Visualization of the Evolution of Layout Metrics for Business Process Models

Cornelia Haisjackl[1], Andrea Burattin[1], Pnina Soffer[2], and Barbara Weber[1,3]

[1] University of Innsbruck, Austria
[2] University of Haifa, Israel
[3] Technical University of Denmark

Abstract. Considerable progress regarding impact factors of process model understandability has been achieved. For example, it has been shown that layout features of process models have an effect on model understandability. Even so, it appears that our knowledge about the modeler's behavior regarding the layout of a model is very limited. In particular, research focuses on the end product or the outcome of the process modeling act rather than the act itself. This paper extends existing research by opening this black box and introducing an enhanced technique enabling the visual analysis of the modeler's behavior towards layout. We demonstrate examples showing that our approach provides valuable insights to better understand and support the creation of process models. Additionally, we sketch challenges impeding this support for future research.

Key words: Process of Process Modeling, Human–Centered Support, Process Model Quality, Understandability, Layout Properties, Visualization

1 Introduction

In recent years, business process models have gained significant relevance due to their critical role for the management of business processes [1]. Advantages of using business process models are, for example, obtaining a common understanding of a company's business processes [2], or enabling the discovery of improvement opportunities [3]. Literature emphasizes that a good understanding of a process model has positive influence on the success of a modeling initiative [4]. Research on impact factors of process model understandability focuses on the product or outcome of the process modeling act [5, 6, 7], identifying features that make models easier to understand [8]. In particular, features that characterize the layout of these models, which are part of their secondary notation, have been shown to have an effect on model understandability [9, 10].

To promote the creation of understandable models and to overcome quality problems right from the start, we need to support humans during the creation of business process models [11]. However, our knowledge about the modeling behavior regarding the model's layout is very limited. Investigations of the process of process modeling [12] identified distinct modeling styles that differ from each other in the amount of editing operations devoted to reconciliation of the model while creating it. Yet, we do not know what layout features (e.g., crossing edges) are addressed at what phase of the modeling process nor the relation between the modeling process and the final features of the resulting model. We believe that to efficiently support the creation of understandable models with good quality, we need to strengthen our understanding of the layouting behavior of modelers while creating process models.

This paper extends existing research on the process of process modeling, i.e., the act of formalizing a process into a process model. Pinggera et al proposed a visualization of the process of process modeling called modeling phase diagrams [13]. The modeling phase diagram is a graphical analysis technique visualizing the low level interactions during the creation of a process model by mapping these onto phases. In this paper we introduce the advanced modeling phase diagram that overlays the modeling phase diagram with layout metrics focusing on layout properties presented in [5], allowing us to visually analyze the evolution of these metrics. We present examples showing how our analysis technique enables us to investigate the modeler's behavior toward selected layout metrics. Hence, this technique contributes insights to understand and support the modeler during the process of process modeling, fostering the creation of understandable models. In addition, we outline challenges regarding the interpretation of advanced modeling phase diagrams for further research.

The remainder of the paper is structured as follows. Sect. 2 gives required information about the process of process modeling (Sect. 2.1) and selected layout metrics (Sect. 2.2). Sect. 3 describes the advanced modeling phase diagram (Sect. 3.1) and demonstrates them along several examples (Sect. 3.2). In addition, it deals with challenges the advanced modeling phase diagram meets (Sect. 3.3) and limitations (Sect. 3.4). Finally, Sect. 5 concludes the paper.

2 Background

Before introducing the advanced modeling phase diagram, this section gives insight to required background knowledge. In particular, Sect. 2.1 introduces process of process modeling features and their operationalization through the modeling phase diagram. Sect. 2.2 gives an overview of selected layout metrics, which are used to demonstrate the use of the advanced modeling phase diagram.

2.1 Process of process modeling

The lifecycle of process model development can be characterized as an iterative and collaborative process, involving an elicitation and formalization phase [14]. In the elicitation phase information is extracted by the domain experts from the domain. This process is described in literature as a negotiation process [15]. The extracted information is then used in the formalization phase by the process modeler for creating the formal process model and validating it. This process is also denoted as process of process modeling. [13] states that the modeler's interaction with the tool, i.e., the process of process modeling, can be seen as a cycle of three successive phases, namely comprehension, modeling, and reconciliation. During the comprehension phase the modeler tries to assess the requirements to be modeled. This knowledge is used during the modeling phase to create the process model. Afterwards, the modeler enhances the understandability of the process model in the reconciliation phase (e.g., he moves modeling elements to a new position in the process model).

To analyze the process of process modeling in a systematic manner, the Cheetah Experimental Platform (CEP) [16] is used. CEP offers a basic process modeling editor and logs any user interaction together with the corresponding time stamp. These event logs describe the creation of the process model step by step, and their replay allows analyzing the process of process modeling at any point in time. [13] offers an overview of all possible interactions and their classification into comprehension, modeling, and reconciliation phases. This classification is also used to visualize the process of process modeling by mapping the model interactions onto the modeling phases (i.e., comprehension, modeling, and reconciliation), resulting into modeling phase diagrams [13].

2.2 Layout metrics

In order to investigate the modeler's behavior regarding the model's layout, we focus on layout metrics presented in [5]. All of these metrics have been implemented [17]. For this paper, we selected two of these layout metrics. In particular, we introduce the crossing edge metric and the orthogonal segments metric. These metrics were selected as examples of clearly visible properties which are independent of one another.

Crossing edge metric. Literature suggests that models should not contain any crossing edges, as models without them are more comprehensible [18, 6]. The crossing edge metric for a process model is described by the number of edge crossings in the model divided by the total number of the model's edges [5]. Therefore, in periods without crossing edges the value of the metric is 0. Note that not all process models can reach a value of 0 upon their completion, as some models cannot be laid out without any edge crossings, i.e., not all process models are planar (cf. [19]). In addition, a trade-off between the organization of modeling elements to patterns or structures and the number of crossing edges needs to be considered [8].

Orthogonal segments metric. [19] proposes that edges should make use of a Manhattan layout to establish a readable layout, i.e., edges should be aligned according an orthogonal layout that consists of horizontal and vertical lines. An edge consists of either one segment or more segments divided by bendpoints. A segment is considered orthogonal if it is parallel to either the horizontal axis or the vertical axis, with a threshold of up to 7 pixels. Fig. 1 shows three orthogonal

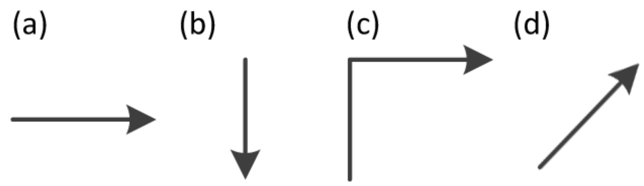

Fig. 1. Examples of orthogonal an non orthogonal segments

segments and one segment that is not orthogonal. Fig. 1 a) depicts a segment that is an orthogonal horizontal line, Fig. 1 b) shows a segment that is an orthogonal vertical line, and Fig. 1 c) is an edge consisting of two orthogonal segments divided by one bendpoint. Fig. 1 d) depicts a segment which is neither a horizontal nor a vertical orthogonal line. The orthogonal segments metric is calculated by the number of orthogonal segments in a process model divided by the number of all segments in the model [5]. Hence, if a process model only contains orthogonal segments, the value of the metric is 1. However, following the Manhattan layout too strictly can lead to a high number of bendpoints, hampering the readability of the model [19].

3 Investigating the evolution of layout metrics

This section presents our approach of investigating the evolution of layout metrics. While Sect. 3.1 introduces the advanced modeling phase diagram, Sect. 3.2 demonstrates several examples. Sect. 3.3 summarizes challenges that appear when using the advanced modeling phase diagram. Finally, Sect. 3.4 lists several generalization limitations.

3.1 Advanced modeling phase diagram

The advanced modeling phase diagram is an enhancement of the modeling phase diagram described in Sect. 2.1. It enables to overlay the traditional modeling phase diagram with selected metrics, allowing us to visually analyze the evolution of these metrics. In particular, each interaction with CEP is used to plot the corresponding metric values, i.e., after each interaction with the process model the new model is computed, triggering the computation of all implemented layout metrics.

For example, Fig. 2 shows an advanced modeling phase diagram with the orthogonal segments metric. Like for the modeling phase diagram, the horizontal axis depicts time while the left vertical axis represents the model's number of elements. The participant's interactions with CEP drawn in black are divided into comprehension, modeling, and reconciliation phases. In addition, the evolution of the orthogonal segments metric is depicted in blue. The value of the metric is shown on the right vertical axis.

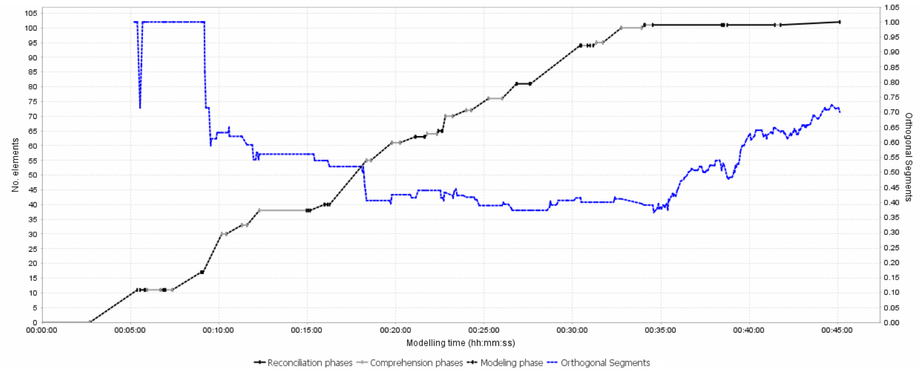

Fig. 2. An advanced modeling phase diagram with the orthogonal segments metric

3.2 Demonstration

This section focuses on the evolution of selected layout metrics described in Sect. 2.2.

Methodology. We analyzed the advanced modeling phase diagrams for a dataset gained from an modeling session described in [20]. The modeling session was conducted in December 2012 with students from Eindhoven University of Technology with CEP. After an initial phase consisting of a demographic survey and a CEP tutorial, the 120 participants were asked to model a process about getting a mortgage with BPMN. CEP logged any user interaction together with the corresponding time stamp. By using the event logs we created advanced modeling phase diagrams for each process model. Afterwards, the dataset was analyzed manually, i.e., we investigated the advanced modeling phase diagrams for selected layout metrics (cf. Sect. 2.2)[1]. These metrics were selected due to their diversity, e.g., while the crossing edge metric usually is only altered from time to time (i.e., when a crossing edge is introduced), the orthogonal segments metric is affected by almost every interaction with the process model that deals with edges (e.g., add edge, delete edge, move edge, etc.). For a more qualified interpretation of the diagrams we also looked at the replays of the event logs of the user interactions (cf. [16]). In the following, we demonstrate some examples showing reoccurring behaviors that we will discuss afterwards.

Crossing edge metric. Fig. 3 shows different examples we could repeatedly observe in our dataset showing different reoccurring behavior of how participants dealt with crossing edges. For readability considerations, we removed the modeling phase plot and left only the one related to the crossing edges metric. In general, the calculation of the crossing edge metric starts with the first edge

[1] A high resolution of all diagrams in this paper including the modeling phases as well as the corresponding process models can be downloaded from: http://bpm.q-e.at/EvolutionOfLayoutMetrics

added to the model. Therefore, the initial value of the metric is typically 0 since there are usually no crossing edges in a model containing only one edge. Most participants introduced crossings edges at some point during process model creation, leading to (temporary) increases of the crossing edge metric. Overall, we observed two reasons why a crossing edge appeared: either the participant introduced it by moving a modeling element that was connected to at least one other modeling element or the tool added it as the participant connected two modeling elements. A tool induced crossing edge can happen, e.g., due to connecting two modeling elements that are on opposite sides of already connected modeling elements. In this case, the tool adds the new edge above the existing ones, without trying to find a way around the existing edges. While crossing edges occurred for almost all participants (for 119 from 120), the reactions of the participants to a crossing edge were quite different.

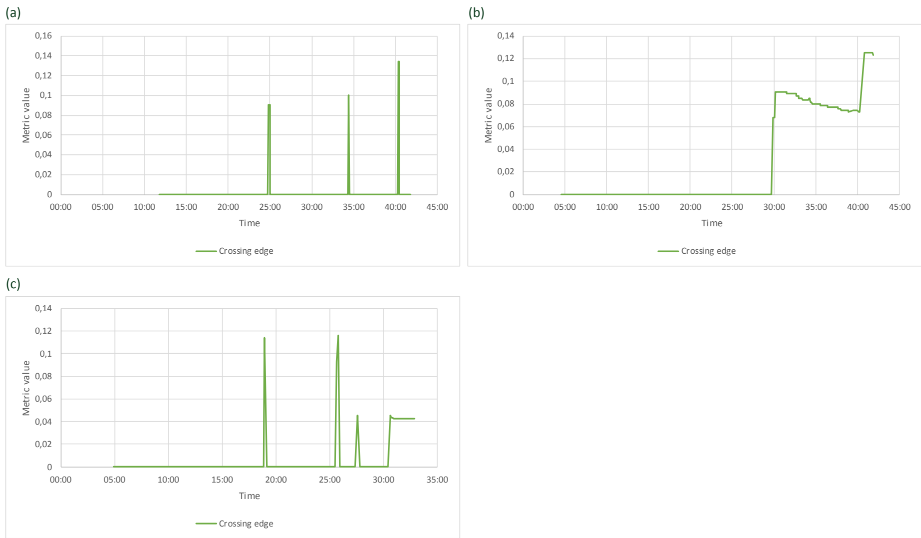

Fig. 3. Examples for the evolution of the crossing edge metric

Fig. 3 (a) shows an example of a participant who immediately resolved crossing edges that appeared throughout the modeling process. This is reflected by the 3 steep spikes (crossing edges appeared but were resolved within a very short timeframe). The value of the crossing edge metric was 0 throughout the whole session except for the 3 spikes. Fig. 3 b), in turn, displays the opposite behavior. Edge crossings that were introduced were never resolved, leading to a continuous increase of the metric. The drop between minutes 30 and 40 is not because of an effort the participant undertook to remove the crossing edges, but can be explained by the fact that the modeler added additional modeling elements, causing the metric to drop slightly till the next crossing edge appeared. Fig. 3 c) depicts another typically reoccurring behavior, showing a participant that initially resolved edge crossings immediately, but who left them unattended towards the end of the modeling process. We found three alternative explanations for this situation when examining the modeling replays. First, due to an error the process model was non-planar and therefore it was not possible to resolve the crossing edge. Second, it was too complex to resolve it, i.e., it would have taken many move operations to resolve one crossing edge. Third, the participant could have easily resolved the crossing edge with only a few move operations, but instead decided to finish the modeling session.

Orthogonal segments metric. Fig. 4 depicts typical reoccurring behaviors of the evolution of the orthogonal segments metric we identified. With the introduction of the first edge, which is typically orthogonal, the initial value of the metric is 1 since all existing edges (i.e., all segments of this one edge) are orthogonal. As CEP does not provide any tool support for achieving orthogonality, a model with completely orthogonal segments was not achieved by any of the 120 participants of the modeling session. Again, different examples of reoccurring behavior are discussed subsequently.

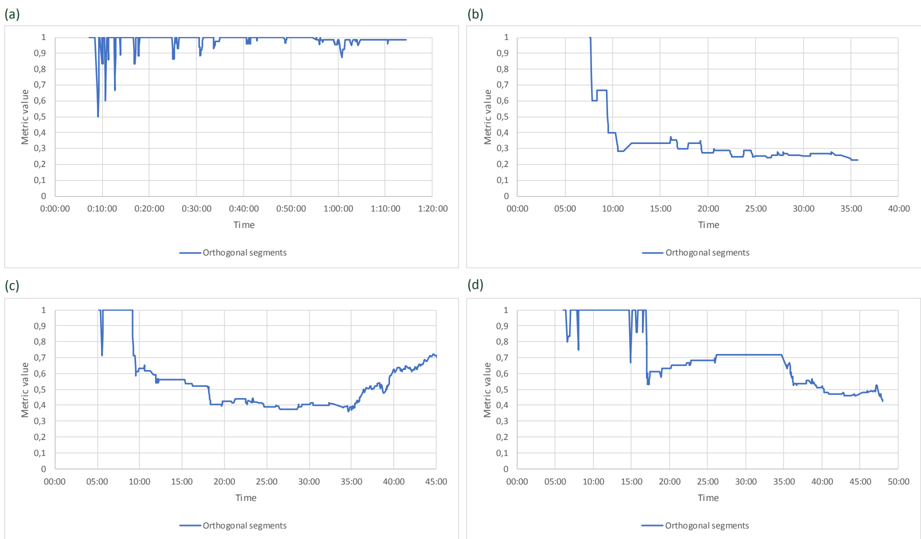

Fig. 4. Examples for the evolution of the orthogonal segments metric

For instance, Fig. 4 a) shows an evolution of the orthogonal segments metric where the participant clearly cared about this metric. The values for the orthogonal segments metric are constantly high (close to 1) except for a few downward spikes. The downward spikes of the metric were due to adding new edges that were not orthogonally aligned by the tool. However, the participant immediately reestablished the orthogonality of the segments. Overall, it can be noted that the spikes in the evolution get smaller throughout the modeling process. If the model already includes many edges, one segment which is not orthogonal does not affect the metric as much as when there are only few edges in the model (i.e., at the beginning of the modeling session). Fig. 4 b) highlights another reoccurring behavior where no noticeable efforts can be observed to improve the orthogonality of the model. After modeling the first modeling elements with high segment orthogonality, the value of the metric was mainly decreasing. Fig. 4 c), in turn, depicts an example of a participant who initially left the segment's orthogonality unattended (i.e., values of the metric have a downward tendency), but at the end of the modeling process made substantial efforts to improve the orthogonality (i.e., the participant applied several move operations to increase orthogonality of the segments). This is also in accordance with the modeling phase diagram of Fig. 4 c) which can be seen in Fig. 2. The modeling phase diagram shows that the increase of the metric happened during reconciliation phases. It was not always the case that only one specific behavior could be noticed, but we observed several modeling sessions with mixed behavior. For example, Fig. 4 d) shows a mix of different reactions toward the orthogonal segments metric. First the participant immediately reestablished the orthogonality of the segments. Then the value decreases, substantially at minute 15, and from then onwards no efforts could be observed to improve the metric. The slight increase after minute 15 can be explained due to adding new edges that only consisted of orthogonal segments, rather than the result of any reconciliation effort.

3.3 Interpretation challenges

The examples shown in Sect. 3.2 demonstrate different reoccurring behaviors which contribute towards a better understanding how modelers layout their models. However, this does not come without several difficulties. This section summarizes these challenges that appeared when using the advanced modeling phase diagram for the interpretation of the evolution of layout metrics.

Challenge C1. There is a vast diversity of layout metrics which all need to be interpreted separately [5]. For example, while a low value for the crossing edge metric indicates a good quality, the opposite is the case for the orthogonal segments metric (cf. Sect. 2.2). Therefore, an automatic interpretation needs to consider each layout metric individually.

Challenge C2. Even within one metric, an increasing or decreasing value does not always indicate the same. For instance, for the orthogonal segments metric an increase usually indicates that a participant changed non orthogonal segments to be orthogonal. However, in Fig. 4 d) the increase of the metric after minute 15 happened due to adding new edges, rather than an effort made by the participant. Therefore, the modeling phase diagram is crucial for interpretation. It shows if a change of the metric happened during a modeling or a reconciliation phase, restricting possible interpretations of the metric (cf. Fig. 2).

Challenge C3. Looking at only one metric at a time does not consider any trade-offs between different layout metrics the modeler has to make. For example, as already mentioned in Sect. 2.2, a modeler might not adhere to a strict Manhattan layout in order to have a low number of bendpoints (cf. broken edges metrics [5]). The implementation of the advanced modeling phase diagram is already able to depict more than one metric at a time, allowing us to detect potential trade-offs. However, analyzing trade-offs automatically is still challenging.

Challenge C4. Some metrics require information about the context in order to be able to interpret the evolution properly. For instance, layout properties relating to the shape or area of the model (cf. [5]) are typically not intentionally handled by the modeler and simply increase. However, as the model grows bigger the modeler might switch to consciously manipulate these metrics in order to fit the process model to the screen size and to avoid scrolling. Therefore, knowing the screen size available to the modeler (i.e., the context) is important for understanding the evolution of metrics relating to the shape or area of a process model.

Challenge C5. The stability of layout metrics might change during the evolution of the metrics due to the way how they are calculated. As already mentioned in Sect. 3.2, this is the case for the orthogonal segments metric, i.e., at the start of the modeling session it is very sensitive to changes, while it gets more stable at the end (cf. Fig. 4 a). Therefore, the interpretation of metrics in initial phases needs to be taken with care.

Challenge C6. Based on the data which is automatically collected by CEP and without additional verbal statements (e.g., think-aloud protocols or interviews), the modeler's intentions cannot always be inferred with certainty. For example, for the crossing edge metric we could identify three possible explanations why one last remaining crossing edge was not resolved (cf. Fig. 3 c). We are able to detect if the model changed to a non-planar one due to an error. However, if that was not the case, it remains unclear why the participant did not resolve the last crossing edge (e.g., if he did not care about it or if he overlooked it).

3.4 Limitations

Our interpretation of the evolution of layout metrics for process models has to be viewed in the light of several generalization limitations. First, in the sense of notational support, we focused on the basics of BPMN, i.e., a notation that is typically taught to a large group of stakeholders [21]. The evolution of layout metrics for the complete set of BPMN or different notations was not considered. Second, all participants were students. However, the participants reported an average of 2.90 years of modeling experience (cf. [20]). Third, our investigation was focused at only one process, i.e., a process about getting a mortgage. Nonetheless we analyzed 120 implementations of this process. Fourth, in this paper we only consider two layouting quality metrics. Fifth, we also do not consider any trade-offs decisions the participants may have taken regarding the layout, influencing the chosen layout metrics. Sixth, participants had no automatic layout support, as layouting algorithms would influence the layout metrics. It should be noted that these decision was made deliberately (cf. Sect. 2.2), highlighting future research directions.

4 Related Work

The work presented in this paper is focused on the model phase diagram described in [13] and [20] (cf. Sect. 2.1). While the modeling phase diagram abstracts the interactions with the process model to comprehension, modeling, and reconciliation phases, [22] introduces PPMCharts (Process of Process Modeling Charts) showing all information available of the recorded operations. These charts are based on the Dotted Chart Analysis plug-in of the process mining framework ProM, providing fine grained details of the construction of the model as well as an overview of the entire modeling process. Apart from this research, [23] investigates the process of modeling focusing on graphical modeling, showing how visualization can aid the development and assessment of computational models for large data sets.

5 Summary and Outlook

While current research focuses on the product or outcome of business process modeling, this paper motivates the need to focus on the act of formalizing a process into a process model itself. In particular, we are interested in the modeler's behavior regarding the layout of a process model. By visualizing the evolution of layout metrics, the advanced modeling phase diagram introduced in this paper contributes toward an in-depth understanding of the process of process modeling, specifically how modelers layout their process models. In particular, our examples show the evolution of two different layout metrics and present how they can be analyzed. Moreover, we list different challenges that need to be considered at future work.

In short term, our follow up research will be concerned about these challenges. For example, we consider investigating how layout metrics influence each other. The implementation of the advanced modeling phase diagram already allows to show more than one metric at a time, enabling us to detect potential trade-offs. In addition, we are planning on providing an automatic interpretation of different evolutions of layout metrics. This allows us, e.g., to quickly detect if certain layout metrics are more important to our participants than others. In the long term, our interest is how we can support the modeler during the creation of a process model. In particular, we plan a recommendation system providing suggestions regarding the layout as the user is modeling. This system is expected to point out specific layout issues the user should improve. Not only should this enhance the layout of the model, but also the overall quality of the end product by influencing other quality dimensions as well. We believe that an efficient support of modelers during the creation of process models fosters understandability and quality of business process models.

Acknowledgements. This research is supported by Austrian Science Fund (FWF): P26140

References

1. Becker, J., Rosemann, M., von Uthmann, C.: Guidelines of business process modeling. In: Business Process Management, Models, Techniques and Empirical Studies. (2000) 30–49
2. Rittgen, P.: Quality and perceived usefulness of process models. In: Proc. SAC'10. (2010) 65–72
3. Scheer, A.W.: ARIS—Business Process Modeling, 3rd ed. Springer (2000)
4. Kock, N., Verville, J., Danesh-Pajou, A., DeLuca, D.: 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
5. Bernstein, V., Soffer, P.: Identifying and quantifying visual layout features of business process models. In: Proc. BPMDS'15, Springer (2015) 200–213
6. Purchase, H.: Which Aesthetic has the Greatest Effect on Human Understanding? In: Proc. GD'97. (1997) 248–261
7. Leopold, H., Mendling, J., Günther, O.: Learning from quality issues of BPMN models from industry. IEEE Software **33**(4) (2016) 26–33
8. Moody, D.L.: The "Physics" of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. IEEE Trans. Software Eng. **35**(6) (2009) 756–779
9. Rosa, M.L., ter Hofstede, A., Wohed, P., Reijers, H., Mendling, J., van der Aalst, W.P.: Managing Process Model Complexity via Concrete Syntax Modifications. IEEE Trans. Industrial Informatics **7**(2) (2011) 255–265
10. Schrepfer, M., Wolf, J., Mendling, J., Reijers, H.A.: The impact of secondary notation on process model understanding. In: The Practice of Enterprise Modeling, Second IFIP WG 8.1 Working Conference, PoEM 2009, Stockholm, Sweden, November 18-19, 2009. Proceedings. (2009) 161–175
11. Mendling, J.: Metrics for Process Models: Empirical Foundations of Verification, Error Prediction and Guidelines for Correctness. Springer (2008)
12. Pinggera, J., Soffer, P., Fahland, D., Weidlich, M., Zugal, S., Weber, B., Reijers, H., Mendling, J.: Styles in business process modeling: an exploration and a model. Software & Systems Modeling (2013)
13. Pinggera, J., Zugal, S., Weidlich, M., Fahland, D., Weber, B., Mendling, J., Reijers, H.: Tracing the Process of Process Modeling with Modeling Phase Diagrams. In: Proc. ER-BPM'11. (2012) 370–382
14. Hoppenbrouwers, S., Proper, H., Weide, T.: A fundamental view on the process of conceptual modeling. In: Proc. ER'05. (2005) 128–143
15. Rittgen, P.: Negotiating Models. In: Proc. CAiSE'07. (2007) 561–573
16. Pinggera, J., Zugal, S., Weber, B.: Investigating the process of process modeling with cheetah experimental platform. In: Proc. ER-POIS'10. (2010) 13–18

17. Burattin, A., Bernstein, V., Neurauter, M., Soffer, P., Weber, B.: Detection and quantification of flow consistency in business process models. CoRR **abs/1602.02992** (2016)
18. Petre, M.: Why Looking Isn't Always Seeing: Readership Skills and Graphical Programming. Communications of the ACM (1995) 33–44
19. Gschwind, T., Pinggera, J., Zugal, S., Reijers, H., Weber, B.: A Linear Time Layout Algorithm for Business Process Models. Technical Report RZ3830, IBM Research (2012)
20. Pinggera, J.: The Process of Process Modeling. PhD thesis, University of Innsbruck, Department of Computer Science (2014)
21. Recker, J.: Opportunities and constraints: the current struggle with bpmn. Business Process Management Journal **16**(1) (2010) 181–201
22. Claes, J., Vanderfeesten, I., Pinggera, J., Reijers, H., Weber, B., Poels, G.: Visualizing the Process of Process Modeling with PPMCharts. In: Proc. TAProViz '12. (2013) 744–755
23. Crapo, A.W., Waisel, L.B., Wallace, W.A., Willemain, T.R.: Visualization and the process of modeling: a cognitive–theoretic view. In: Proc. KDD'00. (2000) 218–226