Decision Support Systems 54 (2013) 1104–1118

Contents lists available at SciVerse ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

A study on the effects of routing symbol design on process model comprehension

Kathrin Figl [a],*, Jan Recker [b], Jan Mendling [c]

[a] Institute for Information Systems & New Media, UZAII, Augasse 2–6, A-1090, Vienna, Austria
[b] Queensland University of Technology, Information Systems School, 2 George Street, Brisbane QLD 4000, Australia
[c] Institute for Information Business, WU, Vienna University of Economics and Business, UZAII, Augasse 2–6, A-1090, Vienna, Austria

ARTICLE INFO

Article history:
Received 30 November 2011
Received in revised form 23 May 2012
Accepted 28 October 2012
Available online 2 November 2012

Keywords:
Process modeling
Notational design
Routing symbols
Comprehension
Cognitive effectiveness

ABSTRACT

Process modeling grammars are used to create models of business processes. In this paper, we discuss how different routing symbol designs affect an individual's ability to comprehend process models. We conduct an experiment with 154 students to ascertain which visual design principles influence process model comprehension. Our findings suggest that design principles related to perceptual discriminability and pop out improve comprehension accuracy. Furthermore, semantic transparency and aesthetic design of symbols lower the perceived difficulty of comprehension. Our results inform important principles about notational design of process modeling grammars and the effective use of process modeling in practice.

1. Introduction

Process models have been recognized as an effective means for documenting and communicating business processes, especially as a means for helping to discuss different viewpoints of stakeholders in projects such as the re-design of business processes [61] or the analysis and design of process-aware information systems [45]. Indications that process models indeed make a solid contribution in this area are, for instance, provided through a study of a large number of redesign projects [26].

Process models are created using process modeling grammars — sets of graphical symbols and rules describing how to connect the graphical symbols [78]. These symbols express relevant aspects of business processes, such as the *tasks* that have to be performed, the *actors* that are involved in the execution of tasks, relevant *data*, and, notably, the *control flow logic* that describes the logical and temporal order in which tasks are to be performed. In essence, the control flow logic of a business process defines those points in the process where parallel or alternative paths might be taken, or where such paths merge. Such routing points characterize the convergence or divergence of process flows.

In process modeling grammars, convergence or divergence semantics are typically expressed through grammatical symbols named "Gateways", "Connectors, or "Splits" and "Joins" [e.g., 62,74]. These routing symbols have been subjected to much academic debate. For instance, some scholars have argued that these symbols are ill-defined from a formal perspective [e.g., 74]. They have also been found to be a key reason for modeling errors such as violation of deadlock and synchronization rules [24], and further argued to lead to understandability problems with practitioners [35].

While all available process modeling grammars support the expression of convergence or divergence semantics in a business process, they utilize different visual symbols for doing so. This difference is of crucial importance for the quality of a grammar. In other domains, it has been found that the form of visual information representation can have a significant impact on the efficiency of information search, explicitness of information, and problem solving [28], the comprehension and recall of graphical models [11,41] and even perceived usability [67].

Our objective in writing this paper, therefore, is to develop insights about the role of routing symbol design in process modeling grammars. We study how model users understand models created with different visual routing symbol designs by drawing on a theory of effective visual notations [39]. We examine four principles of routing symbol design (perceptual discriminability, pop out, semantic transparency and aesthetics) that should lead to improved process model comprehension. We then present an experiment that tests the impact of the four principles of routing symbol design on process model comprehension in terms of accuracy, efficiency and perceived difficulty. The results demonstrate that the symbol design principles affect comprehension accuracy and difficulty in different ways. Comprehension efficiency is not affected by symbol design.

We proceed as follows. First, we review the literature on factors that influence the cognitive load of process model comprehension tasks. We then discuss relevant theoretical considerations pertaining to the visual design of routing symbols in process models and identify four relevant design principles. Next we describe our research model and the experimental design of the study. We then present our data analysis and results. After that, we discuss the results and limitations. We conclude

---

* Corresponding author. Tel.: +43 650 979 48 89; fax: +43 1 31336 90 4467.
E-mail addresses: kathrin.figl@wu.ac.at (K. Figl), j.recker@qut.edu.au (J. Recker), contact@mendling.com (J. Mendling).

0167-9236/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dss.2012.10.037

by summarizing the substantive as well as methodological contributions of this research.

2. Theoretical background

2.1. Cognitive load in comprehending control flow logic in process models

The division of labor in companies poses a considerable challenge to analyzing business processes in a department-spanning manner. Process models have been suggested as a means of abstraction for fostering understanding, transparency and communication of such complex processes. Even though models reduce business processes to their essential components, the creation and understanding of process models still requires high cognitive effort in itself due to the limited information processing capabilities of the human brain [76].

In light of this limitation, the key design principle for process models is to support rather than demand higher-level reasoning processes. This can, for instance, be achieved by conveying visual cues to the next logical step in reasoning about a process-related problem, or by representing process information (e.g., tasks to be performed) in the context of adjacent locations (e.g., in the context of the routing symbols that describe important business rules pertinent to the execution of the task).

Fig. 1 depicts a process model specified in the BPMN grammar [44] to illustrate how visual cueing is typically implemented in process modeling grammars. The model illustrates an E-mail voting process, based on the example given in [43]. The process consists of several activities that are executed according to a pre-defined order to reach the specific process goal (to resolve an issue). Fig. 1 shows that in this order, several divergence and convergence decisions are made, all represented by different types of gateways, in this case using a diamond shape symbol. Modeling "either/or" choices is done via so-called XOR-Split Gateways (e.g. "assess reasons for not voting" *or* skip this activity). After splitting control flow, it may be required to merge it later in the process. Exclusive choices can also be used to model repetition (loop with "election deadline has not yet passed"). Modeling concurrent activities is done via so-called AND-gateways (e.g. "review status of discussion" *and* "moderate E-mail discussion").

As the example shows, the diamond-shaped BPMN gateway symbols are intended to support the end users' interpretation and reasoning about the control flow logic of the process. While this reasoning process is fundamental to understanding the process, the body of literature on error analysis of process models suggests the existence of systematic reasoning fallacies concerning routing symbols [35]. We speculate that this may be traced back to systematic fallacies (so called 'illusory inferences') stemming from the visual design of the model or of the underlying process. These may occur when internally constructing or interpreting mental models on the basis of modeling-level connectives (like conjunctions, inclusive, or exclusive disjunctions) [22]. Concerning the example in Fig. 1, a variety of such cognitive errors could occur. Models readers could, for instance, misinterpret the AND-gateway and think both concurrent activities have to start at the same point of time, or they could confuse XOR and AND gateways if they find these gateway symbols difficult to discriminate perceptually.

Cognitive errors in reasoning about a process model relate to the *cognitive load* associated with the reasoning task. Cognitive load describes how much of the human working memory is used in learning and knowledge acquisition tasks [69]. Its importance stems from its limitations: The human working memory is the main bottleneck for cognitive tasks as its capacity is restricted to only 7+/−2 units of information at any point in time [38]. Recent literature estimates working memory capacity even lower to 3–4 elements [12]. The cognitive load of a task rises if a user has to pay attention to high amounts of relevant units of information, which in turn burdens or even overloads his/her working memory, and consequently impairs problem solving ability, learning and knowledge acquisition [69]. A variety of prior studies in the area of conceptual modeling have demonstrated that a reduction of cognitive load can lead to improvements in objective measures like comprehension [19] as well as in subjective perceptions on ease of understanding [31].

Cognitive load theory distinguishes *intrinsic* and *extraneous* cognitive load. Intrinsic cognitive load is determined by the complexity of information, i.e., the amount of elements, and their relations and interactions. In the process domain, intrinsic load pertains to the complexity of the modeled process, and thus beyond the control of the process analyst modeling a process. In contrast to that, extraneous cognitive load is determined by the way information is represented [25]. Even for exactly the same problem or task, the relative difficulty may vary depending on different problem representations [27]. Therefore, extraneous load pertains to the way a process is modeled and is thus subject to the design choices made when describing a process in a model.

Modeling design choices especially relate to notational aspects — the choice of different visual symbols for describing a process in the model. Precisely, the modifications may relate to the formal rules of a modeling grammar (its *primary notation*) or the way a specific model is visualized (its *secondary notation*) [50]. While the primary notation is normally prescribed by the specification of a modeling grammar, it has been shown that secondary notation influences process model comprehension, for instance, in terms of modularity [60], the grammatical style of text labels [36], or color highlighting [58]. These studies suggest that secondary notation is an important element in determining the extraneous cognitive load in understanding process models. Still, the research to date has focused on the secondary notation of models as a whole as opposed to the secondary notation of specific model elements — such as routing symbols, which is the focus of our work in this paper.

2.2. Effective visual design of notational symbols

To discuss the secondary notation of routing symbols in process models, we turn to a theory of effective visual notations proposed by

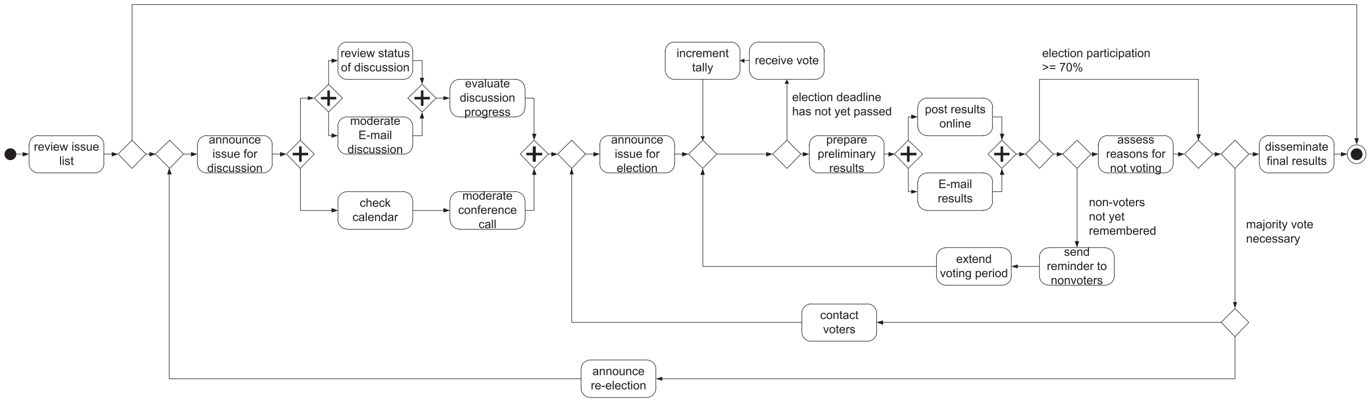


Fig. 1. Example for business process control flow logic (in BPMN).

Moody [39]. He suggests a set of principles for the visual design of notations used in information systems analysis and design.

Moody [39] uses the term "symbol set" as a synonym for the visual vocabulary of a modeling grammar. It compromises 1D, 2D and 3D graphic elements, such as example lines, areas and spatial relationships. For example, for depicting the routing behavior in business processes, process modeling grammars contain *symbols* made up of abstract graphics such as circles and diamond shapes.

In the following we will discuss the criteria which are relevant for discussing symbols and their demand of cognitive load. We build on the concepts from Moody's theory of effective notation design [39] and integrate it with an established framework on symbol characteristics [33].

2.3. Perceptual discriminability and pop out

Moody's theory [39] stipulates that notations of modeling grammars that fully exploit the range of visual variables (spatial dimensions like horizontal and vertical, as well as shape, size, color, brightness, orientation, and texture) have higher *visual expressiveness*. This proposition is based on the principles of *perceptual discriminability* and *pop out*.

Perceptual discriminability is defined as "the ease and accuracy with which graphical symbols can be differentiated from each other" [39]. It concerns the (dis-) similarity of shapes and connecting lines used in process models. The basic argument is that symbols expressing different domain semantics should be perceptually discriminable through the use of different shapes and lines.

A second important factor for the perceptual expressiveness of symbols is the number of feature dimensions on which they differ [80]. According to feature integration theory [72], symbols can be detected most easily amongst other symbols if they differ in one visual variable only (e.g., color but not shape or size). They are detected pre-attentively and hence "*pop out*", which means that they are easy to locate in a model. In contrast, search takes longer if the conjunction of several features is necessary to locate a symbol (e.g. searching a yellow circle among yellow and red squares and circles). Shape is the most important variable in this context [39]. As it is also used predominantly for discriminating objects, it is wise to use it as the primary distinguishing feature among different symbols. Additionally, redundant coding (e.g., a symbol is unique in shape and color) can help to prevent misinterpretations.

In consequence, symbols in a modeling grammar should differ appropriately and sufficiently in terms of visual variables in order to be perceptually discriminable. Further, they should pop out in one visual dimension to be easy to understand.

2.4. Semantic transparency

Semantic transparency describes whether the appearance of a symbol implies its corresponding concept. Moody [39] distinguishes between semantically immediate, opaque/conventional and perverse symbols on this continuum.

Icons, for example, belong to so-called concrete graphics that are easily associated with their referent real-world concepts, because there is a direct relationship between them and their meaning (e.g., a telephone icon to indicate a phone conference) [57]. Iconic representations for classes of activities could improve the understandability of process models as suggested by [34], but they are not yet commonly used. In contrast, abstract symbols typically used in process models have a rather distant relationship with their meaning, which is described as arbitrary [33]. Specifically, routing symbols used in process modeling can be characterized as abstract and not as concrete since they mainly use features such as different shapes (e.g., rectangles or circles) [33].

Semantic transparency closely relates to learnability of symbols. If users can rely on previously-learned associations and symbols are semantically transparent, they will be learned more easily. Recker and Dreiling [53], for instance, showed that the ability to understand and read a model from a specific process modeling grammar to another can be transferred surprisingly easy — partially because the semantic transparency of the two grammars compared was similar. However, many symbols used in modeling notations are abstract and have to be learnt explicitly [39]. In turn, the principle of semantic transparency would suggest that some routing symbols in process models are semantically more immediate than others. This fact should result in improved process model comprehension.

2.5. Aesthetics

Beyond perceptual discriminability, pop out and semantic transparency as important symbol design principles in modeling grammars, we realize that, up to some point, the design and appreciation of a symbol remains subject to subjective evaluation. This is because users may perceive different designs to be more *aesthetic* than others. Aesthetics are relevant for studying the effects of design on human understanding and evaluation, because users rate designs as more usable [71], and prefer it over others [63] if they perceive it as aesthetic. There is a history on exploring aesthetics in conceptual models and graphs [51] and indeed process models [15]; however, this stream of research has mainly focused on layout aesthetics of models as directed graphs. For instance, the authors in [51] propose rules to maximize symmetry and minimize edge crossings and bends to achieve more aesthetic diagrams. Design aesthetics pertaining to modeling grammar symbols, however, have not been examined.

Depending on the evaluated objects, researchers have looked at different criteria to determine aesthetic values, such as, for instance, balance or symmetry [29]. In general, good design should balance complexity and order. For instance, analyses of aesthetically pleasing screen designs revealed measures such as balance, proportions, symmetry or even distribution to be relevant [40]. "Ideal" proportions, combination of parts as a "unity" or "good" form and prototypicality are often considered design features related to aesthetic product design [75]. These studies suggest, in turn, that routing symbols that are perceived as more aesthetic than others will aid process model comprehension.

3. Hypotheses development

Our primary conjecture is that the visual design of routing symbols in process models will affect how well end users will comprehend the control flow of the modeled processes. More precisely, we argue that process model comprehension, measured in terms of accuracy, efficiency and task difficulty [30], is a function of four attributes of routing symbol design, viz., perceptual discriminability, pop out, semantic transparency, and aesthetics.

This conjecture builds on the argument that inefficient visual routing symbol design will impair the comprehension of a process model because it induces additional extraneous cognitive load into the comprehension task. If, however, the notational constructs are designed such that they communicate the meaning of a process more efficiently, model viewers should also be able to understand models better, faster and with more ease due to relatively lower extraneous cognitive load. We now discuss this conjecture in terms of four detailed propositions.

In our initial proposition, we explore how process model comprehension will vary depending on the perceptual discriminability of the routing symbols used in the model. Our argument is that routing symbols that have higher perceptual discriminability will induce lower extraneous cognitive load into the model comprehension task. Perceptual discriminability suggests that it will be easier and faster for model readers to perceptually process and differentiate the different visual components of the process model, thereby lowering the extraneous load of the cognitive processing task [39]. Lower extraneous load has been associated with increased comprehension accuracy [58], and decreased task difficulty perceptions [30]. Also, it has been argued (though not conclusively proven) that comprehension efficiency can be increased [17]. Formally, we state:

H1a. Process model comprehension *accuracy* will be higher when process models contain routing symbols with high perceptual discriminability.

H1b. Process model comprehension *efficiency* will be higher when process models contain routing symbols with high perceptual discriminability.

H1c. Process model comprehension *task difficulty* will be lower when process models contain routing symbols with high perceptual discriminability.

Second, we turn to pop out effects of routing symbols. According to feature integration theory, perceptual processing of visual symbols is strengthened if the symbols contain one visual variable (e.g., shape, or color) with a unique value [52]. Such symbols appear to “pop out” from all other symbols without requiring much cognitive effort. In turn, we can expect that process models that contain routing symbols that are differentiated from all other visual symbols by a unique value in one visual dimension are more easily and perceptually processed faster, in turn aiding the cognitive processing task. Formally, we state:

H2a. Process model comprehension *accuracy* will be higher when process models contain routing symbols that are perceived to pop out.

H2b. Process model comprehension *efficiency* will be higher when process models contain routing symbols that are perceived to pop out.

H2c. Process model comprehension *task difficulty* will be lower when process models contain routing symbols that are perceived to pop out.

In our third proposition, we explore how process model comprehension changes when the semantic transparency of routing symbols varies. The semantic transparency principle suggests that good visual symbols provide cues to the meaning of their content (“form implies content”). Semantically transparent symbols reduce extraneous cognitive load because their meaning can either be directly perceived or easily deduced [49]. This would again suggest a cognitive offloading effect in which perceptual processing of symbols aids the subsequent cognitive interpretation process , which should result in better process model comprehension (in terms of accuracy, efficiency and lower difficulty). We define the following three hypotheses:

H3a. Process model comprehension *accuracy* will be higher when process models contain routing symbols with high semantic transparency.

H3b. Process model comprehension *efficiency* will be higher when process models contain routing symbols with high semantic transparency.

H3c. Process model comprehension *task difficulty* will be lower when process models contain routing symbols with high semantic transparency.

In our last general proposition, we turn to the aesthetic design of routing symbols. Previous studies have demonstrated that users rate designs as more usable [71], and prefer it over others [63] if they perceive it as aesthetic. These findings suggest that task performance (such as the task of comprehending a process model) may be increased if the task artifact is aesthetically pleasing, because it generates a positive affective response [3]. Affective response has been shown to be associated with, for instance, task persistency [5], which relates to task performance. These findings suggest that comprehension performance (in terms of accuracy, efficiency and perceived difficulty) may also be affected by the extent to which routing symbols are perceived to be aesthetically pleasant. We state:

H4a. Process model comprehension *accuracy* will be higher when process models contain routing symbols that are perceived to be aesthetic.

H4b. Process model comprehension *efficiency* will be higher when process models contain routing symbols that are perceived to be aesthetic.

H4c. Process model comprehension *task difficulty* will be lower when process models contain routing symbols that are perceived to be aesthetic.

4. Research method

To test our hypotheses, we chose an experimental method as it affords higher internal validity than other methods [10]. Specifically, we selected a $1*4$ between-groups design that allowed us to focus on the four notational design factors whilst controlling for potentially confounding other variables (e.g., process modeling knowledge or domain complexity). We randomly assigned participants across groups and randomly assigned the order of tasks to control for learning effects.

4.1. Research design

Our design featured one between-subject factor (routing symbol design) and four dependent variables. Additionally, we considered the covariate prior process modeling method knowledge in our design.

The between-subjects factor, the design of routing symbols, had four levels (R_{EPC}, R_{BPMN}, R_{UML}, R_{YAWL}). The routing symbols were derived from four different popular process modeling grammars, viz., EPC, BPMN, UML AD and YAWL. These grammars are typically considered as appropriate representatives for the current set of available grammars in-use [56].

Concerning the manipulation of the variable “routing symbol”, we refrained from inventing artificial notations for routing symbols (e.g., with very low semantic transparency or with inflated perceptual discriminability). This might have maximized the likelihood of significant negative effects on comprehension but would have led to inflated risk of type-2 errors. Instead, we decided to use realistic examples of construct notational design sets based on the design of existing process modeling grammars, so as to maximize relevance for practice and ensure ecological validity of our experimental design. This design allowed us to examine whether different design solutions as used in practice are equally good concerning support of comprehension or whether there are relevant differences.

To operationalize the quality of the design of the four routing symbol sets according to the four notational design attributes semantic transparency, perceptual discriminability, pop out and aesthetics, we collected perceptual rating measurements of the routing symbols alongside these dimensions, which allowed us to separate ‘good’ from ‘bad’ designs.

As dependent variables, we used four measures to examine process model comprehension, consistent with prior work in this area [8,47]. First, we calculated the number of correct answers in a model comprehension task as a measure of comprehension *accuracy*. Second, we collected the task completion time as a measure of comprehension *efficiency*. Third, we measured the perceived difficulty to complete the model comprehension tasks as a measure of the *perceived cognitive load* associated with comprehending process models. Fourth, as routing symbols directly relate to the understanding of the control flow (one particular element of a process model), we added a second measure of perceived difficulty of process model comprehension — a judgment of the *difficulty of control flow comprehension* specifically.

Gemino and Wand [17] differentiate model comprehension tasks between *problem-solving tasks* in which newly developed mental models have to be integrated with deep knowledge structures and *comprehension tasks*. These two levels of measurements have also been referred to as “deep-level understanding” and “surface-level-understanding” [8]. As our research focuses on the effect of symbol design in otherwise informationally equivalent models on effective understanding and not on the different mental models evoked in users, ‘surface-level’ model comprehension tasks were the best choice of measurement. An additional factor for choosing "surface-level" comprehension tasks was that general interpretability of models is the basis for a variety of more specific tasks such as process analysis or redesign [8].

Regarding the covariate, we captured data on *prior method knowledge* in process modeling because it was previously shown to influence model understanding [21]. We measured prior method knowledge by using the set of process modeling method knowledge questions used by Mendling and Strembeck [37], which quizzes respondents' theoretical knowledge of process modeling. Their questions, notably, are grammar-independent and concern grammatical rules of process model routing logic, derived from fundamental work in this area [23], and address control flow criteria such as reachability, deadlocks, liveness and option to complete.

4.2. Procedures and materials

We used a paper questionnaire with five different sections. The Appendix includes examples of the materials.

The first section comprised questions about the participants' demographic data, academic qualifications and prior method knowledge. Participants were asked about the number of years they had worked in the IT-sector and the extent to which they had previously been involved with modeling in education and work. With these demographics, we can describe a sample frame similar to that in other studies on process model comprehension [35,53,59,60].

In the second section of the questionnaire we used the set of process modeling method knowledge questions used by Mendling and Strembeck [37] to measure prior method knowledge, which we used as a covariate in our data analysis. Additionally, we collected self-report data on the estimated amount of hours spent on learning process modeling.

The third section contained a tutorial on the process modeling grammar, in which the treatment was provided (as explained below). The tutorial was specifically tailored to inform participants about the meaning of each symbol in the provided grammar and covered everything the participants needed to know to perform the subsequent comprehension tasks.

The fourth section of the questionnaire displayed four different process models with eight corresponding comprehension tasks for each model (viz., 32 questions in total). Participants in each of the four main study groups (according to the between-subject factor routing symbol design) got all four models in the same routing symbol design (viz., using the same grammatical design of the process model). For each model similar comprehension questions were asked. Additionally, for each set of comprehension questions related to each model, participants also indicate the perceived cognitive load of answering the set of questions, on basis of the perceived cognitive load scale described below. To avoid any order effects e.g., due to fading attention, we used two different sampling strategies. Specifically the models and comprehension questions were arranged in different sequences. Participants were randomly assigned to one of the eight different questionnaires (four treatments of routing symbol design in two different sampling versions each).

The fifth and last section of the questionnaire included questionnaire scales in which participants could rate the routing symbols (AND and XOR) in the models shown in accordance to the four relevant design criteria discussed (viz., semantic transparency, perceptual discriminability, pop out and aesthetics). Additionally, this section included a scale on perceived control flow comprehension as an additional measure of the cognitive load associated with the model comprehension tasks.

Subjects were allowed to spend as much time as desired for the completion of the experimental tasks. On average, the experiment took about 40 min to complete.

4.3. Manipulation of treatment: construction of model sets with different routing symbols

Although process modeling grammars emphasize different viewpoints on processes [66], they share several common elements. Fig. 2 shows the routing symbols selected for the experimental study, which are inspired by existing process modeling grammars EPC (routing symbols R_{EPC}), UML Activity Diagrams (routing symbols R_{UML}) and BPMN (routing symbols R_{BPMN}), and YAWL (routing symbols R_{YAWL}). Relative size is held constant in comparison to further symbols, so that all routing symbols used are of comparable size. Additionally, the orientation of the symbols used is similar, as they are in right angle or directly aligned to the edge flow direction. Therefore, shape ($S_{Grammar}$) remains the main variable that varies amongst the symbols used across the grammars considered. S_{EPC} represents AND using a circle with a logical marker for 'and' in it ("^") and XOR by using a circle and an "X" marker. S_{UML} has different symbols for these concepts: AND is depicted as a filled bar, while XOR is represented by a diamond-shaped symbol. S_{YAWL} uses small rectangles with inscribed triangles. In the AND node, the triangle points inward, in the XOR node outward. S_{BPMN} employs diamond symbols for both node types, using a plus marker for the AND.

Fig. 2 summarizes the notational details of the four process models used in our study. We can see that the models are structurally equivalent, and only different in their use of different symbols for the routing behavior. All further model elements were held constant across the model sets.

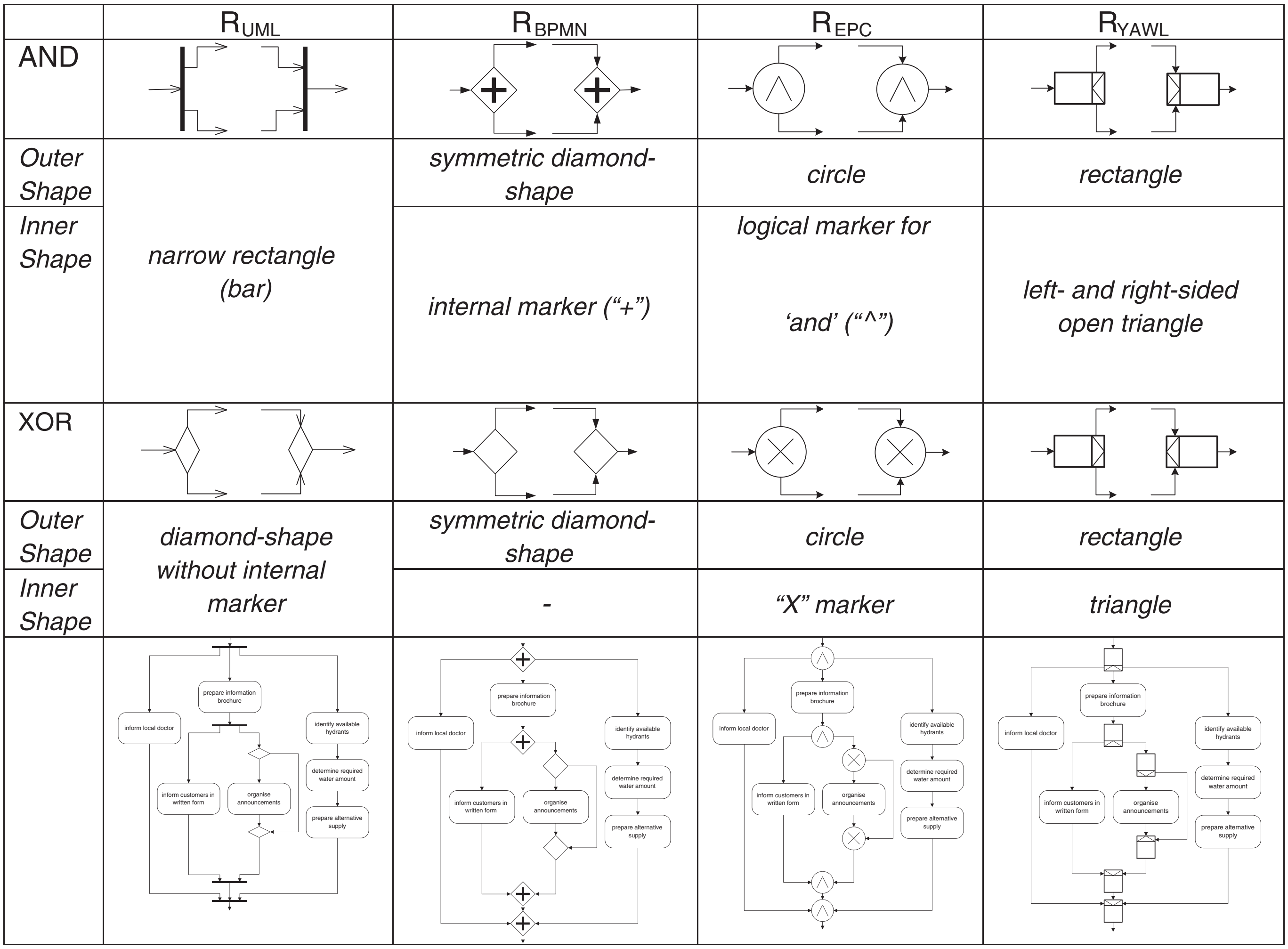
The models were developed as follows: First, we created Visio stencils to be able to model all necessary symbols in one tool providing high flexibility for layout. We directly redraw the exact routing symbol designs from the tool ARIS for EPC (routing symbols R_{EPC}), from examples given in the standard documents for UML Activity Diagrams (routing symbols R_{UML}) and BPMN (routing symbols R_{BPMN}), and from the original research paper on YAWL (routing symbols R_{YAWL}) [73], respectively. Next, the model design was optimized according to process modeling guidelines [35]. Finally, we exchanged the routing symbols in each of the models.

We used four different models so that the selection of the particular domain depicted would not influence results. The four models were selected from different domains such that we could expect that they are understandable for an average student with no special domain knowledge. Two of them stemmed from the business domain (product management and customer support, sales and distribution). The other two stemmed from uncommon domains: an emergency process plan for drinking water pollution and an e-mail election process (the last process was taken from the BPMN standard document [7]). Each of the four models used contained 21 activities. The amount of ANDs varied between 4 and 10, the amount of XORs between 4 and 11. The model size was held constant for all models, because prior research in the area of data modeling has shown that performance decreased in query composition tasks when larger models were used [6]. A sufficient level of complexity is required in empirical studies, because problems concerning the cognitive load may not be present for very small, manageable models, but only appear in more complex models. The models used are realistic examples of "normal" models, as models in practice contain about 19 tasks on average [48].

4.4. Measurement of dependent variables: comprehension accuracy, comprehension efficiency and perceived difficulty

For each model in the questionnaire we posed the same eight types of comprehension questions. The comprehension questions asked participants on four different issues concerning the control flow logic: concurrency, exclusiveness, order and repetition. Our questions were based on the measures developed and used in [35,58,59]. However, in comparison we formulated questions consistently, so that participants always had to consider two model elements (two activities) and their relationship for answering a question. This way, we ensured that the questions all specifically addressed the routing of process activities in a model. We worded each question using day-to-day-language.

We took care that the wording in the questions is understandable, and we ran a pre-test in order to make sure that the participants understood the questions. Questions were selected specifically to concern timely



	R_{UML}	R_{BPMN}	R_{EPC}	R_{YAWL}
AND				
Outer Shape	narrow rectangle (bar)	symmetric diamond-shape	circle	rectangle
Inner Shape		internal marker ("+")	logical marker for 'and' ("^")	left- and right-sided open triangle
XOR				
Outer Shape	diamond-shape without internal marker	symmetric diamond-shape	circle	rectangle
Inner Shape		-	"X" marker	triangle

Fig. 2. Routing symbols derived from existing process modeling grammars and details of a drinking water supply process model containing the different routing symbols.

and logical relationships between tasks in a process, such that participants had to rely on using the diagrams to understand these relationships. In the comprehension questions, participants had a choice of 'right', 'wrong' or 'I don't know' to reduce the probability of guessing.

Despite the use of the same wording, there is a large number of possibilities how to ask these questions, because any two activities can be targeted with the same question. We identified two basic variations: 1) the statement given in the question is correct or wrong and 2) the location of the chosen activities. For varying the location of activities consistently, we decided to use pairs of activities, which are either close (1 activity between them) or distant (>1 activity between them) according to the spatio-visual distance between them. As a consequence, we constructed the test material, varying correct and wrong answers as well as close and distant answers. This measure design allowed us to collect the total number of correct answers as a measure of comprehension accuracy.

To measure comprehension efficiency, we recorded the self-report completion time for the comprehension questions. We asked participants to write down the point of time at the beginning and the end of the questions, similar to [53].

To measure perceived difficulty of the comprehension task, we included the 7-point single-item perceived cognitive load measure (anchored between "very difficult" and "very easy") developed in [32]. This measure was accompanying each model comprehension task.

Additionally, to measure the perceived control flow comprehension difficulty specifically, we constructed a new scale with 4 items, which asked participants whether it had been easy to perceive loops (aspect repetition), concurrency, exclusiveness and sequence of activities (aspect order) in the four models.

4.5. Measurement of independent variable: notational evaluation of the routing symbol designs

To create measures for the independent variables, we constructed new four-item scales for each of the selected criteria: semantic transparency, visual discriminability, pop out and aesthetics. Self-construction of the scales was necessary, because there were no existing measures available for these constructs. More importantly, the nature of these constructs (dimension of the perceptual effectiveness of visual designs) demanded the construction of perceptional scales to evaluate individuals' beliefs about each of the dimensions.

Item construction for semantic transparency, pop out and visual discriminability was theoretically grounded in Moody's framework of desirable properties of effective visual notations [39] and followed established guidelines [54]. First, an item pool was generated with approximately 10 items per evaluated dimension. Then wordings were evaluated in a pre-test with 10 participants. The pre-test consisted of a card-sorting and a questionnaire in which each item candidate was rated according to its wording and its appropriateness to assess the respective dimension on a five-point scale. The content validity of items was checked in an online card sorting test. We used a closed card sorting test with 6 pre-testers, in which they had to arrange the items to given dimensions as well as an open card sorting test with 4 pre-testers in which they could arrange the items to self-named groups. For the final questionnaire the best four item candidates for each dimension were chosen to allow for sufficient reliability of scales. The visual discriminability scale was used to capture the perceptual discriminability of XOR and AND symbols. Additionally, both the XOR and the AND symbol sets were evaluated with the scales for semantic transparency, pop out and aesthetics, respectively. In turn, this approach allowed us to obtain measures for the visual design of the routing symbol sets in each of the models, as perceived by the participants working with the models.

4.6. Participants

Participants in the study were 154 information systems and business students from a European university. Table 1 shows selected demographic data about participants per cell. To account for expert–novice differences [49], we tried to find participants with both high and low experience in modeling and recruited them from different classes with and without prior training in modeling. We selected business school students as they are a realistic proxy of the future end-users of business process models. Table 1 summarizes key demographic variables. We performed analysis of variance tests to screen for possible differences between the experimental groups' demographics, which yielded no problematic differences.

5. Results

5.1. Validity and reliability assessment

We started by assessing validity and reliability of the Likert-type measures for the symbol evaluation. First, we conducted a principal components analysis with all symbol evaluation items as well as the items measuring perceived control flow comprehension difficulty. Five factors emerged with eigenvalues greater than 1, explaining 74.1% of the total variance. The five-factor solution was rotated to simple structure using Varimax. Table D.3 in the Appendix shows all factor loadings, cross-loadings, eigenvalues, and variance statistics. This first analysis demonstrated that the four items from the questionnaire scale control flow comprehension loaded on one factor as expected (factor 5). Additionally, all items evaluating the AND symbol loaded on one factor (factor 1). Factor 2 comprises eight items from the original subscales for evaluating ‘pop out’ of XOR as well as ‘visual discriminability’ of AND and XOR. This is surprising, as visual discriminability items were asked symmetrically for AND and XOR, but seemed to relate more to the design of the XOR than the AND symbol. The factors 3 and 4 resemble further XOR symbol evaluation scales.

As symbol evaluation items were used twice in the questionnaire (to evaluate XOR and AND separately), in a second step, we performed exploratory factor analyses, with extraction and Varimax rotation of solutions with between 2 and 3 factors for XOR and AND items separately. An iteration of the factor analysis was conducted to eliminate problematic measurement items. During this process, it became apparent that the item “The meaning of the XOR/AND-symbol is easy to recognize based on its visual design.” did not load on the expected factor (originally an item for semantic transparency). Therefore, we excluded this item from further analysis. The resulting PCA indicated that there were 2 factors in the dataset with eigenvalues greater than 1; however the gradient of the scree slope suggested that a solution with three factors would be tenable, too. We chose to use the three factor solution, so that the factors could reflect the original questionnaire scales. In comparison to the first factor, which accounted for 62% (XOR), respectively 65% (AND) of the variance, the proportion of variance explained by the further two factors was small (8% to 12%) before the varimax rotation.

To summarize this three-factor solution, it is apparent that the underlying structure displays a fairly unambiguous pattern of item loadings, in line with the postulated questionnaire scales. The factor loadings, cross-loadings, eigenvalues, and variance statistics are presented in the Appendix. Factors are conceptually clear, with 3–4 items loading at 0.7 or above at each factor (labeled “pop out”, “aesthetics” and “semantic transparency”) and exhibit only low cross-loadings. Factor structure and loadings are very similar for the items used to evaluate XOR as well as AND, demonstrating sufficient convergent and discriminant validity of our measurements.

To estimate reliability and internal consistency of our measures, we computed Cronbach's α, which should be greater than or equal to 0.7 to consider items to be uni-dimensional [42].

The reliability for the self-constructed scale “Subjective Difficulty of Control Flow Comprehension” including 4 items was good ($\alpha = 0.77$). Reliability of the newly developed scales for symbol evaluation were also satisfying ($\alpha = 0.87$–0.94). Additionally we calculated Cronbach's α for the subjective cognitive load items belonging to the problem-solving tasks of the four models ($\alpha = 0.95$). Cronbach's α for the general knowledge test on process modeling was 0.74. These results suggest adequate reliability. Deletion of any item produced no marked effect on the instrument's reliability score. In light of these results, we retained all items on these instruments.

Overall, the statistical results confirm that the developed measurement instruments are of appropriate validity and reliability.

6. Hypothesis testing

To examine our hypotheses, we ran four multiple regression analyses implemented in SPSS Version 19.0, one for each dependent variable (comprehension accuracy, comprehension efficiency, and task difficulty

Table 1
Participants demographic data.

	R_{UML} (n = 44; 28%)		R_{BPMN} (n = 48; 31%)		R_{YAWL} (n = 21, 14%)		R_{EPC} (n = 41, 27%)		Total (n = 154)	
	Mean/ amount	SD/ percentage	Mean/ amount	SD/ percentage	Mean/ amount	SD/ percentage	Mean/ amount	SD/ percentage	Mean/ amount	SD/ percentage
Age	23.40	2.62	23.75	3.80	23.00	3.27	24.32	3.16	23.70	3.26
Gender										
Male	27	61%	34	71%	17	81%	24	59%	102	66%
Female	17	39%	14	29%	4	19%	17	41%	52	34%
Highest grade completed										
High school	7	16%	13	27%	3	14%	3	7%	26	17%
Bachelor	34	77%	34	71%	17	81%	36	88%	121	78%
Master	3	7%	1	2%	1	5%	2	5%	7	5%
Participants with work experience in the IT-sector	15	34%	16	33%	3	14%	12	30%	46	30%
Participants with work experience with process models	4	9%	6	13%	2	10%	6	15%	18	12%
Participants with training on modeling basics	31	71%	33	69%	18	86%	36	88%	118	77%
Hours of training on modeling basics at university or school	27.00	36.37	28.00	33.51	31.24	45.76	24.34	23.93	27.12	33.72
Process modeling test score	65%	0.20	60%	0.20	68%	0.20	65%	0.22	64%	0.20

in terms of the two measures — perceived cognitive load and perceived control flow comprehension difficulty), respectively. For all four multiple regression analyses we used the same variables as independent factors, viz., the average total factor scores for perceived semantic transparency, visual discriminability, pop out and aesthetics, as well as the total process modeling method knowledge score.

One assumption behind the use of multiple regression analysis is that the variables are measured on a continuous scale and are normally distributed. Our data screening confirmed that the measures for the dependent variables completion time, perceived cognitive load and perceived control flow comprehension difficulty as well as for the independent variables semantic transparency, pop out and aesthetics met these criteria according to one-sample Kolmogorov–Smirnov tests. The dependent variable comprehension score (skewness = −0.88, kurtosis = 0.07), the process modeling knowledge test score (skewness = −0.30, kurtosis = −0.50) and perceptual discriminability (skewness = −0.90, kurtosis = −0.07) did not meet the criteria of normal distribution in this formal test. However, in large samples over 100 the impact of skewness and kurtosis on the results of the regression analysis diminishes and underestimation of variance disappears [70]. Therefore, we decided to use a linear regression model to analyze our data.

Our initial analyses using ordinary least square (OLS) regression models lead to unstable results of beta coefficients when adding or deleting one of the independent variables. For instance, if using perceived semantic transparency, visual discriminability, pop out and aesthetics as independent factors and comprehension accuracy as dependent variable, the model as a whole was significant and lead to a rejection of the joint hypothesis that these coefficients are zero ($F = 4.14$, $p = 0.02$), but on the other hand all regression coefficients were insignificant. Deleting one of the predictor variables changed the significance of predictors.

Such unstable results are typically an indicator for multicollinearity of predictor variables [16]. And indeed, pop out correlates strongly with the other symbol evaluations (0.61–0.74, $p = 0.00$), and bivariate correlations of around 0.70 can already inflate the size of error terms, weakening the analysis [73]. As determining the relevance of individual predictors is especially relevant for addressing our hypotheses, we thus decided to use ridge regression models to overcome shortcomings of the OLS regression. Ridge regression can handle sets of independent variables with multicollinearity better than OLS regression, because it generates estimators with smaller standard error than OLS regression [13]. Ridge regression was discussed controversially in the 1980s, but recent enhancements of the algorithms lead to robust solutions, stable coefficients and high quality of fit [68]. Table 2 reports the results of the ridge regression analyses showing the standardized beta coefficients and significance levels.

The overall regression model for the dependent variable comprehension accuracy was significant, $R^2 = 0.15$, $F(6, 39) = 3.53$, $p < 0.01$. The data in Table 2 further shows that pop out, perceptual discriminability and process modeling knowledge are significant predictors for process model comprehension, while semantic transparency and aesthetics are not. These results thus support H1a and H2a whereas hypotheses H3a and H4a must be refuted.

A regression of the subjects' responses on the comprehension efficiency measured in terms of time yielded a non-significant overall model. Therefore, hypotheses H1b–H4b were rejected.

In hypotheses H1c–H4c we expected that better symbol design in the dimensions perceptual discriminability, pop out, semantic transparency, and aesthetics will positively influence perceptions of task difficulty of process model comprehension. We used two different measures for task difficulty (perceived cognitive load and perceived control flow comprehension difficulty) which yielded different results. For perceived cognitive load, our predictions were borne out: the overall regression model and all regressing coefficients were significant leading to a 33% explanation rate. However, we found that only pop out, perceptual discriminability and process modeling knowledge were positively related to perceived control flow comprehension difficulty, while semantic transparency and aesthetics were not. These results provide partial support for H3c and H4c and full support for H1c and H2c.

In addition to the regression analyses reported, we ran further analyses to clarify whether individual perceptions of symbol designs or the treatment of actual different routing symbols account for differences in model comprehension. To that end, we ran the same analyses as reported, but included the variable routing symbol design with four levels (R_{EPC}, R_{BPMN}, R_{UML}, R_{YAWL}) as a further independent factor. As the ridge regression is part of the CATREG module in SPSS 19.0, it was possible to include a further variable on a nominal scale. Results demonstrated that the actual routing symbol design was not a significant predictor of process model comprehension. It thus seems that it is rather the individual perception of the symbol design than the property of the symbol itself, which determines the effect on comprehension.

In summary, we obtained strong support for our assertion that the perception of pop out and perceptual discriminability of symbols is positively associated with comprehension accuracy and task difficulty. Semantic transparency and aesthetics were related to perceived cognitive load. Thus, we are confident that our results support the fundamental proposition we have sought to test in our research. We obtained no support for our hypotheses regarding the estimation of process model comprehension efficiency (measured by task completion time). Table 3 summarizes our results.

7. Discussion

Our empirical study set out to test the influence of routing symbol design on process model comprehension in terms of accuracy, efficiency and perceived difficulty.

From a general perspective, the study reported in this paper extends research into the development of a validated notational design theory for modeling grammars. Most notably, we provide a first operationalization, measurement and test of the theory of visual notations by Moody [39].

The results obtained provide support for the nomological validity of Moody's design principles. We observed different levels of significance for our hypotheses on the four design attributes of routing symbols. In particular, pop out and perceptual discriminability showed a stronger effect on process model comprehension than semantic transparency and aesthetics. One possible explanation can be that pop out and perceptual discriminability of visual symbols are dimensions that directly relate to early stages of neural processing; differences in form of symbols can even be processed pre-attentively [79]. Therefore, they have a more direct effect on cognitive load involved in model comprehension tasks. It is not surprisingly that results are similar for both characteristics as they are closely related; symbols that are highly

Table 2
Ridge regression analysis: final model statistics.

Independent factors	Dependent variables			
	Comprehension accuracy	Time	Perceived cognitive load	Perceived control flow comprehension difficulty
	St. beta	St. beta	St. beta	St. beta
Process modeling knowledge	**0.14*****	−0.04	**0.14*****	**0.14*****
Perceptual discriminability	**0.09***	−0.04	**0.14*****	**0.10***
Pop out	**0.09****	−0.2	**0.13*****	**0.13*****
Semantic transparency	−0.02	−0.02	**0.11*****	0.06
Symbol aesthetics	0.04	0.05	**0.09***	0.04
F	**3.69****	0.44	**10.89*****	**6.06*****
R^2	0.15	0.03	0.33	0.22

* $p \leq 0.05$.
** $p \leq 0.01$.
*** $p \leq 0.001$.

Table 3
Summary of hypothesis testing results.

	Independent variable	Dependent variable	Results
H1a	Perceptual discriminability	Comprehension accuracy	Supported
H1b	Perceptual discriminability	Comprehension efficiency	Not supported
H1c	Perceptual discriminability	Task difficulty	Supported
H2a	Pop out	Comprehension accuracy	Supported
H2b	Pop out	Comprehension efficiency	Not supported
H2c	Pop out	Task difficulty	Supported
H3a	Semantic transparency	Comprehension accuracy	Not supported
H3b	Semantic transparency	Comprehension efficiency	Not supported
H3c	Semantic transparency	Task difficulty	Supported for perceived cognitive load, but not for perceived control flow comprehension difficulty
H4a	Aesthetics	Comprehension accuracy	Not supported
H4b	Aesthetics	Comprehension efficiency	Not supported
H4c	Aesthetics	Task difficulty	Supported for perceived cognitive load, but not for perceived control flow comprehension difficulty

discriminable will also be perceived easily and would be rated higher on the pop out scale. In contrast, semantic transparency and aesthetics relate to later stages of the perceptional processing. They reflect a subjective impression of the quality of the symbol design, which is not directly related to the perceptual effectiveness in the cognitive task. We originally postulated that aesthetic design might have an influence on model comprehension due to affective responses, and semantic transparency because of easy associations with existing knowledge structures which would enable cognitive offloading. Both modes of action are more likely to depend on individual graphic preferences and prior experiences stored in long-term-memory in comparison to the pre-attentative perception processes as reflected in pop out and perceptual discriminability.

Turning to the three different dimension of process model comprehension considered, we found that the effects of symbol design perceptions were not equally strong on the different dependent variables. Most notably, there was no effect on the comprehension efficiency. On the forefront, this result suggest that secondary notation effects (visual design choices) do not impact interpretational efficiency [8], viz., on the resource commitment required to gain a faithful understanding of a model. Another interpretation of that result is that participants could freely choose how much time they wanted to spend, and therefore time spent could also reflect their motivation to solve the comprehension tasks. In general, there is always a tradeoff between time spent and correctness of solution in cognitive tasks (referred to as speed-accuracy tradeoff; e.g. [4]). So, as speed was not a set target, participants could maximize accuracy if they wanted to.

Our results regarding comprehension efficiency are in line with other conceptual modeling experiments to a certain extent. For instance, Batra and Davis [2] found no time differences when investigating performance differences between novices and experts, although there was a significant difference in quality of the outcome. Also Reijers et al. [58] found that the treatment “color highlighting” had a significant effect on understanding accuracy, but not on understanding speed. In turn, we believe that our findings, coupled with the body of work to date, clearly point to a gap in understanding the reasons for differences in model comprehension efficiency, and what the consequences are for the effective use of conceptual models for systems analysis and design.

Regarding perceived cognitive difficulty, we note that the reported effects of our independent variables on perceived control flow comprehension difficulty were similar to those on comprehension accuracy but somewhat different from those on perceived cognitive load of the task as a whole. These findings suggest that the perceived control flow comprehension difficulty scale measures objective comprehension difficulty more closely, while perceived cognitive load might reflect a more subjective rating of the general task setting.

Our experimental setting also allowed us to clarify whether notational efficiency is an object property and therefore characteristic of any notation, or dependent on user evaluations. Our results indicate that individual ratings of symbol design were relevant for process model comprehension over and above the actual design of the symbols. Several theories can assist in providing an explanation for this result.

First, research on perception demonstrated in various ways how individual's preferences and motivational states can impact visual processing; so to say people “see what they want to see” [1]. Although most perceptional processes are bottom-up (the brain turns information from sense organs as eyes into a perception), still, top-down influences (driven by higher-level cognitive processes as the users attitude towards the symbols used) can have a great impact and can lead to perceptional biases (e.g. hills appears steeper than they actually are) or even phenomena such as unattentional blindness [65]. In such a case, a negative first impression of the symbols' design could lead to higher cognitive effort in the perception process needed for solving the comprehension tasks, leading both to lower performance and lower symbol evaluation. Additionally, differences in users' judgments of the symbols as well as in actual performances could reflect the fact that the efficiency of visual perception can vary significantly between individuals (see e.g. [77]).

In addition to interpretations on the perceptional point of view, our results can also be interpreted through the lens of motivational theories. Research has shown that individual attitudes as achievement motivation or subjective difficulty mediate the relationship between objective task difficulty and performance [9,20]. For instance, if individuals perceive increased task difficulty, they usually invest higher mental effort to protect performance as long as the tasks doesn't seem impossible to them. So, worse symbol design making the task more difficult per se, could not only lead to a performance loss, but also to an increase of mental effort in motivated participants resulting in similar performance of the experimental groups. However, those individuals who would have had the subjective impression that the symbol design was bad could have been demotivated and performed worse. The evidence in the present analysis is too weak for any real conclusions regarding this issue, however. Future research could include the measurement of intrinsic motivation to clarify the mechanism why individual perceptions of symbol design have a stronger influence on process model comprehension than the symbol design itself.

Finally, we note that our results confirm the relevance of secondary object attributes (perceptions of primary attributes) to understanding object-related behaviors [14]. Recker et al. [55] showed how perceptions of process modeling grammar primary notation characteristics determined behavioral evaluations of the process modeling grammar, such as perceived usefulness and perceived ease of use. While they found that both primary and secondary (perceptual) attributes of the primary notation of the grammar mattered to these evaluations, our results show that in terms of secondary notation, the secondary attributes are important over and above the primary attributes of visual design.

8. Implications and limitations

8.1. Implications for research

We identify several important theoretical and empirical findings in our research.

First, this study is the first to operationalize and measure four principles of effective visual notation design. Our factor analysis of the user evaluations of the symbols confirm that the criteria perceptual discriminability, semantic transparency, pop out and aesthetics are perceived as independent dimensions. Our work thus provides a measurement instrument that can be utilized in future studies on model comprehension; for instance, in studies that examine the interaction effects between ontological (primary notation) principles and visual (secondary notation) principles in conceptual models.

Future research could further extend our approach to investigate cognitive criteria in more detail and identify symbols that represent specific criteria very well and very poorly, and subsequently determine their influence on comprehension. However, our selection of symbols showed that it is very difficult to find symbols, which represent these criteria independently and it might not be possible to find symbols that, for instance, are semantically transparent but not aesthetic.

Our paper additionally encourages the exploration of the amount of variance attributed to individual vs. symbol design factors. Future research could tease out the relevance of personal vs. language factors by measuring further aspects of individual attitudes or perceptional abilities.

Our work uses cognitive load theory and the limitations of working memory to provide a theoretical explanation of the cognitive effectiveness of different routing symbols. The results provide evidence that inefficient design of symbols may place extra extraneous cognitive load on end user. Our results therefore add strength to a growing body of empirical work that applies cognitive load theory to the context of understanding visual models. While subjective measurement of cognitive load has been established as an efficient and reliable instrument [46], it is desirable to adopt techniques developed in cognitive psychology to the study of conceptual model comprehension. First, there is the option to measure cognitive load based on secondary task performance, in which a distracting task is imposed over the original task [46]. Second, brain image processing and other techniques can be applied to directly measure cognitive load on the neurophysiological level. It would be valuable to study these instruments for model comprehension tasks, which might reveal more detailed insights into cognitive load effects in this domain.

8.2. Implications for practice

We believe our findings inform specifically modeling grammar and guideline development. Standards and descriptions for modeling grammars often do not give a reason why specific symbols are chosen. Our study shows the relevance of visual symbol design and aims to motivate future modeling grammar developments to include user evaluation procedures for symbol choice as well. Other areas dependent on the use of symbols for instance conduct detailed user evaluations of different symbol variations as size, orientation, thickness of lines (see for example [64] for a study on prohibitive symbols). Such user evaluations bear the potential to reveal symbol design issues. The design of single symbols is important, but also their combination in a symbol set, so that visual discriminability of symbols can be warranted. Our study shows that the selection of shapes and symbols will affect whether models created with a modeling grammar will be easy and accurately understandable. To achieve, for instance, semantically transparent symbol design, it can be reasonable to use well known symbols known from other domains (e.g., mathematics) to allow for positive transfer effects.

The findings from our study also suggest implications for the choice of a particular modeling notation. For instance, our experiment shows that the YAWL notation suffers from weak discriminability of the XOR and AND routing symbols. The assessments of the other symbols might be used as a source of inspiration for notational rework. Beyond this observation, we do not want to make direct statements about the underlying process modeling grammars used for our experimental routing symbols for two reasons. First, we sometimes had to choose one option when a modeling grammar offered several ones (e.g., in BPMN there are various options for XOR routing symbols). Second, we only considered the routing elements in our experiment. Dimensions such as visual discriminability though have to be considered relative to the whole set of symbols offered by a notation. As we focused on a subset of the symbols of each grammar, our study would not be able to reflect upon weaknesses outside this subset or between the considered symbols and the ignored ones.

8.3. Limitations

We now discuss potential threats to validity and how we addressed them in the experiment. These limitations constrain the interpretation of our results to the context in which we gathered the data.

The participants of our study were students who were familiar with process modeling in general, although they were not experts in this area. The results might differ if the experiment is replicated with experts in business process modeling or with experts with a stronger software engineering background. However, as the participants had received training in information technology, their level of modeling experience was likely to be equivalent to those of typical business users of process models in many organizations. A recent study even found graduate students to perform better in a process model comprehension experiment [59]. Also, our study concerned basic cognitive principles of visual design and comprehension, and would have been confounded by using participants with high level of domain knowledge as these participants would have relied on background domain knowledge to answer the comprehension tasks; which would have obscured the effects of the visual designs [47]. Still, future research could investigate the different cognitive demands and preferences of experts and novices for symbol choice. Here, it may well be the case that certain designs are better for untrained people than trained experts who know what to look for.

In designing our treatments, we had to trade off internal and external validity in light of ecological validity considerations. We chose for a treatment with symbols from real modeling languages and with models with realistic textual descriptions to warrant ecological validity. While the results are in line with our predictions and showed in the correct direction, the use of realistic models over models with artificially inflated visual differences probably contributed to the rather low level of explained variance we observed in our results. This is because, for instance, users could have been able to partially derive information of symbols from the context or the semantic content of a process model. For instance, for the construction of a loop an XOR symbol is needed but not an AND symbol.

Finally, we discuss the experimental task order as a potential source of bias. In our study, participants had to rate the control flow symbols after completing the model comprehension tasks. To determine if they might have given different answers when asked about the visual design of the symbols without performing the comprehension tasks, we gathered additional data points on symbol rating through a follow-up study with a student population comparable with the initial population (viz., students in the same courses in the following year). In this follow-up study, students only rated the symbols without having to perform a comprehension test. A comparison of the symbol evaluations with and without comprehension tasks is displayed visually in Appendix E. A correlation analysis revealed high similarity of the variances observed in the ratings across the two groups ($r = 0.94$, $p = 0.000$), a further MANCOVA analysis showed no interaction effect between symbol ratings and point of time of rating. These results indicate that our study results about the influence of symbol ratings on the variance in model comprehension remains valid, albeit we note a slight bias exhibited by the experiment task order, in that ratings were slightly lowered when performed after comprehension tasks. Still, we do not see evidence that this bias could have influenced the regression results in any way, as this bias was consistent over all evaluations.

9. Conclusion

Our study provides empirical evidence of the importance of symbol design on (process) model comprehension. We found that notational characteristics such as perceptual discriminability and pop out are significantly associated with perceived cognitive load and model comprehension accuracy but not comprehension efficiency.

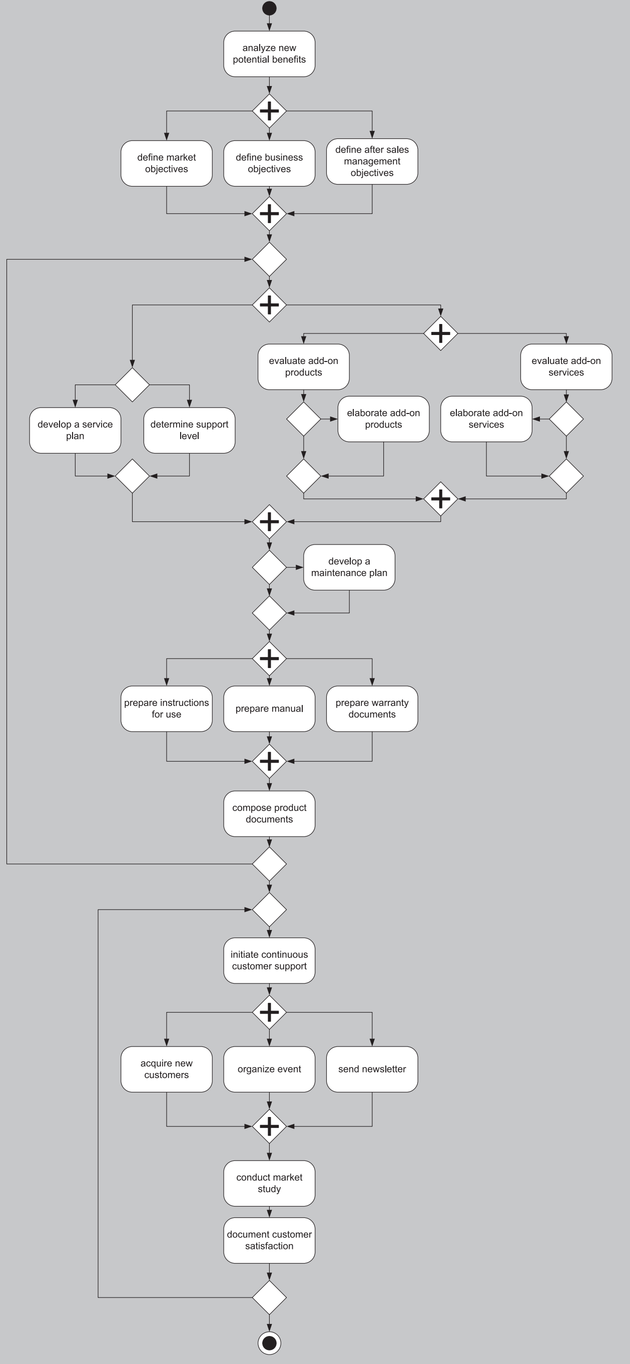
In a broader sense, the results provide evidence for the utility of the theory of effective visual design of notations to the study of process modeling in practice and the management of process modeling initiatives in practice. In turn, our research adds to the growing body of experimental research on conceptual modeling practices, and adds to the inventory of relevant theories, complementing relevant, established principles based on ontological considerations [47], multimedia learning considerations [18] and secondary notation considerations [58]. The cumulative tradition of research of these studies, in turn, advances our understanding of the issues and challenges in an important field of information systems practice, the effective use of modeling notations for the analysis and design of organizational and technological systems.

Acknowledgements

Dr. Recker's contributions to this research have been supported by a grant from the Australian Research Council (ARC DE120100776) and by a Fellowship from the Alexander-von-Humboldt Foundation.

Appendix A. Example of model with comprehension questions

Table A.1
Example of model with comprehension questions.

	• **Concurrency 1:** “Prepare instructions for use” and “prepare manual” can be executed at the same point of time. (correct, close)
	• **Concurrency 2:** “Evaluate add-on products” and “send newsletter” can be executed in parallel. (wrong, distant)
	• **Exclusiveness 1:** In one process instance “develop a service plan” as well as “determine support level” can be executed. (correct, close)
	• **Exclusiveness 2:** The process steps “define after sales management objectives” and “organize event” are mutually exclusive. (wrong, distant)
	• **Sequence 1:** If “conduct market study” as well as “elaborate add-on services” are executed in a process instance, then “conduct market study” has to be finalized before “elaborate add-on services” can start. (wrong, distant)
	• **Sequence 2:** If “define market objectives” as well as “analyze new potential benefits” are executed in a process instance, then “define market objectives” is executed before “analyze new potential benefits”. (wrong, close)
	• **Repetition 1:** “Document customer satisfaction” can be executed more often than “define after sales management objectives”. (correct, distant)
	• **Repetition 2:** In each process instance “acquire new customers” is executed exactly as often as “initiate continuous customer support”. (correct, close)

Appendix B. Subjective difficulty of control flow comprehension

(5 point scale from "totally disagree" to "totally agree"). [$\alpha = 0.765$]

It was easy to perceive in the models, which process steps …
- …are executed in a loop.
- …are executed in parallel (AND).
- … are mutually exclusive (XOR).
- … are executed in sequence.

Appendix C. User evaluation of symbols

In the following, we ask you to rate the visual appearance of the AND as well as XOR symbols: [Cronbach's Alphas in brackets for XOR and AND items]

- Perceptual Popout [$\alpha_{XOR} = 0.93/\alpha_{AND} = 0.93$]
  - XOR/AND-symbols can be found quickly in a model.
  - XOR/AND-symbols are easy to find in a model.
  - XOR/AND-symbols can be recognized immediately in a model.
  - It is easy to recognize XOR/AND-symbols in a model.

- Semantic transparency [$\alpha_{XOR} = 0.87/\alpha_{AND} = 0.89$]
  - XOR/AND-symbols are intuitively understandable in a model.
  - Even without explanation it is clear what a XOR/AND-symbol represents.
  - One doesn't have to learn the meaning of the XOR/AND-symbol, to understand it.

- Aesthetics [$\alpha_{XOR} = 0.93/\alpha_{AND} = 0.91$]
  - The XOR/AND-symbol is optically pleasing.
  - The XOR/AND-symbol is visually appealing.
  - The XOR/AND-symbol is visually esthetic.
  - The XOR/AND-symbol is well-designed.

- Visual discriminability [$\alpha = 0.94$]

AND and XOR symbols are …..
  - …difficult to distinguish in a model.
  - …well to distinguish in a model.
  - …easy to confuse in a model.
  - …easy to discriminate in a model.

Appendix D. Factor analysis for symbol evaluation

Table D.1
Factor loadings for XOR symbol evaluation items (PCA with varimax rotation).

Symbol evaluation items (XOR)	Factor 1 — perceptual pop out	Factor 2 — aesthetics	Factor 3 — semantic transparency
The XOR-symbol is visually aesthetic. (XOR, aesthetics)	0.17	**0.86**	0.27
The XOR-symbol is optically pleasing. (XOR, aesthetics)	0.40	**0.77**	0.24
The XOR-symbol is well-designed. (XOR, aesthetics)	0.43	**0.74**	0.25
The XOR-symbol is visually appealing. (XOR, aesthetics)	0.36	**0.80**	0.33
XOR-symbols can be found quickly in a model. (XOR, pop out)	**0.83**	0.24	0.28
XOR-symbols are easy to find in a model. (XOR, pop out)	**0.86**	0.27	0.22
It is easy to recognize XOR-symbols in a model. (XOR, pop out)	**0.81**	0.33	0.16
XOR-symbols can be recognized immediately in a model. (XOR, pop out)	**0.86**	0.30	0.18
Even without explanation it is clear what a XOR-symbol represents. (XOR, semantic transparency)	0.28	0.31	**0.79**
XOR-symbols are intuitively understandable in a model. (XOR, semantic transparency)	0.27	0.24	**0.80**
One doesn't have to learn the meaning of the XOR-symbol, to understand it. (XOR, semantic transparency)	0.12	0.23	**0.89**
Eigenvalue	3.47	3.05	2.53
Percentage of variance	31.55	27.68	23.03

Total variance accounted for = 82%.
Item loadings at or above 0.50 are shown in bold for clarity.

Table D.2
Factor loadings for AND symbol evaluation items (PCA with varimax rotation).

Symbol evaluation items (AND)	Factor 1 — perceptual pop out	Factor 2 — aesthetics	Factor 3 — semantic transparency
The AND-symbol is visually aesthetic. (AND, aesthetics)	0.25	**0.83**	0.23
The AND-symbol is optically pleasing. (AND, aesthetics)	0.43	**0.76**	0.27
The AND-symbol is well-designed. (AND, aesthetics)	0.41	**0.68**	0.29
The AND-symbol is visually appealing. (AND, aesthetics)	0.34	**0.82**	0.27
AND-symbols can be found quickly in a model. (AND, pop out)	**0.83**	0.35	0.23
AND-symbols are easy to find in a model. (AND, pop out)	**0.82**	0.33	0.27
It is easy to recognize AND-symbols in a model. (AND, pop out)	**0.78**	0.31	0.28
AND-symbols can be recognized immediately in a model. (AND, pop out)	**0.81**	0.32	0.29
Even without explanation it is clear what a AND-symbol represents. (AND, semantic transparency)	0.31	0.35	**0.77**
AND-symbols are intuitively understandable in a model. (AND, semantic transparency)	0.31	0.37	**0.77**
One doesn't have to learn the meaning of the AND-symbol, to understand it. (AND, semantic transparency)	0.22	0.15	**0.89**
Eigenvalue	3.38	3.10	2.54
Percentage of variance	30.74	28.21	23.05

Total variance accounted for = 82%.
Item loadings at or above 0.50 are shown in bold.

Table D.3
Factor loadings for all symbol evaluation items (PCA with varimax rotation).

	Factor 1 — AND-symbol	Factor 2 — XOR perceptual discriminability and pop out	Factor 3 — XOR aesthetics	Factor 4 — XOR-semantic transparency	Factor 5 — subjective difficulty of control flow comprehension
It was easy to perceive in the models, which process steps are executed in a loop. (subjective difficulty of control flow comprehension)	0.21	0.16	0.14	0.10	**0.75**
It was easy to perceive in the models, which process steps are executed in parallel (AND). (subjective difficulty of control flow comprehension)	0.25	0.18	0.02	0.17	**0.78**
It was easy to perceive in the models, which process steps are mutually exclusive (XOR). (subjective difficulty of control flow comprehension)	0.10	0.15	0.24	0.17	**0.75**
It was easy to perceive in the models, which process steps are executed in sequence. (subjective difficulty of control flow comprehension)	0.09	0.14	−0.03	−0.04	**0.74**
AND and XOR symbols are difficult to distinguish in a model (recoded). (perceptual discriminability)	0.17	**0.81**	0.09	−0.00	0.16
AND and XOR symbols are well to distinguish in a model. (perceptual discriminability)	0.24	**0.82**	0.08	0.02	0.22
AND and XOR symbols are easy to confuse in a model. (perceptual discriminability)	0.19	**0.81**	0.10	−0.01	0.18
AND and XOR symbols are easy to discriminate in a model. (perceptual discriminability)	0.19	**0.84**	0.08	0.06	0.16
The XOR-symbol is visually aesthetic. (XOR, aesthetics)	0.22	0.12	**0.82**	0.23	0.05
The XOR-symbol is optically pleasing. (XOR, aesthetics)	0.07	0.35	**0.77**	0.27	0.12
The XOR-symbol is well-designed. (XOR, aesthetics)	0.07	0.35	**0.73**	0.31	0.09
The XOR-symbol is visually appealing. (XOR, aesthetics)	0.19	0.31	**0.76**	0.33	0.14
XOR-symbols can be found quickly in a model. (XOR, pop out)	0.14	**0.69**	0.30	0.40	0.05
XOR-symbols are easy to find in a model. (XOR, pop out)	0.17	**0.71**	0.35	0.31	0.11
It is easy to recognize XOR-symbols in a model. (XOR, pop out)	0.16	**0.71**	0.38	0.21	0.13
XOR-symbols can be recognized immediately in a model. (XOR, pop out)	0.20	**0.69**	0.37	0.30	−0.02
Even without explanation it is clear what a XOR-symbol represents. (XOR, semantic transparency)	0.18	0.23	0.34	**0.75**	−0.01
XOR-symbols are intuitively understandable in a model. (XOR, semantic transparency)	0.19	0.16	0.28	**0.72**	0.28
The meaning of the XOR-symbol is easy to recognize based on its visual design. (XOR, semantic transparency)	0.17	0.46	0.44	0.43	0.11
One doesn't have to learn the meaning of the XOR-symbol, to understand it. (XOR, semantic transparency)	0.20	0.05	0.28	**0.78**	0.14
The AND-symbol is visually aesthetic. (AND, aesthetics)	**0.65**	0.08	0.55	0.02	0.04
The AND-symbol is optically pleasing. (AND, aesthetics)	**0.73**	0.14	0.49	−0.01	0.16
The AND-symbol is well-designed. (AND, aesthetics)	**0.69**	0.16	0.40	0.04	0.20
The AND-symbol is visually appealing. (AND, aesthetics)	**0.71**	0.12	0.53	0.03	0.10
AND-symbols can be found quickly in a model. (AND, pop out)	**0.68**	0.52	0.19	0.02	0.14
AND-symbols are easy to find in a model. (AND, pop out)	**0.73**	0.46	0.16	0.01	0.11
It is easy to recognize AND-symbols in a model. (AND, pop out)	**0.68**	0.44	0.12	0.05	0.19
AND-symbols can be recognized immediately in a model. (AND, pop out)	**0.71**	0.46	0.14	0.02	0.18
Even without explanation it is clear what a AND-symbol represents. (AND, semantic transparency)	**0.81**	0.07	0.00	0.29	0.04
AND-symbols are intuitively understandable in a model. (AND, semantic transparency)	**0.78**	0.05	0.03	0.36	0.19
The meaning of the AND-symbol is easy to recognize based on its visual design. (AND, semantic transparency)	**0.76**	0.28	0.08	0.11	0.11

Table D.3 (*continued*)

	Factor 1 — AND-symbol	Factor 2 — XOR perceptual discriminability and pop out	Factor 3 — XOR aesthetics	Factor 4 — XOR-semantic transparency	Factor 5 — subjective difficulty of control flow comprehension
One doesn't have to learn the meaning of the AND-symbol, to understand it. (AND, semantic transparency)	**0.74**	0.03	−0.18	0.36	0.09
Eigenvalue	6.90	6.41	4.54	3.03	2.83
Percentage of variance	21.57	20.03	14.17	9.48	8.85

Total variance accounted for = 74.10%.
Item loadings at or above 0.50 are shown in bold.

Appendix E. Symbol evaluation

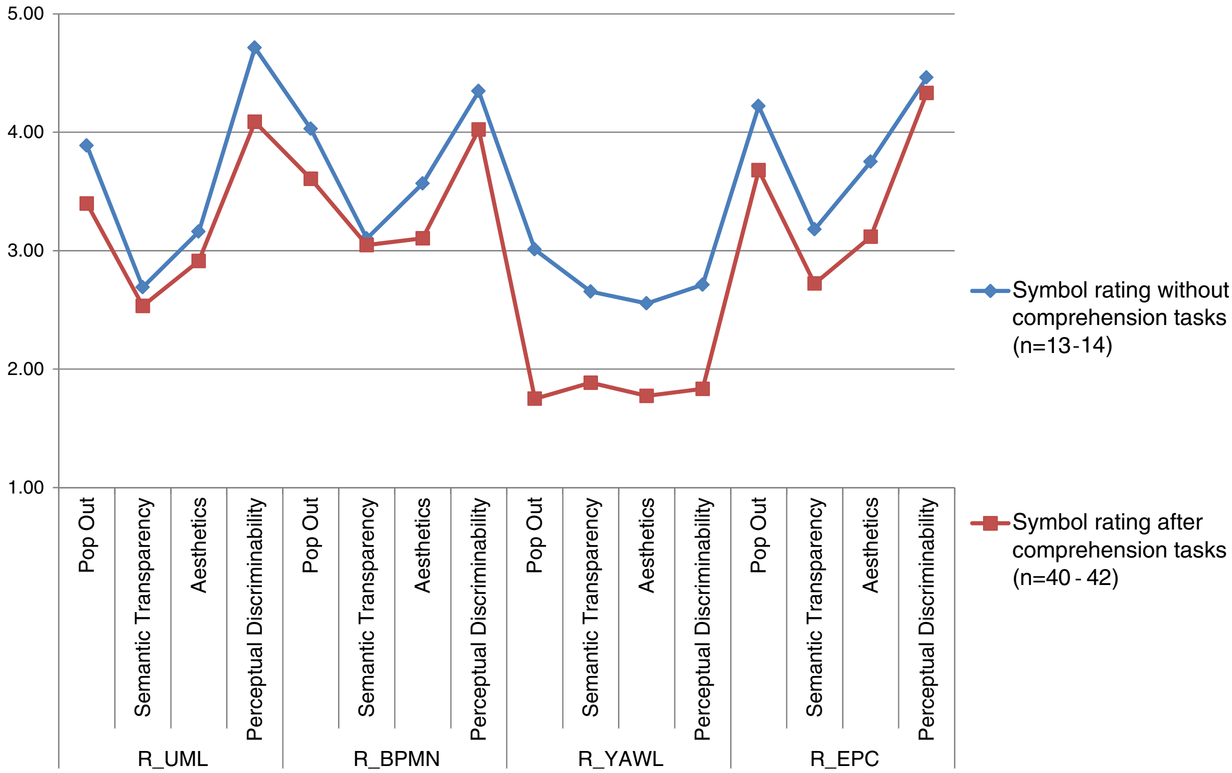


Fig. 3. Results of symbol rating in experimental and post-hoc group.

References

[1] E. Balcetis, D. Dunning, See what you want to see: motivational influences on visual perception, Journal of Personality and Social Psychology 91 (4) (2006) 612–625.

[2] D. Batra, J.G. Davis, Conceptual data modelling in database design: similarities and differences between expert and novice designers, International Journal of Man–machine Studies 37 (1) (1992) 83–101.

[3] P.H. Bloch, Seeking the ideal form: product design and consumer response, The Journal of Marketing 59 (3) (1995) 16–29.

[4] R. Bogacz, E.-J. Wagenmakers, B.U. Forstmann, S. Nieuwenhuis, The neural basis of the speed–accuracy tradeoff, Trends in Neurosciences 33 (1) (2010) 10–16.

[5] P.C. Bottger, M.A. Woods, Different determinants of task persistence and growth satisfaction: affective responses to performance, planning and job characteristics, Australian Journal of Management 13 (2) (1988) 303–317.

[6] P.L. Bowen, R.A.O. Farrell, F.H. Rohde, An empirical investigation of end-user query development: the effects of improved model expressiveness vs . complexity, Information Systems Research 20 (4) (2009) 565–584.

[7] BPMI.org, OMG, Business Process Modeling Notation Specification. Final Adopted Specification, in, Object Management Group, 2006.

[8] A. Burton-Jones, Y. Wand, R. Weber, Guidelines for empirical evaluations of conceptual modeling grammars, Journal of the Association for Information Systems 10 (6) (2009) 495–532.

[9] R.L. Capa, M. Audiffren, S. Ragot, The interactive effect of achievement motivation and task difficulty on mental effort, International Journal of Psychophysiology 70 (2) (2008) 144–150.

[10] T.D. Cook, D.T. Campbell, Quasi-Experimentation: Design and Analysis Issues, Houghton Mifflin, Boston, Massachusetts, 1979.

[11] K. Corral, D. Schuff, R.D. St, The impact of alternative diagrams on the accuracy of recall: a comparison of star-schema diagrams and entity-relationship diagrams, Decision Support Systems 42 (2006) 450–468.

[12] N. Cowan, The magical mystery four: how is working memory capacity limited, and why? Current Directions in Psychological Science 19 (1) (2010) 51–57.

[13] W.H. Crown, Statistical Models for the Social and Behavioral Sciences: Multiple Regression and Limited-Dependent Variable Models, Praeger, Westport, 1998.

[14] G.W. Downs Jr., L.B. Mohr, Conceptual issues in the study of innovation, Administrative Science Quarterly 21 (4) (1976) 700–714.

[15] P. Effinger, M. Siebenhaller, M. Kaufmann, An Interactive Layout Tool for BPMN, in: IEEE Conference on Commerce and Enterprise Computing, IEEE, Vienna, Austria, 2009, pp. 399–406.

[16] D. Farrar, R. Glauber, Multicollinearity in regression analysis: the problem revisited, The Review of Economics and Statistics 49 (1) (1967) 92–107.

[17] A. Gemino, Y. Wand, A framework for empirical evaluation of conceptual modeling techniques, Requirements Engineering 9 (4) (2004) 248–260.

[18] A. Gemino, Y. Wand, Complexity and clarity in conceptual modeling: comparison of mandatory and optional properties, Data & Knowledge Engineering 55 (3) (2005) 301–326.

[19] M. Genero, G. Poels, M. Piattini, Defining and validating metrics for assessing the understandability of entity-relationship diagrams, Data & Knowledge Engineering 64 (3) (2008) 534–557.

[20] M.S. Humphreys, W. Revelle, Personality, motivation, and performance: a theory of the relationship between individual differences and information processing, Psychological Review 91 (2) (1984) 153–184.

[21] V. Khatri, I. Vessey, V. Ramesh, P. Clay, P. Sung-Jin, Understanding conceptual schemas: exploring the role of application and IS domain knowledge, Information Systems Research 17 (1) (2006) 81–99.
[22] S. Khemlani, P.N. Johnson-Laird, Disjunctive illusory inferences and how to eliminate them, Memory & Cognition 37 (5) (2009) 615–623.
[23] B. Kiepuszewski, A.H.M. ter Hofstede, W.M.P. van der Aalst, Fundamentals of control flow in workflows, Acta Informatica 39 (3) (2003) 143–209.
[24] E. Kindler, On the semantics of EPCs: resolving the vicious circle, Data & Knowledge Engineering 56 (1) (2005) 23–40.
[25] P.A. Kirschner, Cognitive load theory: implications of cognitive load theory on the design of learning, Learning and Instruction 12 (1) (2002) 1–10.
[26] N. Kock, J. Verville, A. Danesh-Pajou, D. DeLuca, Communication flow orientation in business process modeling and its effect on redesign success: results from a field study, Decision Support Systems 46 (2) (2009) 562–575.
[27] K. Kotovsky, J.R. Hayes, H.A. Simon, Why are some problems hard? Evidence from Tower of Hanoi, Cognitive Psychology 17 (2) (1985) 248–294.
[28] J.H. Larkin, H.A. Simon, Why a diagram is (sometimes) worth ten thousand words, Cognitive Science 11 (1) (1987) 65–100.
[29] T. Lavie, N. Tractinsky, Assessing dimensions of perceived visual aesthetics of web sites $, Information Systems 60 (2004) 269–298.
[30] D. Leutner, C. Leopold, E. Sumfleth, Cognitive load and science text comprehension: effects of drawing and mentally imagining text content, Computers in Human Behavior 25 (2) (2009) 284–289.
[31] A. Maes, G. Poels, Evaluating quality of conceptual modelling scripts based on user perceptions, Data & Knowledge Engineering 63 (3) (2007) 701–724.
[32] N. Marcus, M. Cooper, J. Sweller, Understanding instructions, Journal of Educational Psychology 88 (1) (1996) 49–63.
[33] S.J.P. McDougall, M.B. Curry, O.D. Bruijn, Measuring symbol and icon characteristics: norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols, Behavior Research Methods, Instruments, & Computers 31 (3) (1999) 487–519.
[34] J. Mendling, J. Recker, H.A. Reijers, On the usage of labels and icons in business process modeling, International Journal of Information System Modeling and Design 1 (2) (2010) 40–58.
[35] J. Mendling, H. Reijers, W.M.P. van der Aalst, Seven process modeling guidelines (7PMG), Information and Software Technology 52 (2) (2010) 127–136.
[36] J. Mendling, H.A. Reijers, J. Recker, Activity labeling in process modeling: empirical insights and recommendations, Information Systems 35 (4) (2010) 467–482.
[37] J. Mendling, M. Strembeck, J. Recker, Factors of process model comprehension — findings from a series of experiments, Decision Support Systems 53 (1) (2012) 195–206.
[38] G.A. Miller, The magical number seven, plus or minus two: some limits on our capacity for processing information, Psychological Review 63 (1956) 81–97.
[39] D.L. Moody, The "physics" of notations: toward a scientific basis for constructing visual notations in software engineering, IEEE Transactions on Software Engineering 35 (6) (2009) 756–779.
[40] D.C.L. Ngo, Measuring the aesthetic elements of screen designs, Displays 22 (3) (2001) 73–78.
[41] J.C. Nordbotten, M.E. Crosby, The effect of graphic style on data model interpretation, Information Systems Journal 9 (2) (1999) 139–155.
[42] J.C. Nunnally, I.H. Bernstein, Psychometric Theory, 3rd ed. McGraw-Hill, New York, New York, 1994.
[43] Object Management Group, BPMN 2.0 by Example, in, 2010.
[44] OMG, Business Process Modeling Notation, V1.2, in, Object Management Group, 2009.
[45] C. Ouyang, W.M.P. van der Aalst, M. Dumas, A.H.M. ter Hofstede, J. Mendling, From business process models to process-oriented software systems, ACM Transactions on Software Engineering Methodology 19 (1) (2009) 2–37.
[46] F. Paas, J.E. Tuovinen, H. Tabbers, V. Gerven, P.W. M, Cognitive load measurement as a means to advance cognitive load theory, Educational Psychologist 38 (2003) 63–72.
[47] J. Parsons, An experimental study of the effects of representing property precedence on the comprehension of conceptual schemas, Journal of the Association for Information Systems 12 (6) (2011) 401–422.
[48] S. Patig, V. Casanova-Brito, B. Vögeli, IT Requirements of Business Process Management in Practice — An Empirical Study, in: R. Hull, J. Mendling, S. Tai (Eds.), Business Process Management, Springer, Berlin/Heidelberg, 2010, pp. 13–28.
[49] M. Petre, Why looking isn't always seeing: readership skills and graphical programming, Communications of the ACM 38 (6) (1995) 33–44.
[50] M. Petre, Cognitive dimensions 'beyond the notation', Journal of Visual Languages & Computing 17 (4) (2006) 292–301.
[51] H.C. Purchase, D. Carrington, J.-A. Allder, Empirical evaluation of aesthetics-based graph layout, Empirical Software Engineering 7 (3) (2002) 233–255.
[52] P.T. Quinlan, Visual feature integration theory: past, present and future, Psychological Bulletin 129 (5) (2003) 643–673.
[53] J. Recker, A. Dreiling, The effects of content presentation format and user characteristics on novice developers' understanding of process models, Communications of the Association for Information Systems 28 (6) (2011) 65–84.
[54] J. Recker, M. Rosemann, A measurement instrument for process modeling research: development, test and procedural model, Scandinavian Journal of Information Systems 22 (2) (2010) 3–30.
[55] J. Recker, M. Rosemann, P. Green, M. Indulska, Do ontological deficiencies in modeling grammars matter? Management Information Systems Quarterly 35 (1) (2011).
[56] J. Recker, M. Rosemann, M. Indulska, P. Green, Business process modeling: a comparative analysis, Journal of the Association for Information Systems 10 (4) (2009) 333–363.
[57] J. Recker, N. Safrudin, M. Rosemann, How novices design business processes, Information Systems 37 (6) (2012) 557–573.
[58] H.A. Reijers, T. Freytag, J. Mendling, A. Eckleder, Syntax highlighting in business process models, Decision Support Systems 51 (2011) 339–349.
[59] H.A. Reijers, J. Mendling, A study into the factors that influence the understandability of business process models, IEEE Transactions on Systems, Man, and Cybernetics — Part A 41 (2011) 449–462.
[60] H.A. Reijers, J. Mendling, R.M. Dijkman, Human and automatic modularizations of process models to enhance their comprehension, Information Systems 36 (5) (2011) 881–897.
[61] J. Sarkkinen, H. Karsten, Verbal and visual representations in task redesign: how different viewpoints enter into information systems design discussions, Information Systems Journal 15 (3) (2005) 181–211.
[62] A.-W. Scheer, ARIS — Business Process Modeling, 3rd ed. Springer, Berlin, Germany, 2000.
[63] B.N. Schenkman, F. Jonsson, Aesthetics and preferences of web pages, Behaviour Information Technology 19 (5) (2000) 367–377.
[64] K.-K. Shieh, S.-M. Huang, Factors affecting preference ratings of prohibitive symbols, Applied Ergonomics 34 (6) (2003) 581–587.
[65] D. Simons, C. Chabris, Gorillas in our midst: sustained inattentional blindness for dynamic events, Perception 28 (9) (1999) 1059–1074.
[66] P. Soffer, Y. Wand, Goal-driven multi-process analysis, Journal of the Association for Information Systems 8 (3) (2007).
[67] A. Sonderegger, J. Sauer, The influence of design aesthetics in usability testing: effects on user performance and perceived usability, Applied Ergonomics 41 (3) (2010) 403–410.
[68] L. Stan, Enhanced ridge regressions, Mathematical and Computer Modelling 51 (5–6) (2010) 338–348.
[69] J. Sweller, Cognitive load during problem solving: effects on learning, Cognitive Science: A Multidisciplinary Journal 12 (2) (1988) 257–285.
[70] B.G. Tabachnick, L.S. Fidell, Using Multivariate Statistics, Pearson Education, Inc., Boston, 2007.
[71] N. Tractinsky, A.S. Katz, D. Ikar, What is beautiful is usable, Interacting With Computers 13 (2) (2000) 127–145.
[72] A. Treisman, G. Gelade, A feature-integration theory of attention, Cognitive Psychology 12 (1) (1980) 97–136.
[73] W.M.P. van der Aalst, A.H.M. ter Hofstede, YAWL: Yet Another Workflow Language, Information Systems 30 (4) (2005).
[74] H.M.V. Verbeek, W.M.P. van der Aalst, A.H.M. ter Hofstede, Verifying workflows with cancellation regions and OR-joins: an approach based on relaxed soundness and invariants, The Computer Journal 50 (3) (2007) 294–314.
[75] R.W. Veryzer, Aesthetic response and the influence of design principles on product preferences, Advances in Consumer Research Volume 20 (1993) 224–228.
[76] I. Vessey, Cognitive fit: a theory-based analysis of the graphs versus tables literature, Decision Sciences 22 (2) (1991) 219–240.
[77] E.K. Vogel, M.G. Machizawa, Neural activity predicts individual differences in visual working memory capacity, Nature 428 (6984) (2004) 748–751.
[78] Y. Wand, R. Weber, Research commentary: information systems and conceptual modeling — a research agenda, Information Systems Research 13 (4) (2002) 363–376.
[79] C. Ware, Information Visualization, 2 ed. Elsevier. Morgan Kaufmann, San Francisco, 2004.
[80] W. Winn, Encoding and retrieval of information in maps and diagrams, IEEE Transactions on Professional Communication 33 (3) (1990) 103–107.

Dr. Kathrin Figl is an Assistant Professor in the Institute for Information Systems and New Media at the Vienna University of Economics (WU). She received her Doctoral (awarded with the Dr. Maria Schaumayer Award) and two Master's in Information Systems and Psychology, both with honours, from the University of Vienna. Most of her applied research and teaching focuses on the intersection between information systems and psychology, including research on information systems education, human-computer-interaction and cognitive aspects of modeling. In 2010 she was awarded the excellent teaching award from the Vienna University of Economics for her lecture on information systems. She has authored more than 40 papers, including four best paper awards, in peer-reviewed journals and conference proceedings.

Dr. Jan Recker is Alexander-von-Humboldt Fellow, Professor for Information Systems and Woolworths Chair of Retail Innovation at Queensland University of Technology, Brisbane, Australia. His research focuses on usage of process design in organizational practice as well as IT-enabled business transformations and business innovations. He has written over 100 books, journal articles and conference proceedings, including publications in *MIS Quarterly*, *Journal of the Association for Information Systems*, *Information Systems*, *European Journal of Information Systems*, *Information & Management*, *Scandinavian Journal of Information Systems*, *Decision Support Systems* and others. He is Senior Editor for the Journal of IT Theory and Application, Associate Editor for Communications of the AIS, a member of the editorial board of several international journals and serves on the program committee of various conferences.

Jan Mendling is a Full Professor with the Institute for Information Business at Wirtschaftsuniversität Wien (WU Vienna), Austria. His research areas include Business Process Management, Conceptual Modelling and Enterprise Systems. He has published more than 100 research papers and articles, among others in *ACM Transactions on Software Engineering and Methodology*, *Information Systems*, *Data & Knowledge Engineering*, *Decision Support Systems*, *Formal Aspects of Computing*, and *IEEE Transactions on Software Engineering*. He is member of the editorial board of three international journals. His Ph.D. thesis has won the Heinz-Zemanek-Award of the Austrian Computer Society and the German Targion-Award for dissertations in the area of strategic information management. He is one of the founders of the Berlin BPM Community of Practice (http://www.bpmb.de) and organizer of several academic events on process management. He was program co-chair of the International Conference on Business Process Management 2010.