pinggera, J., Zugal, S., & Weber, B. (2010a). Investigating the Process of Process Modeling with Cheetah Experimental Platform. In B. Mutschler, J. Recker, R. Wieringa, J. Ralyte, & P. Plebani (Eds.), *ER-POIS 2010. Proceedings of the 1st International Workshop on Empirical Research in Process-Oriented Information Systems*. (Vol. CEUR -WS 6, pp. 13–18).
- Pinggera, J., Zugal, S., Weber, B., Fahland, D., Weidlich, M., Mendling, J., & Reijers, H. A. (2010b). How the structuring of domain knowledge helps casual process modelers. In J. Parsons, M. Saeki, P. Shoval, C. Woo, & Y. Wand (Eds.), *Conceptual modeling – ER 2010. 29th international conference on conceptual modeling, vancouver, BC, Canada, November 1–4, 2010. Proceedings*. (Vol. LNCS 6412, pp. 445–451). Springer Berlin Heidelberg.
- Pinggera, J., Soffer, P., Fahland, D., Weidlich, M., Zugal, S., Weber, B., ... Mendling, J. (2013). Styles in Business Process Modeling: An Exploration and a Model. *Software & Systems Modeling*, (published online, May 2013).
- Pithers, R. T. (2002). Cognitive learning style: a review of the field dependent-field independent approach. *Journal of Vocational Education & Training*, 54(1), 117–132.
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12(1), 61–86.
- Popper, K. (2005). *The logic of scientific discovery*. Electronic version: Taylor & Francis.
- Recker, J. C., Mendling, J., & Hahn, C. (2013). How collaborative technology supports cognitive processes in collaborative process modeling: a capabilities-gains-outcome model. *Information Systems*, 38(8), 1031–1045. doi:10.1016/j.is.2013.04.001.
- 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: Systems and Humans*, 41(3), 449–462.
- Rockwell, S., & Bajaj, A. (2005). COGEVAL: applying cognitive theories to evaluate conceptual models. *Advanced Topics in Database Research*, 4, 255–282.
- Rogers, Y., & Rutherford, A. (1992). *Models In the mind - theory, perspective, and application*. London: Academic Press.
- Sánchez-González, L., García, F., Ruiz, F., & Piattini, M. (2013). Toward a quality framework for business process models. *International Journal of Cooperative Information Systems*, 22(01), 1350003–1–15.
- Shaft, T. M., & Vessey, I. (2006). The role of cognitive fit in the relationship between software comprehension and modification. *MIS Quarterly*, 30(1), 29–55.
- Simon, H. A. (1996). *The sciences of the artificial*. MIT Press.
- Sinha, A. P., & Vessey, I. (1992). Cognitive fit: an empirical study of recursion and iteration. *IEEE Transactions on Software Engineering*, 18(5), 368–379.
- Soffer, P., Kaner, M., & Wand, Y. (2012). Towards understanding the process of process modeling: Theoretical and empirical considerations. In F. Daniel, K. Barkaoui, & S. Dustdar (Eds.), *Business process management workshops. BPM 2011 international workshops, clermont-ferrand, France, August 29, 2011, revised selected papers, part I*. (Vol. LNBIP 99, pp. 357–369). Springer Berlin Heidelberg.
- Sperling, G. (1963). A model for visual memory tasks. *Human Factors*, 5(1), 19–31.
- Stark, J., & Esswein, W. (2012). Rules from cognition for conceptual modelling. In P. Atzeni, D. Cheung, & S. Ram (Eds.), *Conceptual modeling. 31st international conference ER 2012, florence, Italy, October 15–18, 2012. Proceedings*. (Vol. LNCS 7532, pp. 78–87). Springer Berlin Heidelberg.
- Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, 12(2), 257–285.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12(3), 185–233.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.
- Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Englewood Cliffs: Educational Technology.
- Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review*, 17(2), 147–177.
- Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: instructional design for complex learning. *Educational Psychologist*, 38(1), 5–13.
- Vanderfeesten, I., Cardoso, J., Mendling, J., Reijers, H. A., & Van der Aalst, W. M. P. (2007). Quality Metrics for Business Process Models. In *BPM and Workflow handbook* (pp. 179–190). Future Strategies Inc.
- Vessey, I. (1991). Cognitive fit: a theory-based analysis of the graphs versus tables literature. *Decision Sciences*, 22(2), 219–240.
- Vessey, I., & Galletta, D. (1991). Cognitive fit: an empirical study of information acquisition. *Information Systems Research*, 2(1), 63–84.
- Vessey, I., & Weber, R. (1986). Structured tools and conditional logic: an empirical investigation. *Communications of the ACM*, 29(1), 48–57.
- Weber, R. (2012). Evaluating and developing theories in the information systems discipline. *Journal of the Association for Information Systems*, 13(1), 1–30.
- Weick, K. E. (1989). Theory construction as disciplined imagination. *Academy of Management Review*, 14(4), 516–531.
- Weske, M. (2007). *Business process management: Concepts, languages, architectures* (1st ed.). Springer Berlin Heidelberg.
- Whetten, D. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490–495.
- Wiedenbeck, S., & Ramalingam, V. (1999). A comparison of the comprehension of object-oriented and procedural programs by novice programmers. *Interacting with Computers*, 11(3), 255–282.
- Witkin, H. A., & Goodenough, D. R. (1981). *Cognitive styles: Essence and origins, field dependence and field independence*. Michigan: International Universities Press.
- Xiao, L., & Zheng, L. (2012). Business process design: process comparison and integration. *Information Systems Frontiers*, 14(2), 363–374.
- Zugal, S., Soffer, P., Haisjackl, C., Pinggera, J., Reichert, M., & Weber, B. (2013). Investigating expressiveness and understandability of hierarchy in declarative business process models. *Software & Systems Modeling*, (published online, June 2013).
- Zur Muehlen, M., & Recker, J. C. (2008). How much language is enough? Theoretical and practical use of the business process modeling notation. In Z. Bellahsene & M. Léonard (Eds.), *Advanced information systems engineering. 20th international conference, CAiSE 2008 montpellier, France, June 16–20, 2008 proceedings. LNCS 5074* (Vol. LNCS 5074, pp. 465–479). Springer Berlin Heidelberg.

Jan Claes is a PhD student and teaching assistant at Ghent University (UGent) since 2009. He is member of the UGent MIS research group, which is part of the Faculty of Economics and Business Administration. During his research visit in 2012 at Eindhoven University of Technology (TU/e), he started a close collaboration with the Information Systems subdepartment of which he became a member in 2013. As a result, he will complete his PhD in 2015 in a joint setting between UGent and TU/e. After obtaining a master degree in Industrial Engineering (University College of Ghent, 2005) and in Business Economics (UGent, 2006), he has worked consecutively as a programmer, analyst and team leader in the human resources and retail sector. Besides the research into the process of process modeling presented in this paper, he has also worked on event log merging in the process mining research domain. More information can be found at www.janclaes.info.

Irene Vanderfeesten is an assistant professor in Business Process Management and Information Systems at Eindhoven University of Technology (TU/e). She received her MSc degree in Computer Science and her PhD degree in Industrial Engineering from TU/e in 2004 and 2009, respectively. Before joining the Information Systems group of the department of Industrial Engineering and Innovation Sciences at TU/e in 2010 she also worked as an IT consultant in the banking and insurance sector. Her current research interests include business modeling; business process redesign and improvement; workflow management; and human aspects of business process management and information systems.

Frederik Gailly is an assistant professor at Ghent University. He is member of the UGent MIS research group, which is part of the Faculty of Economics and Business Administration. Prof. Gailly's research focuses on enterprise ontologies and on how they can be applied in the context of conceptual modeling. This research focus has resulted in the reengineering of existing enterprise ontologies (e.g., the Resource Event Agent Enterprise ontology), the development of ontology-based conceptual modeling approaches and the evaluation and integration of enterprise modeling languages. Currently Prof Gailly is (co-)supervising 5 PhD projects of which 2 focus on ontology-driven conceptual modeling. Prof Gailly has published in the Journal of Information Systems (JIS) and Information Systems Journal (ISJ) and has been guest editor for a special issue on Process Modeling for the International Journal of Accounting

Information Systems (IJAIS). He has presented his work in some well-known high-quality academic conferences in Business Informatics (e.g. CAISE, Conceptual Modeling, ICEIS, Business Process Management).

Paul Grefen is a full professor in the School of Industrial Engineering at Eindhoven University of Technology since 2003. He chaired the Information Systems subdepartment from 2006 to 2014. He received his Ph.D. in 1992 from the University of Twente and held assistant and associate professor positions in the Computer Science Department. He was a visiting researcher at Stanford University in 1994. He has been involved in various European research projects as well as various projects within the Netherlands. He is an editor of the International Journal of Cooperative Information Systems. He is an editor and author of the books on the WIDE and CrossWork projects, and has authored books on workflow management and e-business. He is a member of the Executive Board of the European Supply Chain Forum. His current research covers architectural design of business information systems, inter-organizational business process management, and service-oriented business design and support. He teaches at the M.Sc. and Ph.D. levels at TU/e and at the executive level for TIAS business school.

Geert Poels is full professor at Ghent University's Faculty of Economics and Business Administration. He is founder and director of the UGentMIS Business Informatics research group. He was (co-)supervisor of 7 completed PhD's and is (co-)supervisor of 10 ongoing PhD's. He (co-)authored 91 publications listed in Web of Science, including 39 journal articles. In 2012 he was program chair of the sixth International IFIP Conference on Research and Practical Issues in Enterprise Information Systems (CONFENIS 2012). In 2014 he was conference chair of the eight European Conference on Information Systems Management and Evaluation (ECIME 2014). Prof. Poels is the developer of the DISTANCE approach to software measurement. He published widely on topics related to functional size measurement and quality of conceptual models. At Ghent University he initiated the Business Informatics research and developed the conceptual modeling, business ontology, and business modeling research lines. His current research also involves Enterprise Architecture, Business Process Management, and Service Science. He is editorial (review) board member of three academic journals.