

# Towards Guiding Process Modelers Depending upon Their Expertise Levels

Jonas Bulegon Gassen<sup>1</sup>(✉), Jan Mendling<sup>1</sup>, Lucineia Heloisa Thom<sup>2</sup>,  
and José Palazzo M. de Oliveira<sup>2</sup>

<sup>1</sup> Wirtschaftsuniversität Wien, Vienna, Austria  
{jonas.gassen,jan.mendling}@wu.ac.at

<sup>2</sup> Universidade Federal do Rio Grande do Sul, UFRGS - PPGC, Porto Alegre, Brazil  
{lucineia,palazzo}@inf.ufrgs.br

**Abstract.** Business process modeling is an important task for supporting business process management. One challenge of process modeling is the diversity of expertise from novice modelers to expert designers, which defines the need to provide different sorts of guidance during modeling. In this paper, we discuss these expertise differences based on the cognitive load theory and present an experiment that tested how people with different levels of expertise work with different instructional material presented as automatic feedback of a tool. Our conclusions suggest that guidance on reworking models needs to take different levels of expertise into account.

**Keywords:** Business process · Process modeling · Cognitive load theory · Process of process modeling

## 1 Introduction

The quality of business process modeling is of crucial importance for business process management. Various works discuss techniques and concepts for improving the quality of process models covering modeling guidelines [1, 2], labeling guidelines [3, 4], and pitfalls of process modeling [5]. Also the importance of the way of how a model is created, which is called the process of process modeling (PPM), is discussed in this context [6, 7]. However, many challenges are still open [8], a.o. to take the diverse levels of modeling expertise into account.

It is generally agreed that suitable feedback and guidance during PPM (e.g. by means of a tool) can help to raise the quality of process models, in particular for *novices*. These are modelers with little experience that do not yet fully understand how correct and meaningful models can be created. The problem in this context is that *experts* will likely be annoyed by seeing basic hints, which distract them from the actual modeling task. Indeed, it has been found in experiments that what helps novices to increase performance can actually decrease performance of experts [9].

The objective of this paper is to investigate this trade-off in more detail. We formulate the idea that guidance has to be adapted to the level of expertise in order to be fully effective. To substantiate the benefits of such an strategy, an experimental research approach is required. We describe the cognitive foundations of such an experiment in terms of cognitive load theory [10]. We complement our argument with initial insights from a pre-study.

Against this background, this paper is structured as follows. Section 2 presents the background of effective guidance in terms of cognitive research. Section 3 describes our overall research design and the way how the experiment was conducted. Section 4 presents the study results along with a discussion of the findings. Section 5 concludes the paper.

## 2 Background

The central difference between novices and experts is that the latter will find tasks easy that the former find difficult. Cognitive load theory [10] helps to discuss this matter in more detail by distinguishing intrinsic and extraneous cognitive load. In essence, the intrinsic source of effort is related to the information itself and the extraneous load is related to the presentation of the information. Both can vary due to many factors: such as knowledge of the modeler, instructional procedures, or complexity of information.

Both types of cognitive load are related to differences between novices and experts while solving problems, such as the use of means-ends analysis versus domain-specific knowledge, respectively. Means-ends analysis, which is natural for humans, requires considering the current state and the goal state. While moving from the current state in the direction of the goal-state, humans verify differences between them and try to apply problem-solving strategies to change the state. Another difference according to cognitive load theory is that novices need more effort, since they use thinking, different from experts, which use knowledge. Instructional data may help novices. This bigger effort spent by a novice is clarified by recent research with functional magnetic resonance imaging, which demonstrates that the cerebral adaptation modifies the brain of an established professional such that energy consumption is reduced [11]. Confirmatory experiments have been conducted with different conceptual tasks such as translation or interpretation [12].

These findings have implications for guidance in process modeling, specifically the so-called worked example effect, the expertise reversal effect, or the guidance fading effect. These effects imply that guidance for novices should present more detailed instructional information and fill the knowledge gaps on novice's knowledge. Designers with higher level of expertise should work with problem-solving based on prior knowledge. Thus, the instructional information should follow the level of expertise of the designer; otherwise it may compromise the PPM.

These observations are complementary to recent research on process modeling that identify user characteristics as factors of performance. Recker and Dreiling [13] analyzed the comprehension of models and find that prior knowledge of a modeling grammar and of business process management showed to

be significant factors. Process modeling of novices is investigated in [14], suggesting that a hybrid representation of text plus abstract graphical shapes lead to better semantic quality. Theoretical arguments in this direction had been provided before by Moody [15]. The observation that users approach a modeling task differently is described by Pinggera et al. [16]. The cognitive phases of comprehension, modeling and reconciliation are identified in the record of their modeling experiment. Similar observations are made by Soffer et al. [17].

In summary, the need for differentiating guidance for novices and experts can be deduced from the literature. In the following, we describe a research design to clarify the benefits of such an approach.

### 3 Research Design

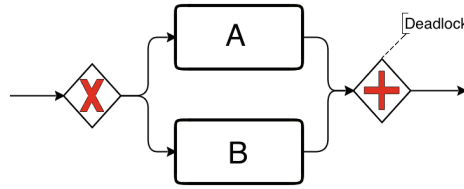
This section presents our idea of guidance concerning different expertise levels, our experiment and its threats to validity.

#### 3.1 Guidance

We divide guidance in two types: (i) to help the modeler immediately during the process modeling and (ii) after the modeling task is completed in order to enhance learning, for instance regarding the complexity and the compliance with formal guidelines [18] and number of errors per model. For this paper we focus on (i).

Considering a situation of immediate feedback in terms of instructional material such as diagrams with detailed textual explanations. According to Sweller, regarding the expertise reversal effect, [10], each of these materials would be suitable for novices, as they are new knowledge for them. However, the same instruction presented to experts may require additional cognitive resources to process redundant material, compared to the same material without detailed text. An example in the context of PPM occurs after the detection of a deadlock. If the designer has a low expertise level, a worked example is shown, describing why the deadlock may occur. A worked example exposes a step-by-step solution, which provides the person with the problem-solving schemas that need to be stored in long-term memory [10]. This may fill the gap of knowledge of the user, because he may not know why deadlocks occur in process models. If the user has a higher level of expertise, the elements causing the error may be marked, allowing the designer to fix it. However, since the user knows why the deadlock happens, the detailed explanation and even the worked example could create an extraneous cognitive load and should not be shown.

Figure 1 presents an example for that case. A novice could be shown a red marking in case of a deadlock. The image shows a simplified case, because the situation where it occurred in the user's modeling environment could be more complicated, causing extraneous cognitive load. Besides the figure, a text explaining why it occurs would be shown, such as: "The XOR Split, by its semantic, will execute only ONE outgoing flow. The AND Join will wait for ALL incoming flows before continuing the execution. Thus, the presented use will never



**Fig. 1.** Worked example for a deadlock

complete, generating a deadlock”. This would help the novice designer to learn during the modeling. If the user has a mid-low expertise level, the same image with the short text “deadlock” could be shown. For users with a higher expertise level the markers could be done in the current model and a short text “deadlock” could be shown. For experts, only the markers in the current model could be more suitable. Empirical tests should be done in order to understand the threshold and level of instructional material to use for each case. Besides guidance during process modeling, based on data regarding modeler’s actions during PPM, a personalized plan of study could be drawn for each modeler. Thus, the modelers could follow the plan to learn about their mistakes outside of the modeling activity.

In psychology the act of following common norms for a group, such as: staying in line or stopping before a red signal is called conformity. Asch [19] and other works present experiments that show the conformity behavior in humans and discuss how this is important for life in society. We assume the process modeling community to be a society. Thus, an operationalization for conformity could be done by means of guidelines and best practices used by modelers. Data in this context is used to measure conformity of the user within the society.

### 3.2 Experiment Description

In this work we tested, through an online experiment, how designers react to some instructional material. The experiment was performed as within-subjects, all participants received the same treatment. However, there are possible differentiation amongst the participants regarding expertise level on process modeling. The purpose of the experiment was to analyze the influence of different levels of instructional material during the PPM. First of all, we asked the designers to answer demographics (based on [6]), in order to try to assess their knowledge on business process modeling and Business Process Modeling and Notation (BPMN). The demographics are presented later on. Secondly, we inserted errors intentionally in three process models - mortgage application, company resupplying and travel booking (based on models from an online collection<sup>1</sup>). The designers had to describe which changes they would suggest to each model to improve or fix them, nothing could be added to or removed from the models, only changed. The problems applied were: for Model A, a deadlock, for Model B

<sup>1</sup> <http://bpmmai.org>.

a livelock and for Model C, verb-object style misuse. Assuming that a tool would show instructional information about the detected problems, we organized the same procedure. Models A and C have instructions, in two different levels. They were presented in the following order:

1. **Model A1:** no instruction.
2. **Model B1:** no instruction.
3. **Model A2:** worked example (Fig. 1).
4. **Model B2:** no instruction.
5. **Model A3:** worked example plus textual description on the problem.
6. **Model B3:** no instruction.
7. **Model C1:** no instruction.
8. **Model C2:** red marks in the model to spot the problem, only.
9. **Model C3:** red marks plus examples and textual description on the problem.

Each of these steps was composed by a set of elements, some of them allowing the participant to enter information (inputs). They are presented on Table 1.

**Table 1.** Elements of each step of the experiment, after demographics

Element	Type	Occurence	Input
Model name plus description	Text	All	No
Image of the process model	BPMN model	All	No
Instruction level 1	Example	A2	No
Instruction level 2	Example plus text	A3, C3	No
Marks in the model	Red marks	C2, C3	No
Question for changes in the model	Free text	All	Yes
Why the change should be made	Free text	All	Yes
Mental effort for answering	Likert 1–5	After each	Yes
Confidence on the answer	Likert 1–5	After each	Yes
Wheter the participant knew VO	Yes-No	After C3	Yes

The subjective labels of the Likert scales varied from very low to very high mental effort/confidence on checking the model quality, which were transformed to 1 to 5, as presented on Table 1. To support this approach we consider the work of Paas [20] that pointed out self-evaluation as valuable asset for mental effort evaluation. The questions answered as free text require human evaluation, which will be explained further on in this paper. Showing the same model could cause a control problem. One could look different for the model because it is showed repeatedly. We defined three variants (one per step) of models A and B. The structure was kept the same, but we applied changes in domain, layout and used synonyms where terms repeated amongst models. In this way, the models look different but are actually the same regarding flow and mistakes presented. We did not create variants of Model C because we spot the problem in the second

and third times that it was shown. Model B does not have an instruction in order to test if the instruction of Model A will help a similar problem.

In the steps where instruction was provided, the participant had to at least open the instruction, only then the questions (the same for other steps) were made available. Thus, he keeps the instruction open as much as was needed. Based on answers of two questions for the same model we can infer the following about guidance:

- **Wrong and Wrong:** necessary and not sufficient;
- **Wrong and Correct:** necessary and sufficient;
- **Correct and Correct:** not necessary and sufficient.

The reason why experts may find necessary to have marks showing the problems could be explained by the Anchoring & Adjustment phenomenon [21]. Besides the questions presented above, we also stored the overall time to perform the experiment, time that each instruction remained open and amount of times that each instruction was opened. We conjectured that experts will keep the instructions open by much less time than novices.

### 3.3 Threats to Validity

Regarding the experiment, a possible drawback is the evaluation of experts vs. novices. In addition, some measurements might not be possible, such as additional effort of experts caused by detailed instructional material. Our current sample do not allow to perform statistical significance tests, a bigger and more balanced sample is required. Lastly, the evaluation of answers is performed by humans and thus it has to be done carefully to avoid causing bias.

## 4 Experiment Execution

This section presents the execution of the experiment altogether with analysis and results.

### 4.1 Data Cleansing

A total of 33 persons participated of the experiment, from which 2 participants have answered in Portuguese and were removed even before the evaluation of answers. A translation could lead to bias and then was avoided. Another criteria applied was to remove participants that applied less than 10 min in the experiment, four participants were removed. Lastly, participants who stated having more than 6 days of self-education or formal training were removed, as they could have misunderstood the question, this happened to 5 participants. The questionnaire presented that one semester of classes would result in 3 days of training, as the questions referred to self-education or training within the last 12 months, is unlikely to have more than 6 days within 12 months. The possibility of having two courses at the same time was not considered for our sample. Therefore, 22 participants out of 33 stand after data cleansing.

## 4.2 Demographics

The participants are from academy (undergraduate, master and Ph.D. students) and were chosen intentionally and not statistically at random. We sent invitation for a known sample, some of them answered and some did not. In order to have information about the participants we asked demographic questions based on Pinggera (2012), which resulted 10 variables. They are presented, altogether with a normalization, in Table 2:

**Table 2.** Descriptive for demographics variables and normalization

Variable	Mean	SD	Scale	Normalization	New scale
Years of modeling	1.05	2.17	0–inf	value * 2	0–inf.
Models analyzed	3.36	10.68	0–inf	0;25;50;75;100, +1 per level	0–5
Models Created	1.59	4.28	0–inf	0;25;50;75;100, +1 per level	0–5
Activities average	4.50	7.36	0–inf	0;5;10;15;20, +0.3 per level	0–1.5
Formal training	1.14	1.86	0–inf	Value / 2	0–3
Self-education	0.59	0.80	0–inf	Value / 2	0–3
Self-evaluation	1.86	1.39	1–7	IF (Value>2; Value-2; 0)	0–5
BPMN familiarity	1.95	1.50	1–7	IF (Value>2; Value-2; 0)	0–5
BPMN competence	1.86	1.39	1–7	IF (Value>2; Value-2; 0)	0–5
BPMN months	3.05	8.96	0–inf	Value /12 * 2	0–inf.

The proper categorization of expert and novice users requires a deeper evaluation and we do not discuss this specific point here. In order to analyze our data we performed a categorization based on years of modeling, transforming the other variables to a value related to a year of modeling. For the variables normalized by means of levels (i.e. models analyzed), for each level a certain value was add, for example, reading 40 models gives the participant 2 points. For variables arising from Likert scales (i.e. self-evaluation), points were given only to participants that have value greater than 2. The first two options of the scale were “strongly disagree” and “disagree” on questions such as “Are you an experienced business process modeler?”. In this sense, selecting any of these questions gives 0 points for the participant. After the normalization, we sum each variable and participants with a score higher than 10 were considered experts, the remainder novices. Based on that, 5 participants were selected as experts and 17 as novices. This sample has an imbalance, which difficult the realization of statistics and should be solved by having more participants.

To achieve 10 points the user needs at least almost one point in each variable. It is not possible to achieve 10 points with only one variable, besides years of modeling or BPMN months, which would give a certain degree of expertise in case they are high. Also, it is very unlikely to have some variable and to not have another, as some of them are connected.

### 4.3 Answers Evaluation

The experiment required free text answers in some points and an evaluation of those answers is necessary. We performed a peer review evaluation. Each reviewer gave a score to each of the questions of all participants. The reviewer had three options: wrong, somewhat correct and correct. After the subjective answers, they were transformed to 0, 1 and 2, respectively. The mean was performed for the scores and was used as score for each participants answer.

### 4.4 Results

The experiment allows us to compare many different aspects. First, we compared the means of scores for each questions. A t-test would be suitable to test the significance of the differences between the means of each group, however due to the imbalance of our sample, novices group with 17 participants and experts group with 5 participants, we did not perform significance tests. Table 3 show the means for each question, comparing novices and experts.

**Table 3.** Means of scores based on answers evaluation (0–2)

	A1	A2	A3	B1	B2	B3	C1	C2	C3
Novices	0.00	0.29	1.21	0.12	0.47	0.56	0.00	0.00	0.59
Experts	0.80	0.80	0.80	1.20	0.80	0.80	0.20	0.60	0.80

It is possible to see in Table 3 that only in question 3 of model A the group of novices achieved better scores than experts. Novices having higher scores than experts on A3 may be due to issues such as incorrect categorization of groups (novice vs. expert), novices applying more attention and effort than experts or that some experts did not detected the problems due to anchoring and adjustment phenomena. Also, it is possible to note that aside of Models A and B for experts, all groups increased scores, as the questionnaire evolves. In addition, another interesting point here is that model B did not present an instruction, it was expected that the instructions of model A, being similar, would help the participants in questions of model B. They were showed in a paired fashion with questions from model A, as presented on Sect. 3.2. Table 3 shows the means increasing as instructions of model A were shown, for novices group, this may suggest retention and transfer [22]. Table 4 show the means for model C using as group the participants that did not know and that knew the guideline.

**Table 4.** Scores separated by knowledge about verb-object style (0–2)

	C1	C2	C3
Did not know	0.06	0.06	0.59
Knew	0.00	0.40	0.80



For all cases besides the first question, which did not have any help to spot the problem, the group that knew the guideline achieved better scores. The second question had marks in the activities containing the problematic labels, but only the last question (instruction 3) presented knowledge about the verb-object style. It is possible to note how complete knowledge elevates score for novices. Suggesting that proper material is fundamental as feedback during modeling. Besides the scores, we also analyzed time, Table 5 exposes the means of time and count of instruction access for our experiment.

**Table 5.** Time and access means

	Novices	Experts
Total time for experiment	1343.6	977.4
Time for Instruction 1	48.7	9.4
Time for Instruction 2	70.6	43.6
Time for Instruction 3	78.1	50.6
Access count for Instruction 1	3.1	1.2
Access count for Instruction 2	2.4	1.4
Access count for Instruction 3	1.65	1.4

The novices group needed a greater amount of time to complete the whole experiment. We used simple measurements of time, thus it is not possible to have a great precision in this regard. Some participants did not go through all the experiment at once. Thus, participants that used more than 50 min to perform the experiment were set as missing values (2 participants). Also, participants that used more than 5 min in any of the instructions were set as missing values (1 participant). Novices had greater means for all instructions, both for time spent and counting times opening each instruction, as conjectured.

As presented on Sect. 3.2, based on wrong and correct answers we infer whether instructions were necessary and sufficient for participants. We performed a percentage for each conclusion. It is important to note that we used only full scores. For example, necessary and not sufficient (N&NS) means  $Q(N)$  and  $Q(N+1)$  as wrong or score 0 for both. Necessary and sufficient (N&S) means wrong (score = 0) and correct (score = 2). Finally, unnecessary and sufficient (UN&S) means both correct (score = 2). Figure 2 shows the percentages for each inference to all instructions, separated by novices and experts.

Figure 2 allows to verify that instructional material is necessary and should be well studied as an improvement for process modeling tools. All instructions on their first apparition (A1, B1 and C1) were at least 40% necessary and not sufficient for experts and at least 71% necessary and not sufficient for novices. Second apparitions had smaller percentages, thus the previous material or the second material completeness might have helped.

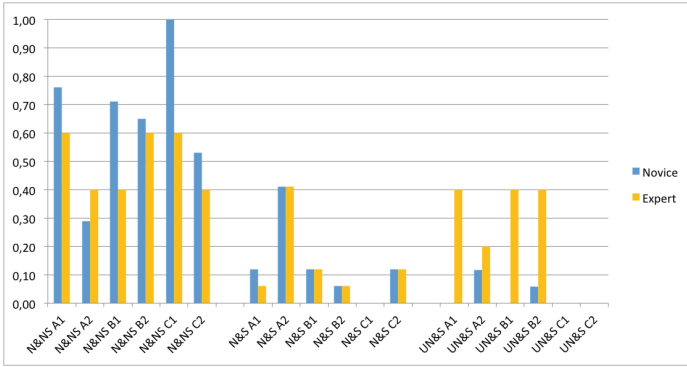


Fig. 2. Necessary and sufficient conditions

### 4.5 Discussion

Due the imbalance of the current sample, we are not able to perform statistical tests regarding significance, in this sense we cannot accept or reject our hypotheses. However, the results allows us to have an idea on differences between experts and novices and their interaction with instructional material regarding process modeling. The classification of one as expert or novice is an important point. Referring to instructional material presented by tools, in order to improve process modeling, the use of the term experienced instead of expert might provide a better semantics. Which means that the person in question has a wider range of knowledge about process modeling. Yet, to provide better instruction to the user, it would be interesting to gather data about a variety of different aspects concerning the process of process modeling. It is possible to see the difference on results between questions concerning a problem occurring in process models, such as deadlock, and a guideline, such as verb-object style. Thus conformity should also be considered. An expert modeler might not know specific guidelines, what does not make of him a novice. Hence a fine-grained approach, might be more suitable than only classifying users as experts vs. novices. Web services could be made available to tools that are willing to present this personalization. They would have to call the web services in modeling time, based on modelers' actions, and send the related data. Then the web services could provide suitable feedback accordingly. In addition, improvement plans could be generated to the modeler, describing aspects that should be studied by each modeler.

## 5 Conclusions

We presented discussions about expertise and conformity level of business models designers and reactions based on that. Complementary empirical research has to be done in order to test which specific aspects should be considered to suggest the expertise level of users. Also, the instructional material for different levels of

expertise has to be calibrated. However, defending the expertise level of modelers is a very complex task. In this sense, gathering a wide range of data in a fine-grained fashion might be a better solution. The feedback can be focused in each fine-grained aspect. For example, an expert modeler might not know about the verb-object style and it does not make of him a novice.

Other experiments in many directions should be performed. Companies may have their own guidelines and best practices, which would create a local society. The conformity is measured based on the global society, which refers to the process modeling community in general. Thus it is possible to match the designer with the local and global societies. Also, the conformity of the local society (company) can be matched against the global society (BPM community).

**Acknowledgments.** This work has been partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal (CAPES), Brazil.

## References

1. Becker, J., Rosemann, M., von Uthmann, C.: Guidelines of business process modeling. In: van der Aalst, W., Desel, J., Oberweis, A. (eds.) *Business Process Management*. LNCS, vol. 1806, pp. 30–49. Springer, Heidelberg (2000)
2. Mendling, J., Reijers, H.A., Cardoso, J.: What makes process models understandable? In: Alonso, G., Dadam, P., Rosemann, M. (eds.) *BPM 2007*. LNCS, vol. 4714, pp. 48–63. Springer, Heidelberg (2007)
3. Mendling, J., Reijers, H.A., Recker, J.: Activity labeling in process modeling: empirical insights and recommendations. *Inf. Syst.* **35**(4), 467–482 (2010)
4. Leopold, H., Smirnov, S., Mendling, J.: On the refactoring of activity labels in business process models. *Inf. Syst.* **37**(5), 443–459 (2012). Best papers from DOLAP 2010
5. Rosemann, M.: Potential pitfalls of process modeling: part a. *Bus. Proc. Manage. J.* **12**(2), 249–254 (2006)
6. Pinggera, J., Zugal, S., Weber, B.: Investigating the process of process modeling with cheetah experimental platform. *EMISA Forum* **30**(2), 25–31 (2010)
7. Claes, J., Vanderfeesten, I., Reijers, H.A., Pinggera, J., Weidlich, M., Zugal, S., Fahland, D., Weber, B., Mendling, J., Poels, G.: Tying process model quality to the modeling process: the impact of structuring, movement, and speed. In: Barros, A., Gal, A., Kindler, E. (eds.) *BPM 2012*. LNCS, vol. 7481, pp. 33–48. Springer, Heidelberg (2012)
8. Indulska, M., Recker, J., Rosemann, M., Green, P.: Business process modeling: current issues and future challenges. In: van Eck, P., Gordijn, J., Wieringa, R. (eds.) *CAiSE 2009*. LNCS, vol. 5565, pp. 501–514. Springer, Heidelberg (2009)
9. Schewpe, J., Rummer, R.: Attention, working memory, and long-term memory in multimedia learning: an integrated perspective based on process models of working memory. *Educ. Psychol. Rev.* **26**(2), 285–306 (2014)
10. Sweller, J., Ayres, P., Kalyuga, S.: *Cognitive Load Theory. Explorations in the Learning Sciences, Instructional Systems and Performance Technologies*. Springer, New York (2011)

11. Nigmatullina, Y., Hellyer, P.J., Nachev, P., Sharp, D.J., Seemungal, B.M.: The Neuroanatomical Correlates of Training-Related Perceptuo-Reflex Uncoupling in Dancers. *Cerebral Cortex*, New York (2013)
12. Moser-Mercer, B.: The search for neuro-physiological correlates of expertise in interpreting. *Translation and Cognition*, pp. 263–287 (2010)
13. Recker, J.C., Dreiling, A.: The effects of content presentation format and user characteristics on novice developers' understanding of process models. *Commun. Assoc. Inf. Syst.* **28**(6), 65–84 (2011)
14. Recker, J., Safrudin, N., Rosemann, M.: How novices design business processes. *Inf. Syst.* **37**(6), 557–573 (2012)
15. Moody, D.: The physics of notations: toward a scientific basis for constructing visual notations in software engineering. *IEEE Trans. Softw. Eng.* **35**(6), 756–779 (2009)
16. Pinggera, J., Zugal, S., Weber, B., Fahland, D., Weidlich, M., Mendling, J., Reijers, H.A.: How the structuring of domain knowledge helps casual process modelers. In: Parsons, J., Saeki, M., Shoval, P., Woo, C., Wand, Y. (eds.) *ER 2010. LNCS*, pp. 445–451. Springer, Heidelberg (2010)
17. Soffer, P., Kaner, M., Wand, Y.: Towards understanding the process of process modeling: theoretical and empirical considerations. In: Daniel, F., Barkaoui, K., Dustdar, S. (eds.) *BPM Workshops 2011, Part I. LNBIP*, vol. 99, pp. 357–369. Springer, Heidelberg (2012)
18. Mendling, J., Reijers, H.A., van der Aalst, W.M.P.: Seven process modeling guidelines (7pmsg). *Inf. Softw. Technol.* **52**(2), 127–136 (2010)
19. Asch, S.: Opinions and social pressure. *Sci. Am.* **193**(5), 31–35 (1955)
20. Paas, F.: Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *J. Educ. Psychol.* **84**, 429–434 (1992)
21. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. *Science* **185**(4157), 1124–1131 (1974)
22. Mayer, R.: *Multimedia Learning*. Cambridge University Press, Cambridge (2001)