

# Modeling Styles in Business Process Modeling

Jakob Pinggera<sup>1</sup>, Pnina Soffer<sup>2</sup>, Stefan Zugal<sup>1</sup>, Barbara Weber<sup>1</sup>, Matthias Weidlich<sup>3</sup>, Dirk Fahland<sup>4</sup>, Hajo A. Reijers<sup>4</sup>, and Jan Mendling<sup>5</sup>

<sup>1</sup> University of Innsbruck, Austria

{jakob.pinggera, stefan.zugal, barbara.weber}@uibk.ac.at

<sup>2</sup> University of Haifa, Israel

spnina@is.haifa.ac.il

<sup>3</sup> Eindhoven University of Technology, The Netherlands

d.fahland|h.a.reijers@tue.nl

<sup>4</sup> Technion - Israel Institute of Technology

weidlich@tx.technion.ac.il

<sup>5</sup> Vienna University of Economics and Business

jan.mendling@wu.ac.at

**Abstract.** Research on quality issues of business process models has recently begun to explore the process of creating process models. As a consequence, the question arises whether different ways of creating process models exist. In this vein, we observed 115 students engaged in the act of modeling, recording all their interactions with the modeling environment using a specialized tool. The recordings of process modeling were subsequently clustered. Results presented in this paper suggest the existence of three distinct modeling styles, exhibiting significantly different characteristics. We believe that this finding constitutes another building block toward a more comprehensive understanding of the process of process modeling that will ultimately enable us to support modelers in creating better business process models.

**Key words:** business process modeling, process of process modeling, modeling styles, cluster analysis

## 1 Introduction

Considering the heavy usage of business process modeling in all types of business contexts, it is important to acknowledge both the relevance of process models and their associated quality issues. However, actual process models display a wide range of problems [1]. Following the SEQUAL framework [2], quality dimensions of models include syntactic, semantic, and pragmatic quality. Syntactic and semantic quality relate to model construction, and address the correct use of the modeling language and the extent to which the model truthfully represents the real world behavior it should depict, respectively. Pragmatic quality addresses the extent to which a model supports its usage for purposes such as understanding behavior or developing process aware systems. Considering process models whose purpose is to develop an understanding of real world behavior, pragmatic quality is typically related to the understandability of the model [3]. Clearly, an

in-depth understanding of the factors influencing the various quality dimensions of process models is in demand.

Most research in this area puts a strong emphasis on the product or outcome of the process modeling act (e.g., [4, 5]). For this category of research, the resulting model is the object of analysis. Many other works—instead of dealing with the quality of individual models—focus on the characteristics of modeling languages (e.g., [6, 7]). Recently, research has begun to explore another dimension presumably affecting the quality of business process models by incorporating the process of creating a process model into their investigations (e.g., [8, 9]). In particular, the focus has been put on the formalization phase in which a process modeler is facing the challenge of constructing a syntactically correct model reflecting a given domain description (cf. [10]). Our research can be attributed to the latter stream of research.

This paper contributes to our understanding of the process of process modeling (PPM) by investigating whether different ways of process modeling can be identified, i.e., can we observe different modeling styles when modelers create process models? Knowledge about different modeling styles will support the creation of customized process modeling environments, supporting modelers in creating high quality models. Similarly, a more comprehensive understanding of the PPM can be exploited for teaching students in how to create process models of high quality. We conducted a modeling session with 115 students, recording all their interactions with the modeling environment using a specialized tool. To identify different modeling styles the collected PPM instances were automatically clustered suggesting the existence of three different modeling styles. The modeling styles were subsequently analyzed using a series of measures for quantifying the PPM to validate differences between the three groups.

The paper is structured as follows. Section 2 presents backgrounds on the PPM and introduces measures for quantifying this process. Section 3 describes data collection and cluster analysis. Section 4 presents the results, followed by their discussion in Section 5. The paper is concluded with a discussion of related work in Section 6 and a brief summary in Section 7.

## 2 Backgrounds

This section provides background information on the PPM and explains how this process can be captured and quantified using a series of measures.

### 2.1 The Process of Process Modeling

During the formalization phase process modelers are working on creating a syntactically correct process model reflecting a given domain description by interacting with the process modeling tool [10]. This modeling process can be described as an iterative and highly flexible process [11, 12], dependent on the individual modeler and the modeling task at hand [13]. At an operational level, the modeler’s interactions with the tool would typically consist of a cycle of the three successive phases of (1) comprehension (i.e., the modeler forms a mental model

of domain behavior), (2) modeling (i.e., the modeler maps the mental model to modeling constructs), and (3) reconciliation (i.e., the modeler reorganizes the process model) [9, 8].

**Comprehension.** Research on human cognition and problem solving has shed light on comprehension. According to [14], when facing a task, the problem solver first formulates a mental representation of the problem, and then uses it for reasoning about the solution and which methods to apply for solving the problem. In process modeling, the task is to create a model which represents the behavior of a domain. The process of forming mental models and applying methods for achieving the task is not done in one step applied to the entire problem. Rather, due to the limited capacity of working memory, the problem is broken down to pieces that are addressed sequentially, chunk by chunk [8, 9].

**Modeling.** The modeler uses the problem and solution developed in working memory during the previous comprehension phase to materialize the solution in a process model (by creating or changing it) [8, 9]. The modeler’s utilization of working memory influences the number of modeling steps executed during the modeling phase before forcing the modeler to revisit the problem for acquiring more information [9].

**Reconciliation.** After modeling, modelers typically reorganize the process model (e.g., renaming of activities) and utilize the process model’s *secondary notation* (e.g., notation of layout, typographic cues) to enhance the process model’s understandability [15, 16]. However, the number of reconciliation phases in the PPM is influenced by a modeler’s ability of placing elements correctly when creating them, alleviating the need for additional layouting [9].

## 2.2 Capturing Events of the Process of Process Modeling

To investigate the PPM, actions taken during modeling have to be recorded and mapped to the phases described above. When modeling in a process modeling environment, process modeling consists of adding nodes and edges to the process model, naming or renaming activities, and adding conditions to edges. In addition to these interactions a modeler can influence the process model’s secondary notation, e.g., by laying out the process model using move operations for nodes or by utilizing bendpoints to influence the routing of edges, see [9] for details.

To capture modeling activities, and for obtaining a closer look on how process models are created in a systematic manner, we instrumented a basic process modeling editor to record each user’s interactions together with the corresponding time stamp in an event log, describing the creation of the process model step by step. Editor and event recording are available within Cheetah Experimental Platform (CEP) [17].

## 2.3 Quantifying the Process of Process Modeling

A log of modeling events allows quantitative analysis of a PPM. Based on the conceptual background, comprehension (C), modeling (M), and reconciliation

(R) phases can be identified by grouping events into respective phases (see [9] for details). Then, a PPM can be divided into *modeling iterations* [9]. One iteration is assumed to comprise a comprehension (C), modeling (M), and reconciliation (R) phase in this respective order. The iterations of a modeling process are identified by aligning its phases to the CMR-pattern. If a certain phase of this pattern is not present in the modeling process, the respective phase is skipped for the observed iteration and the process is considered to continue with the next phase of the pattern. In the following we present five measures quantifying the process of process modeling.

**Number of Iterations.** This measure counts the modeling iterations per PPM reflecting how often a modeler had to interrupt modeling for comprehension or reconciliation.

**Share of Comprehension.** When comprehending, a mental model of the problem and a corresponding solution is developed which is then formalized in modeling phases. Differences in the amount of time spent on comprehension can be expected to characterize modeling styles and to impact on the modeling result. We quantify this aspect as the ratio of the average length of a comprehension phase in a process to the average length of an iteration. The initial comprehension phase is neglected as it is typically subject to various influences unrelated to problem solving (e.g., the modeler did not start immediately).

**Iteration Chunk Size.** Modelers can be assumed to conduct modeling in chunks of different sizes. We quantified *chunk size* as the average number of create and delete operations executed in one iteration. This measure reflects the ability to model large parts of a model without the need to comprehend or reconcile.

**Reconciliation Breaks.** A steady process of modeling is assumed to be a sequence of iterations following the CMR-pattern. Reconciliation can sometimes be skipped if the modeler can place all model elements directly at the right spot clearly alleviating the need for reconciliation. However, some processes may even show iterations of CR-patterns, i.e., an iteration without a modeling phase, where a modeler interrupts the common flow of modeling for additional reconciliation. We quantified this aspect by the relative share of iterations that comprise unexpected reconciliation (without modeling) out of all iterations.

**Delete Iterations.** From time to time, modelers are required to remove content from the process model. This might happen when modelers identify errors in the model, which are subsequently resolved by removing some of the modeling constructs and implementing the desired functionality. This measure describes the number of iterations of the PPM containing delete operations relative to the total number of iterations of the PPM.

### 3 Clustering

To be able to make generalizations, we have used cluster analysis to a set of PPM instances. Cluster analysis allows us to identify groups of modelers exhibiting

similar modeling styles. This section describes the modeling session, data pre-processing and cluster analysis.

### 3.1 Data Collection

The modeling session was designed to collect PPM instances of students creating a formal process model in BPMN from an informal description. The object that was to be modeled is a process describing the activities a pilot has to execute prior to taking off an aircraft<sup>1</sup>.

To mitigate the risk that the PPM instances were impacted by complicated tools or notations [11], we decided to use a subset of BPMN for our experiment. In this way, modelers were confronted with a minimal number of distractions, but the essence of how process models are created could still be captured. A pre-test was conducted at the University of Innsbruck to ensure the usability of the tool and the understandability of the task description. This led to further improvements of CEP and minor updates to the task description.

The modeling sessions were conducted in November 2010 with students of a graduate course on Business Process Management at Eindhoven University of Technology and in January 2011 with students from Humboldt-Universität zu Berlin following a similar course. The modeling session at each university started with a demographic survey, followed by a modeling tool tutorial explaining the basic features of CEP. After that, the actual modeling task was presented in which the students had to model the above described “Pre-Flight” process. This was done by 102 students in Eindhoven and 13 students in Berlin. By conducting the experiment during class and closely monitoring the students, we mitigated the risk of falsely identifying comprehension phases due to external distractions. No time restrictions were imposed on the students.

### 3.2 PPM Profile for Clustering

When trying to identify different types of PPM instances using clustering, the question arises how to represent such a process to make clustering possible. Based on our previous experience with the PPM we decided to focus on four aspects. The adding of content, the removal of content, reconciliation of the model and comprehension time, i.e., the time when the modeler does not work on the process model. To also reflect that modeling is a time-dependent process, we do not just look at the total amount of modeling actions and comprehension, but on their *distribution* over time as follows. We sampled every process into segments of 10 seconds length. For each segment, we compute its *profile*  $(a, d, r, c)$ , i.e., the numbers  $a$ ,  $d$ , and  $r$  of add, delete, and reconciliation events, and the time  $c$  spent on comprehension. The *profile* of one PPM is then sequence  $(a_1, d_1, r_1, c_1)(a_2, d_2, r_2, c_2) \dots$  of its segments’ profiles. The  $a$ ,  $d$ , and  $r$  are obtained per segment by classifying each event according to Table 1. Adding

<sup>1</sup> Material download: <http://pinggera.info/experiment/ModelingStyles>

Interaction	Classification	Interaction	Classification
CREATE NODE	Adding	RENAME ACTIVITY	Reconciliation
DELETE NODE	Deleting	UPDATE CONDITION	Reconciliation
CREATE EDGE	Adding	MOVE NODE	Reconciliation
DELETE EDGE	Deleting	MOVE EDGE LABEL	Reconciliation
RECONNECT EDGE	Adding/Deleting	MODIFY EDGE BENDPOINT	Reconciliation

**Table 1.** Classification of CEP’s User Interactions

a condition to an edge was considered being part of creating an edge. Comprehension time  $c$  was computed as follows. Group events to intervals: an interval is a sequence of events where two consecutive events are  $\leq 1$  second apart, its duration is the time difference between its first and its last event (intervals of 1 activity got a duration of 1 second). Then  $c$  is the length of the segment (10 secs) minus the duration of all intervals in the segment. For example, if the modeler moved activity A after 3 secs, activity B after 3.5 secs and activity C after 4.2 secs the comprehension time in this segment would be  $10 - 1.2 = 8.8$  seconds. To give all PPM profiles equal length, shorter profiles were extended with segments of no interaction to reach the length of the longest PPM (required for clustering).

### 3.3 Clustering

The PPM profiles were exported from CEP [17] and subsequently clustered using Weka<sup>1</sup>. The KMeans algorithm, first proposed in [18], utilizing an euclidean distance measure was chosen for clustering as it constitutes a well known and easy to use means for cluster analysis. As KMeans might converge in a local minimum [19], the obtained clustering has to be validated. If the identified clusters exhibit significant differences with regard to the measures described in Section 2, we conclude that different modeling styles were identified. KMeans requires the number of cluster to be known a priori. As this was not the case we gradually increased the number of clusters starting from 2, resulting in only one major cluster. Setting the number of expected clusters to 3 revealed two major clusters and one cluster of 2 PPM instances. Most promising results were achieved by setting the number of clusters to be generated to 4 and starting with a seed of 10, returning 3 major clusters and one small cluster of 2 PPM instances. We considered these 3 major clusters for further analysis; increasing the number of expected clusters only generated additional small clusters.

## 4 Results

In this section we present results of the cluster analysis and validate the difference among the clusters using the measures described in Section 2.

<sup>1</sup> <http://www.cs.waikato.ac.nz/ml/weka>

Measure	C1	C2	C3
Number of instances	42	22	49
Avg. no. of adding operations	61.36	52.91	52.57
Avg. no. of deleting operations	10.81	3.91	4.55
Avg. no. of reconciliation operations	76.26	42.00	39.27
Avg. no. operations	148.43	98.82	96.39

**Table 2.** Statistics per cluster

#### 4.1 Three Clusters

We identified three major clusters of 42, 22 and 49 instances, called C1, C2, and C3 in the sequel. In order to visualize the obtained clusters we calculated the average number of adding, deleting and reconciliation operations per segment for each cluster. Additionally, we calculated the moving average of six segments, i.e., one minute, providing us with a smoother representation of the modeling processes presented in Figures 1, 2, and 3 for C1, C2, and C3 respectively. The horizontal axis denotes the segments into which the PPM instances were sampled. The vertical axis indicates the average number of operations that were performed in this segment. For example, a value of 0.8 for segment 9 (cf. Fig. 2) indicates that all modelers in this cluster averaged 0.8 adding operations within this 10 second segment.

C1 (cf. Fig. 1) is characterized by long PPM instances, as the first time the adding series reaches 0 is after about 205 segments. Additionally, the delete series indicates more delete operations compared to the other clusters. Several fairly large spikes of reconciliation activity can be observed, the most prominent one after about 117 segments.

C2, as illustrated in Fig. 2, is characterized by a fast start as a peak in adding activity is reached after 13 segments. In general, the adding series is most of the time between 0.5 and 0.9 operations, higher compared to the other two clusters. The fast modeling behavior results in short PPM instances as the adding series is 0 for the first time after about 110 segments.

On first sight, C3 (cf. Fig. 3) seems to be between C1 and C2. The adding curve is mostly situated between 0.4 and 0.7, a little lower than for C2, but still higher compared to C1. Similar values can be observed for the reconciliation curve. The deleting curve remains below 0.1. The duration of the PPM instances is also between the duration of C1 and C2 as the adding series is 0 for the first time after about 137 segments.

Table 2 presents general statistics on the number of adding operations, the number of deleting operations, the number of reconciliation operations and the total number of operations for each cluster. Interestingly, modelers in C1 had more adding operations, more deleting operations and, probably most notable, almost twice as many reconciliation operations compared to C2 and C3. At a first glance, the numbers for C2 and C3 appear to be very similar.

The following procedure for conducting the statistical analysis was used. If the data was normally distributed and homogeneity of variances was given we used Oneway ANOVA to test for differences among the groups. Pairwise comparisons

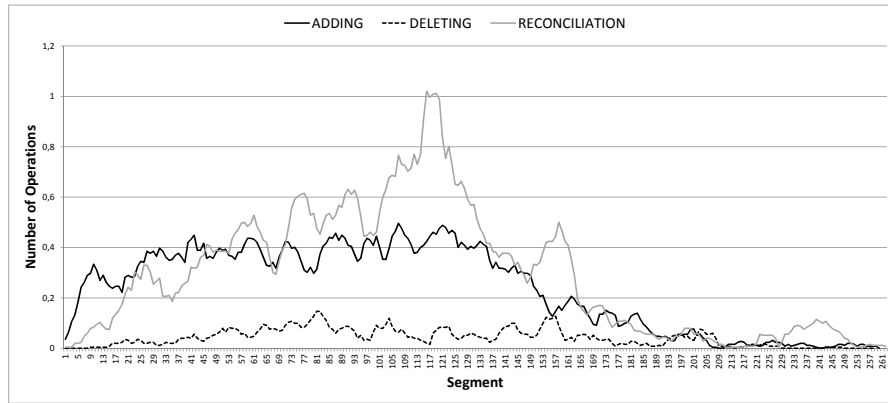


Fig. 1. Cluster C1

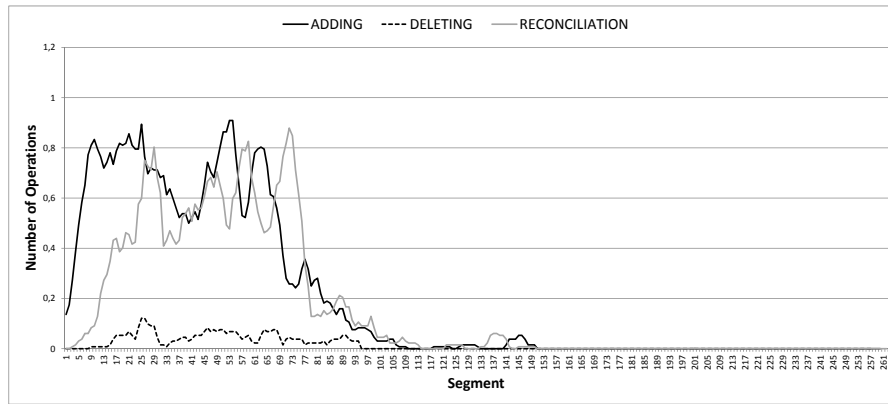


Fig. 2. Cluster C2

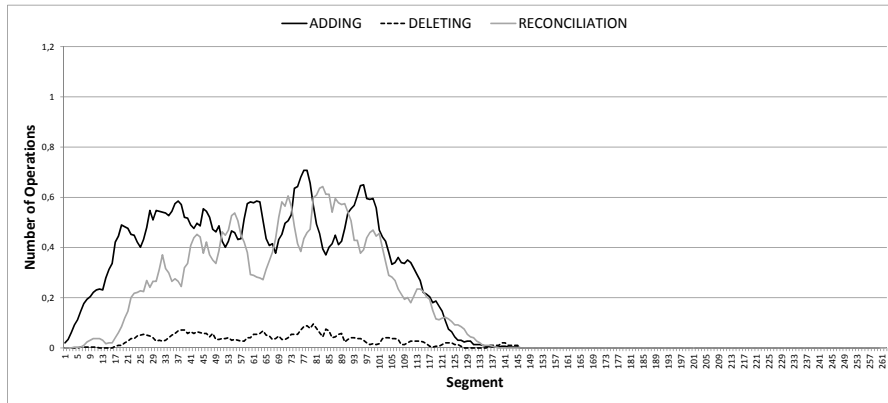


Fig. 3. Cluster C3



Statistic		All groups	Pairwise comparison		
			1-2	1-3	2-3
Number of Adding Operations	Sig.	0.000 <sup>a</sup>	0.003 <sup>a</sup>	0.000 <sup>a</sup>	
	test	Oneway ANOVA	Bonferroni post-hoc test		
Number of Deleting Operations	Sig.	0.000 <sup>a</sup>	0.000 <sup>b</sup>	0.000 <sup>b</sup>	
	test	Kruskall-wallis	Mann-whitney test		
Number of Reconciliation Operations	Sig.	0.000 <sup>a</sup>	0.000 <sup>b</sup>	0.000 <sup>b</sup>	
	test	Kruskall-wallis	t-test for unequal variances		
Number of Total Operations	Sig.	0.000 <sup>a</sup>	0.000 <sup>b</sup>	0.000 <sup>b</sup>	
	test	Kruskall-wallis	t-test for unequal variances		

<sup>a</sup>  $p < 0.05$     <sup>b</sup>  $p < 0.05/3$

**Table 3.** Significant differences for statistics

were done using the bonferroni post-hoc test. Note that the bonferroni post-hoc test uses an adapted significance level. Therefore, p-values less than 0.05 are considered to be significant, i.e., there is no need to divide the significance level by the number of groups, i.e., clusters. In case normal distribution or homogeneity of variance was not given a non-parametric alternative to ANOVA, i.e., kruskall-wallis, was utilized to test for differences among the groups. Pairwise comparisons were done using the t-test for (un)equal variances (depending on the data) if normal distribution was given. If no normal distribution could be identified the mann-whitney test was utilized. In either case, i.e., t-test or mann-whitney test, the bonferroni correction was applied, i.e., the significance level was divided by the number of clusters.

The results are summarized in Table 3, indicating significant differences between C1 and C2 and C1 and C3, but not between C2 and C3. Only significant differences are stated.

## 4.2 Applying Measures

In order to further distill the properties of the three clusters, we calculated the measures described in Section 2.3 for each PPM. Table 4 provides an overview presenting the average values for each measure in each cluster. As indicated in Fig. 1, C1 constitutes the highest number of PPM iterations. Tightly connected to this observation is the average iteration chunk size. Modelers in C2 added by far the most content per iteration to the process model. Also the number of iterations containing delete iterations is higher for C1 than for the other clusters, which is consistent with the higher number of delete operations (cf. Table 2). The amount of time spent on comprehending the task description and developing the plan on how to incorporate them into the process model seems to be far larger for C1 compared to C2, which has the lowest share of comprehension, but also larger compared to C3. When considering reconciliation breaks C3 sets itself apart posting the lowest number of reconciliation breaks. C2 is somewhere in between and C1 has the highest number of reconciliation breaks.

Statistical analysis of the differences between the groups was performed following the procedure described in the previous section. An overview of the results

Measure	C1	C2	C3
Avg. no. of PPM iterations	21.50	12.32	14.69
Avg. iteration Chunk Size	3.66	5.28	4.24
Avg. share of comprehension	49.88	39.28	45.02
Avg. reconciliation breaks	21.37	18.14	13.85
Avg. delete iterations	17.06	10.07	10.83

**Table 4.** Measures per cluster

Measure		All groups	Pairwise comparison		
			1-2	1-3	2-3
Iteration Chunk Size	Sig.	0.000 <sup>a</sup>	0.000 <sup>b</sup>	0.000 <sup>b</sup>	0.007 <sup>b</sup>
	test	Kruskall-wallis	t-test for unequal variances		
Number of Iterations	Sig.	0.000 <sup>a</sup>	0.000 <sup>b</sup>	0.000 <sup>b</sup>	0.004 <sup>b</sup>
	test	Kruskall-wallis	t-test for unequal variances		
Share of Comprehension	Sig.	0.000 <sup>a</sup>	0.000 <sup>a</sup>	0.036 <sup>a</sup>	0.045 <sup>a</sup>
	test	Oneway ANOVA	Bonferroni post-hoc test		
Delete Iterations	Sig.	0.005 <sup>a</sup>	0.026 <sup>a</sup>	0.011 <sup>a</sup>	
	test	Oneway ANOVA	Bonferroni post-hoc test		
Reconciliation Breaks	Sig.	0.005 <sup>a</sup>	0.004 <sup>a</sup>		
	test	Oneway ANOVA	Bonferroni post-hoc test		

<sup>a</sup>  $p < 0.05$     <sup>b</sup>  $p < 0.05/3$

**Table 5.** Significant differences for measures

is presented in Table 5. Only significant differences are stated. In contrast to the statistics presented in Table 3, we were able to identify significant differences between C2 and C3.

## 5 Discussion

In this section we present our insights when comparing the identified clusters and we discuss the lessons learned in this work and how they influence our future work. Additionally, limitations of this work are described.

### 5.1 Cluster C1

C1 can be clearly distinguished from C2 and C3. This becomes evident on visual inspection of Fig. 1, but also when considering the number of adding operations, the number of deleting operations, the number of reconciliation operations and the total number of operations. We identified statistically significant differences between C1 and C2 and between C1 and C3 for all statistics (cf. Table 3).

In general, modelers in C1 had rather long PPM instances, i.e., the number of PPM iterations was significantly higher compared to C2 and C3. In addition, modelers in C1 spent more time on comprehension compared to C2. Modelers started rather slowly, not eclipsing 0.5 adding operations. The slow modeling speed is underlined by the significantly lower chunk size compared to C2 and C3. During the whole process, adding operations are accompanied by a relatively high

amount of delete operations. This is underlined by the significant differences in the number of delete iterations between C1 and C2 and C1 and C3 (cf. Table 5). Also, we observed a fairly large amount of reconciliation operations, culminating in a massive peak after about half of the PPM instances.

The results suggest that modelers in C1 were not as goal oriented as their colleagues in other clusters, since they spent a great amount of time on comprehension, added more modeling elements which were subsequently removed and put significantly more effort into improving the visual appearance of the process model. There might be multiple reasons for this behavior. On the one hand, it could point toward modelers having trouble executing the modeling task and therefore needed more reconciliation to facilitate their understanding of the process model at hand. On the other hand, their focus on layouting might have acted as a distraction from the modeling task, resulting in the higher number of adding operations and deleting operations. Still, other techniques will be required for further investigating this claim, e.g., think aloud protocols (cf. [20]).

### 5.2 Cluster C2

When inspecting Fig. 2 the very steep start of the adding curve strikes the eye, indicating that modelers started creating the process model right away. When focusing on reconciliation operations, several spikes in the layouting curve can be identified, notably one last spike right after the number of adding operations decreases. As already mentioned above, C2 is statistically significant different compared to C1 for all statistics presented in Table 2. No differences can be identified between C2 and C3.

Considering the measures described in Section 2.3, C2 has a significantly higher chunk size compared to C1 and C3. Similarly, we observed the lowest number of PPM iterations. This means that modelers add a lot more content per PPM iteration. In addition, modelers in C2 did not spend as much time on comprehension compared to modelers in C1 and C3.

In a nutshell, modelers of C2 are very focused and goal oriented following a straight path when creating the process model. They are quick in making decisions about how to proceed and only slow down their modeling endeavor from time to time for some reconciliation, resulting in short PPM instances.

### 5.3 Cluster C3

Fig. 3 shows the PPM instances for C3. The processes are shorter compared to C1 and longer compared to C2. It is lacking the fast start of the adding curve we identified for C2. The reconciliation curve is more or less following the adding curve. Notably, this is the only curve without a reconciliation spike once the number of adding operations decreases.

The calculated measures indicate clear differences to C1 when it comes to chunk size, number of iterations, share of comprehension, but also number of delete operations and reconciliation breaks. C2 and C3 differ in chunk size, the number of iterations and the time spent on comprehension.

When comparing C2 and C3, the question arises whether modelers in C3 followed the same strategy as modelers in C2, just a little slower. We believe that, in contrary to C2, modelers of C3 followed a more systematic approach to process modeling. They continuously reconciled their process model, alleviating the need for dedicated reconciliation breaks. This is indicated by the lack of a reconciliation spike after the decrease of adding operations in Fig. 3. Additionally, reconciliation breaks points into this direction (18.16 for C2 vs. 13.85 for C3). Fig. 4 depicts the reconciliation breaks box plot, hinting at a difference in reconciliation breaks between C2 and C3. Still, the difference did not turn out to be statistically significant leaving us with some future work on investigating whether this claim actually holds.

Additionally, we believe that different reasons for reconciliation breaks exist. On the one hand, modelers are forced to stop their modeling endeavour and layout the process model when they are overwhelmed by the complexity at hand. On the other hand, some modelers might stop modeling at strategic points to reconcile the process model in order to avoid situations like the one mentioned above before they even arise. Even though this explanation would fit the boxplot depicted in Fig. 4, further investigations are in demand to fully understand the reconciliation behavior of modelers.

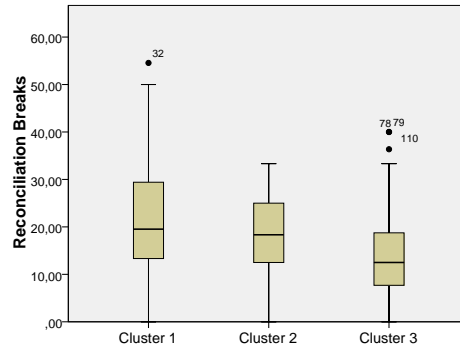


Fig. 4. Reconciliation Breaks

#### 5.4 Lessons Learned

We were able to identify three different modeling styles using cluster analysis. Differences among the clusters were subsequently validated using a series of measures quantifying the PPM. Note that these measures were defined prior to performing the cluster analysis. The measures are based on the detected iterations of the PPM, approaching the PPM from a different angle. Therefore they enable us to validate the differences among the three clusters.

The detected modeling styles contribute to our understanding of the process of process modeling, as, to our knowledge, this is the first systematic attempt to establish a categorization of PPM instances in the domain of business process modeling. We believe that further refinements of the categorization will emerge, ultimately enabling us to create personalized modeling environments based on their observed modeling behavior. In addition, these findings can be exploited for teaching purposes. For example, teachers might be able to identify students facing difficulties during a modeling assignment based on their modeling behavior and provide them with additional support. Still, some research questions emerging from these findings have to be addressed first. The most pressing might be whether a modeler's personal style persists over several different modeling

tasks or if the modeling style is determined by the modeling task at hand. To answer this question further empirical investigations are in demand. Based on some preliminary observations we would assume that the influence of the modeling task cannot be neglected. Even though a modeler might like to create a process model in a straight forward, goal oriented way, the complexity of the modeling task might force her to reduce the modeling speed and switch to a more conservative modeling style.

On a long-term basis questions on how to exploit this knowledge to improve the quality of the resulting process model become evident. Unfortunately, the naive assumption that one modeling style is superior to the others could not be confirmed. All clusters contained excellent process models and process models of low quality. This is not surprising though. Even modelers in C1 who face difficulties, exhibiting long PPM instances, can still come up with good process models if they succeed in overcoming the adversity they are facing.

### 5.5 Limitations

The interpretation of our findings is presented with the explicit acknowledgement of a number of limitations to our study. First of all, our respondents represented a rather homogeneous and inexperienced group. Although relative differences in experience were notable, the group is not representative for the modeling community at large. At this stage, in particular, the question can be raised whether experienced modelers also exhibit the same style elements as skillful yet inexperienced modelers. In other words, will experienced modelers display similar characteristics of style or can other styles be observed within their approaches? Note that we are mildly optimistic about the usefulness of the presented insights on the basis of modeling behavior of graduate students, since we have established in previous work that such subjects perform equally well in process modeling tasks as some professional modelers [21].

Secondly, the influence of the modeling task—more precisely, the modeling task’s complexity (cf. [22])—on the PPM is not fully understood. All students in our modeling session were working on the same modeling assignment. Hence, the observed clusters might be specific to modeling tasks of this complexity level. Further investigations will be necessary to let sunlight fall on the influence of the modeling task, which might result in the emergence of additional clusters. Preliminary results of a different modeling task suggest the existence of modeling styles comparable to the results presented in this paper.

Thirdly, we can not rule out that KMeans identified a local minimum, resulting in a suboptimal clustering. To counter this threat we validated the clustering using a series of measures quantifying the PPM and identified significant differences among the three groups.

## 6 Related Work

Our work is essentially related to model quality frameworks and research on the process of modeling.

There are different frameworks and guidelines available that define quality for process models. Among others, the SEQUAL framework uses semiotic theory for identifying various aspects of process model quality [3], the Guidelines of Process Modeling describe quality considerations for process models [23], and the Seven Process Modeling Guidelines define desirable characteristics of a process model [24]. While each of these frameworks has been validated empirically, they rather take a static view by focusing on the resulting process model, but not on the act of modeling itself. Our research takes another approach by investigating the process followed to create the process model.

Research on the process of modeling typically focuses on the interaction between different parties. In a classical setting, a system analyst directs a domain expert through a structured discussion subdivided into the stages elicitation, modeling, verification, and validation [10, 25]. The procedure of developing process models in a team is analyzed in [26] and characterized as a negotiation process. Interpretation tasks and classification tasks are identified on the semantic level of modeling. Participative modeling is discussed in [27]. These works build on the observation of modeling practice and distill normative procedures for steering the process of modeling towards a good completion. Our work, in turn, focuses on the formalization of the process model, i.e., the modeler's interactions with the modeling environment when creating the formal business process model.

## 7 Summary

This paper contributes to our understanding of the PPM as it constitutes the first systematic attempt to identify different modeling styles in the domain of business process modeling. We conducted a modeling session with 115 students of courses on business process management, collecting their PPM instances. We were able to identify three different modeling styles using cluster analysis and validated the retrieved clusters using a series of measures for quantifying the PPM. We believe that a better understanding regarding the PPM will be beneficial for future process modeling environments and will support teachers in mentoring their students on their way to professional process modelers.

**Acknowledgements.** This research was funded by the Austrian Science Fund (FWF): P23699-N23. The research stay of Dr. Pnina Soffer in Innsbruck, Austria was funded by BIT School.

## References

1. Mendling, J.: Metrics for Process Models: Empirical Foundations of Verification, Error Prediction, and Guidelines for Correctness. Springer (2008)
2. Lindland, O.I., Sindre, G., Sølvsberg, A.: Understanding Quality in Conceptual Modeling. *IEEE Softw.* **11** (1994) 42–49
3. Krogstie, J., Sindre, G., Jørgensen, H.: Process models representing knowledge for action: a revised quality framework. *EJIS* **15** (2006) 91–102

4. Van der Aalst, W., Ter Hofstede, A.: Verification of workflow task structures: A petri-net-based approach. *IS* **25** (2000) 43–69
5. Gruhn, V., Laue, R.: Complexity metrics for business process models. In: Proc. ICBIS '10. (2006) 1–12
6. Siau, K., Rossi, M.: Evaluation techniques for systems analysis and design modelling methods—a review and comparative analysis. *ISJ* **21** (2011) 249–268
7. Moody, D.L.: The "Physics" of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. *IEEE Trans. Software Eng.* **35** (2009) 756–779
8. Soffer, P., Kaner, M., Wand, Y.: Towards Understanding the Process of Process Modeling: Theoretical and Empirical Considerations. In: Proc. ER-BPM '11. (2011) 357–369
9. Pinggera, J., Zugal, S., Weidlich, M., Fahland, D., Weber, B., Mendling, J., Reijers, H.A.: Tracing the process of process modeling with modeling phase diagrams. In: Proc. ER-BPM '11. (2012) 370–382
10. Hoppenbrouwers, S., Proper, H., Weide, T.: A fundamental view on the process of conceptual modeling. In: Proc. ER '05. (2005) 128–143
11. 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
12. Morris, W.T.: On the Art of Modeling. *Management Science* **13** (1967) B–707–B–717
13. Willemain, T.R.: Model Formulation: What Experts Think about and When. *Operations Research* **43** (1995) 916–932
14. Newell, A., Simon, H.: *Human problem Solving*. Prentice Hall (1972)
15. Petre, M.: Why Looking Isn't Always Seeing: Readership Skills and Graphical Programming. *Commun. ACM* (1995) 33–44
16. Mendling, J., Reijers, H.A., Cardoso, J.: What Makes Process Models Understandable? In: Proc. BPM '07. (2007) 48–63
17. Pinggera, J., Zugal, S., Weber, B.: Investigating the process of process modeling with cheetah experimental platform. In: Proc. ER-POIS'10. (2010) 13–18
18. MacQueen, J.: Some methods of classification and analysis of multivariate observations. In: Proc. Berkeley Symposium on Math., Stat., and Prob. (1967) 281–297
19. Hamerly, G., Elkan, C.: Alternatives to the k-means algorithm that find better clusterings. In: Proc. CIKM '02. (2002) 600–607
20. Ericsson, K.A., Simon, H.A.: *Protocol analysis: Verbal reports as data*. MIT Press (1993)
21. Reijers, H., Mendling, J.: 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
22. Cardoso, J.: Business process control-flow complexity: Metric, evaluation, and validation. *JWSR* **5** (2008) 49–76
23. Becker, J., Rosemann, M., von Uthmann, C.: Guidelines of business process modeling. In: BPM. Volume 1806 of LNCS. Springer (2000) 241–262
24. Mendling, J., Reijers, H.A., van der Aalst, W.M.P.: Seven process modeling guidelines (7pmg). *Information & Software Technology* **52** (2010) 127–136
25. Frederiks, P., Weide, T.: Information modeling: The process and the required competencies of its participants. *DKE* **58** (2006) 4–20
26. Rittgen, P.: Negotiating Models. In: Proc. CAiSE '07. (2007) 561–573
27. Stirna, J., Persson, A., Sandkuhl, K.: Participative Enterprise Modeling: Experiences and Recommendations. In: Proc. CAiSE '07. (2007) 546–560