CrossMark

ORIGINAL ARTICLE

Evaluation of collaborative modeling processes for knowledge articulation and alignment

Stefan Oppl[1]

Abstract The evaluation of modeling processes for years has focused on assessing the outcome of modeling, while the process of modeling itself hardly has been subject of examination. Only in recent years, with rising interest in collaborative conceptual modeling approaches, the process of modeling has gained more attention. Different streams of research have focused on examining the sequence of model manipulations or have considered the collaborative modeling to be a negotiation process and shaped evaluation criteria accordingly. This article proposes to add the perspective of articulation and alignment of the topic of modeling to the potential foci of evaluation. Its contribution is to show, how existing research originating in theories of co-construction of knowledge can be adapted for this purpose and how the adopted perspective is complementary to those already available in related work. A comparative field study applying different evaluation methods on a single real world showcase is presented to show the practical applicability of the approach and allows to identify potential for the combination of previously unrelated analytical dimensions.

Keywords Collaborative conceptual modeling · Process of modeling · Articulation and alignment · Evaluation methodology

1 Introduction

Examining the process of conceptual modeling is an area of research that has gained momentum in the last years (Claes et al. 2013; Soffer et al. 2012). Research in the area of process management traditionally has taken the existence of appropriate

✉ Stefan Oppl
   stefan.oppl@jku.at

[1] Department of Business Information Systems, Communications Engineering, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria

Springer

models for granted (Hjalmarsson et al. 2015), and has hardly addressed the questions of how models should be created, who should be involved during creation, and how an appropriate representation of the real world could be assessed.

Over the last years, however, interest has risen in researching the facilitation of the process of model elicitation and creation (Hjalmarsson et al. 2015). One notable trend in this area is the active involvement of domain experts in the modeling process (Antunes et al. 2013). This strategy is based on the hypothesis that direct involvement of people with relevant knowledge can help to avoid modelers' bias and consequently lead to models that are more useful for informing actual work processes (Krogstie et al. 2006). End-user involvement has been successfully pursued in the area of requirements engineering (Mullery 1979) and educational research (Mayer 1989) since decades. It also has been deployed in several research efforts in the area of social-technical systems modeling (Herrmann et al. 2000). Approaches following this strategy, however, usually have adopted measures related to the generated outcome to assess their impact and hardly have evaluated the process of modeling itself.

In recent years, it has been recognized that the added value of domain-expert-driven modeling not only is generated via the resulting models, but also by creating common ground for the involved people (Hoppenbrouwers et al. 2005). Research has started to examine how facilitation of modeling processes can support the evolution of common ground (Hoppenbrouwers and Rouwette 2012). In this line of research, several efforts have been made to qualitatively describe the effects occurring during modeling (Rittgen 2007; Seeber et al. 2012). The modeling process is considered as a series of negotiation acts among actors, with the model being an artifact generated as an outcome. Evaluations of the process of modeling consequently focus on depicting and analyzing the observed negotiation acts and their impact on the model. Other approaches focus on the process of model creation and collect their data solely from observing the manipulation of the model (Pinggera et al. 2012; Sedrakyan et al. 2014).

As will be shown in more detail in Sect. 2, these approaches examine different aspects of how the actors negotiate to reach a common understanding about content of the model. Existing research (Gemino and Wand 2003; Krogstie et al. 2006; Mayer 1989) indicates that further potential for evaluating collaborative modeling processes can be found in examining an actor's understanding about a modeled topic and how this its development is facilitated during modeling. These works argue for the need of assessing the development of understanding about the modeled topic, as the understanding of a topic ultimately affects the actions performed in the real world. Several authors have addressed this issue from a cognitive perspective, focusing on the development of understanding on the subject of modeling for individual modelers (Claes et al. 2015; Recker et al. 2013; Soffer et al. 2012). They, however, either focus on examining the modeling process of individual modelers confronted with artificial tasks in an experimental setting (Claes et al. 2015; Soffer et al. 2012) or examine collaborative modeling processes ex-post using question-naires (Recker et al. 2013). The process of collaborative modeling itself not yet has been investigated with respect to how the involved actors articulate their individual views about the modeled topic and how they develop a common understanding.

This article introduces an approach that addresses this gap and considers collaborative modeling as a process of knowledge articulation and alignment. We adapt an evaluation methodology that has been proposed in the area of technology enhanced learning (Weinberger and Fischer 2006), where articulation and alignment has already been adopted as perspective when evaluating collaboration processes. Taking this perspective on evaluation allows to fill the gap identified above. A comparative field study has been conducted for comparing the proposed methodology with other evaluation approaches identified in related work as a baseline to show that it allows to derive information previously not addressed in data analysis. The results of the field study are also used to discuss the implications of using diverse objects of investigation to analyze complementary aspects of a collaborative modeling process.

This article is structured as follows: In the next section, the current state-of-the-art for evaluating collaborative process modeling processes is presented. The paper reviews those approaches regarding the information that can be drawn from their results for assessing modeling techniques and identifies a gap in the current state-of-the-art. It then introduces an evaluation approach, which investigates the process of collaborative modeling from a knowledge articulation and alignment perspective. Subsequently, this approach is reflected against existing approaches by applying them on a real world case and assessing the qualities of the different outcomes. The article closes with the discussion of the implications of these findings and an account on the next steps towards a comprehensive evaluation framework for collaborative modeling processes.

2 Related work

The empirical analysis of conceptual model quality has been subject of research for more than two decades (Sindre et al. 1994). Based on an extensive literature review, Gemino and Wand (2004) derive a framework for empirically evaluating conceptual modeling techniques. Their analysis shows that most approaches use the modeling results as their objects of investigation and omit the process of modeling itself. This diagnosis remains true also in more recent research (Nelson et al. 2011; Claes et al. 2013). In this light, Gemino and Wand (2004) propose to also to assess the effectiveness of the process of model creation and model interpretation. While they identify potential useful variables, they do not address the operationalization of these variables.

This prior research has been picked up in the emerging research field of collaborative modeling in recent years (Mendling et al. 2011; Wilmont et al. 2013). When focusing on collaborative aspects of modeling, the process of model creation and understanding is intertwined, as different actors dynamically switch between active and observational roles during the process of modeling. The observable interaction phenomena thus are an important aspect of evaluation.

The empirical analysis of the process of collaborative modeling has been addressed in various disciplines, from system dynamics (Rouwette et al. 2002) over requirements engineering (Konaté et al. 2013) to process modeling (Rittgen 2007). In the area of collaborative problem solving, existing research has proposed approaches to describe social processes that can be observed during collaborative

work (Weinberger et al. 2007). Those approaches are not limited to collaborative modeling tasks. In the following, we focus on approaches that have been applied in a modeling context.

2.1 Available approaches for evaluating collaborative modeling

Rittgen (2007) proposes a coding scheme to describe negotiation during collaborative modeling sessions. The analysis is based on utterances of the actors make during the modeling session collected using a think-aloud methodology (Van Someren et al. 1994). The utterances are analyzed along "levels of organizational semiotics" (social, pragmatic, semantic, and syntactic level. These levels allow to describe a modeling process regarding the interaction among the modelers from different perspectives—regarding their collaboration (social, pragmatic) as well as how they build a model about their topic of discourse (semantic, syntactic).

This approach has been adopted by Seeber et al. (2012) to investigate collaboration in group settings and develop support for such evaluations in the CoPrA toolset. They use the pragmatic dimension of Rittgen (2007) for data preparation and have developed IT-support to generate high-level analytics of collaboration processes. More specifically, they analyze the collaboration along the flow of modeling, i.e. related to the sequence of the performed actions, and along their distribution among the involved actors. This allows to study actor's behavior in the group, creating a comprehensive picture of the collaboration process.

Taking a similar perspective, Hoppenbrouwers et al. (2005) argue for considering the process of collaborative conceptual modeling to be a dialogue among the involved actors. In contrast to the approaches described above, the object of investigation here is the whole social setting the modeling session is conducted in. Hoppenbrouwers and Rouwette (2012) adopt this approach and use the concept of "focused conceptualizations" [FoCons (Hoppenbrouwers and Wilmont 2010)] to split the modeling process into its different conversational topics. FoCons are analyzed regarding different aspects of the modeling process (e.g. required input, desired outcome, participants, guidance measures, type of abstraction activity, etc.).

Forster et al. (2013) focus on observing the process of model creation. Observations are derived automatically from logs that are created by computer-supported modeling environments. The object of investigation consequently are observable changes in the model created by the actors. They propose different visualizations to indicate activities of modelers. This includes heat-map overlays of the model to indicate areas that have been subject to more frequent changes or timeline-based modeler activity diagrams, which can be used to identify turn-taking in modeling activities. Pinggera et al. (2012) propose to use modeling phase diagrams to visualize the evolution of the model over time, identifying the sequence of different types of modeling activities. In further research, Pinggera et al. (2013) examined, how to use eye movement analysis to investigate the activities of individuals involved in a collaborative modeling process. Recker et al. (2013) propose an approach that focusses on cognition of the actors involved in collaborative modeling and collect data using an ex-post questionnaire (i.e. do not observe the modeling process directly).

In a third strain of research, efforts have been made to describe evaluation approaches on a meta level to make them adaptable to different foci of evaluation. Ssebuggwawo et al. (2013) propose to use the analytical hierarchical process (Saaty 1990) to examine collaborative modeling processes. By implementing their framework, modeling process evaluation can be conducted from different perspectives and using different objects of investigation, such as the modeling language used, the modeling procedure, the produced model and the medium used for modeling (i.e., tool support). This approach thus can be considered an operationalization of the framework proposed by Gemino and Wand (2004). It, however, does not offer new approaches to examining the operative procedure of collaborative modeling itself but refers to methods that have already been discussed above.

2.2 Review of objects of investigation

In an attempt to systematically review the current state-of-the-art, the approaches described above, which directly observe the process of collaborative modeling, have been examined regarding which object(s) of investigation they use to derive claims about aspects of an observed collaborative modeling session (summarized in Fig. 1). All reviewed approaches derive their claims based on a single object of investigation. They differ in where they draw their data from—Rittgen (2007) and Seeber et al. (2012) use the utterances of the actors as their main source of information, Forster et al. (2013) and Pinggera et al. (2012) draw their information from the sequence of model changes, and Hoppenbrouwers and Rouwette (2012) observe the social setting of the modeling session and the interaction among the actors as a whole. These different objects of investigation to the best knowledge of the author so far have not yet explicitly been used in combination to examine collaborative modeling sessions. This leads to the question, whether a combination could provide added value and allow to derive claims that go beyond the currently proposed aspects. We will get back to this question at the end of Sect. 5.

2.3 Review of addressed analytical dimensions

Using the analytical dimensions of an approach as a means for structuring related work furthermore reveals that not all potentially relevant aspects are equally covered by existing approaches. The structure depicted in Fig. 2 has been
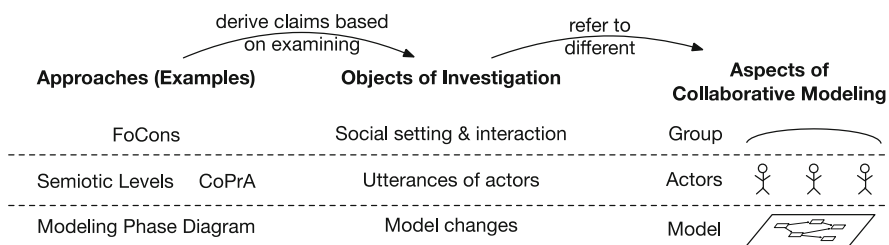

Fig. 1 Different objects of investigation used in reviewed related work
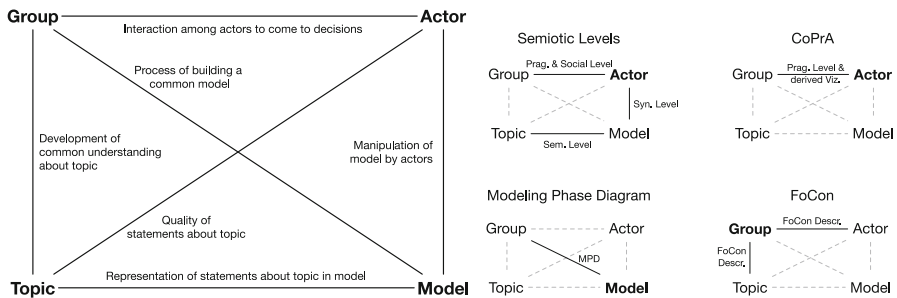
Fig. 2 Analytical dimensions for collaborative modelling sessions (*left*), coverage of analytical dimensions in reviewed related work (*right*)

inductively derived from the reviewed approaches. It was validated and augmented with the framework proposed by Krogstie et al. (2006). Related work analyses collaborative modeling processes along the following dimensions:

- The manipulation of the *model* by the *actors* is analyzed on the syntactic level of Rittgen (2007).
- The interaction among the *actors* to come to decisions as a *group* is analyzed on the pragmatic and social level of Rittgen (2007) and also is addressed in CoPrA (Seeber et al. 2012) in more detail. FoCons (Hoppenbrouwers and Rouwette 2012) take a different perspective on this dimension and analyze the roles of the involved actors.
- The process of building a *model* as a *group* is examined in Modeling Phase Diagrams (Pinggera et al. 2012).
- The way of representing statements about the *topic* to be modeled (i.e. domain knowledge) in the *model* is addressed on the semantic level of Rittgen (2007).
- The development of a common understanding within the *group* about the *topic* of modeling is only explicitly addressed in FoCons (Hoppenbrouwers and Rouwette 2012) by describing guidance measures, observed abstraction activities and used information resources.

The cornerstones of the structure (*model, actors, group, topic*) remain stable throughout all reviewed approaches. This is in line with the main elements of the framework for assessing process model quality presented by Krogstie et al. (2006), which puts particular focus how *actors* could use *modeling* to extend their knowledge about the target domain (i.e. the *topic*) and how *models* could enable them to act within the domain.

2.4 Gap analysis

When reviewing the links between the cornerstones (cf. Fig. 2, right), it is obvious that the link between *actor* and *topic* currently is not addressed in any of the available approaches. Krogstie et al. (2006), however, indicate that this link would

be high relevance when examining modeling processes, as it could be used to analyze the quality of statements actors make about the topic of modeling and how it changes over time, pointing at potential knowledge transfer about the topic among the different actors. This claim also is coherent with Gemino and Wand (2003), who argue for considering modeling techniques as facilitators of learning processes and consequently evaluating them regarding their effects on actors' understanding about the topic.

An analytical approach filling this gap thus has to explicitly consider the quality of the statements articulated by actors about the topic of modeling in a collaborative setting. To the best knowledge of the author, no existing approach in collaborative modeling addresses this requirement. Search has thus been extended to domains that examine social processes in collaborative work on a common topic. An approach potentially filling the identified gap has been proposed by Weinberger and Fischer (2006). Their methodology has been developed to assess argumentative construction of knowledge in collaborative learning settings. Its analytical dimensions can be mapped to the structure proposed above and explicitly address all links between *actor*, *group*, and *topic*. The next section describes how the links are instantiated in the proposed analytical dimensions and how the approach has been adapted to be suitable to analyze collaborative conceptual modeling processes.

3 Evaluating collaborative modeling from an articulation and alignment perspective

The review of related work has shown a lack of support for evaluating collaborative modeling activities from an articulation and alignment perspective. Articulation and alignment here are to be understood as individual and collaborative activities related to externalizing actors' views on a given topic (Strauss 1993) and aligning them to avoid problems during operative collaboration (Vennix et al. 1996). As such, articulation has also been recognized as an important component during the process of conceptual modeling (Krogstie et al. 2006).

One approach that explicitly addresses the quality of articulation and alignment activities in the evaluation of collaborative problem solving processes is proposed by Weinberger and Fischer (2006). It has been successfully deployed and validated in collaborative learning settings and allows to examine how articulate their views about a given topic and how they interact as a group to develop a common understanding. The dimensions they propose for the classification of the actors' articulation and alignment activities are derived from literature, which describes observable phenomena in the process of the collaborative construction of knowledge. Due to original scope of the approach, it does not explicitly consider modeling activities. Still, the proposed dimensions are applicable, as modeling can be considered a sub-class of generic "problem solving" activities (Lesh and Harel 2003). In the following, we review the original analytical dimensions and describe their adaptation for examining collaborative conceptual modeling sessions.

3.1 Object of investigation

The object of investigation in Weinberger and Fischer (2006) are textual statements made by actors in a collaborative (online) environment. The content of the statements is reviewed and classified in the different dimensions described below. In collaborative modeling, the units of analysis are not necessarily bound to verbally articulated statements but can also be constituted by modeling activities. In the following, we thus discuss which object(s) of investigation can be used to appropriately collect data for the application of the approach.

As noted above, some modeling evaluation approaches [e.g. (Pinggera et al. 2012)] use data generated by IT-supported modeling environments to identify their units of analysis. The object of investigation are the models themselves, actors—if at all—are only indirectly considered in the analysis. Approaches focusing on actors' behaviors [e.g. (Rittgen 2007; Seeber et al. 2012)] often use transcriptions of utterances of actors, partially in combination with model changes as their objects of investigation. The transcriptions are either drawn from technologically mediated communication tools (e.g. chats) or retrieved from video recordings of co-located modeling sessions. Approaches targeting to analytically describe the overall social setting [e.g. (Hoppenbrouwers and Rouwette 2012)] use similar approaches. Ssebuggwawo (2012) gives an overview about potential approaches to collect data about collaborative modeling sessions and how to prepare them for analysis.

For the present work, the main objects of investigation are the statements made about the topic by the actors during collaborative modeling. In order for the approach of Weinberger and Fischer (2006) to be applicable, the units of analysis must be segmented to form epistemologically self-contained statements, i.e. refer to a single aspect of the topic. In addition, in the the context of collaborative modeling, acts of modeling, which not necessarily are accompanied by verbally articulated statements, also need to constitute distinct units of analysis. Consequently, units of analysis are separated if one of the following events occur:

- Persons finishing the articulation of a self-contained statement (i.e. a statement that can be interpreted without considering other utterances). Finishing can be identified by turn-taking (i.e. another person taking over) or semantic distinction (i.e. continuing with a statement referring to a different topic). Utterances made by several persons at the same time on the same topic, where no clear semantic distinction can be made, do not constitute distinct units of analysis.
- Persons finishing a manipulation of the model. A manipulation is finished as soon as a person continues with the manipulation of a different area of the model or turn-taking occurs. Simultaneous manipulations do not constitute distinct units of analysis.

The identified units of analysis are classified along different analytical dimensions as described in the following.

3.2 Analytical dimensions

The interaction of the involved actors during collaborative problem solving is assessed in the following dimensions (Weinberger and Fischer 2006):

- Actor participation
- Epistemic nature of statements
- Argumentative quality of claims
- Social modes of co-construction

The classification categories specified in these dimensions are interpreted with respect to their application in collaborative modeling settings in the following. In addition, the present article extends the methodology to also assess manipulations of the model. This allow to put the social interactions among the actors in the context of their modeling activities.

The *participation dimension* refers to the amount of contributions made by the actor. This includes two aspects: the quantity of participation for each actor and the heterogeneity of participation, i.e. the amount of turn taking happening during the modeling process. Participation is not limited to utterances (verbal or written, depending on the source of the analyzed material) but also includes manipulations of the model. During analysis, the actually involved persons are identified for each unit of analysis. Each actor is assigned a unique identifier that allows tracking of the amount participation and involvement in turn-taking for each individual.

The *epistemic dimension* refers to the quality of contributions made in one unit of analysis. Each unit of analysis is classified in a single category. The following scheme is used for classification: An initial distinction is made between on- and off-task statements. Off-task statements comprise all statements which are content-wise not related to the topic of modeling. On-task statements are distinguished based on their content. Following Weinberger and Fischer (2006), statements can refer to: (a) the problem space. Statements in this category refer to the concrete case that is currently articulated or discussed; (b) the conceptual space. Statements in this category refer to generalizations of a concrete case and cover theoretical considerations about the generic aspects of the current issue; (c) the relationships between problem and conceptual space and their adequacy. Statements in this category link case-specific and generic statements. Judging whether the uttered relationship is adequate or inadequate requires a coder to have domain-specific knowledge. As this knowledge not necessarily is available, considering this distinction is optional but allows to conduct a more informed analysis of the modeling process regarding processes of developing an understanding of the topic of modeling; and (d) the relationships between the problem space and prior knowledge. Statements made in this category link case-specific statements to prior knowledge of an actor.

The *argumentative dimension* focusses on contributions to problem inquiry and resolution observable in the units of analysis. In a first analytical step, claims made by the actors are identified. Each unit of analysis either constitutes a non-argumentative move or an argumentative claim. Claims can be qualified or

grounded. Actors explicitly limit the validity of qualified claims validity through describing the context in which the claim is assumed to be valid. Grounded claims are argumentatively backed by the actors through further justifications, which explain why they are assumed to be valid. Claims can also have both qualities, or exhibit neither of them. The latter cases are considered "simple claims". In a second analytical step, according to Weinberger and Fischer (2006), the claims should be analyzed regarding their role in a sequence of arguments. Claims can act as arguments, counterarguments, or integrative statements. In collaborative modeling sessions, the identification of such chains can be challenging due to the dynamic nature of discourse. It thus might not always be feasible to perform this second coding step.

The final dimension of the original approach addresses the *social modes of co-construction*. It classifies the observed discourse with respect to how the actors as a group create align their understanding about the topic and formulate arguments together. Units of analysis that contain content referring to the topic of modeling (as identified in the epistemic dimension) here are distinguished into externalization, elicitation, and consensus-building activities. Externalization refers to units during which actors contributes its own view on the current topic of discourse. Elicitation activities refer to actors questioning others or provoking reactions. Consensus-building can again take different forms. Their identification is described in detail by Weinberger and Fischer (2006) and summarized in the following: In "quick consensus building", contributions of one actor are accepted by the group implicitly or explicitly without any modification and any "indication that the peer perspective has been taken over" (Weinberger and Fischer 2006) by the other learners. Quick consensus-building does not give any indication, if knowledge alignment has taken place. "Integration-oriented consensus building" means that actors take over positions of other actors and extend and validate these positions with own input. A unit rated in this category must show statements that "significantly differ(s) from a juxtaposition of perspectives, but indicates a further development of the analysis" (Weinberger and Fischer 2006) by an actor. "Conflict-oriented consensus building" is characterized by actors, who not accept contributions of others as they are, but challenge. They require adaptation of the articulated positions in order to achieve a common understanding. Units that should be rated in this category are indicated by "rejection, exclusion or negative evaluation of peer contributions" (Weinberger and Fischer 2006), either explicitly or implicitly by ignorance or replacement of a contribution.

The *modeling dimension* describes model manipulations performed by the actors. These manipulations can take different forms, which are informed by those described by Rittgen (2007) for the syntactic level of modeling analysis: (a) adding elements to the model, (b) changing the layout of the model (i.e. rearranging elements), (d) merging duplicate modeling elements or removing them (which is common, when actors contribute individually prepared model elements to a shared model). Further model manipulations can be added to the available analytical categories in this dimension, if they are considered relevant for the applied modeling methodology at hand (e.g., an analysis of the generation of BPMN models

Fig. 3 Coverage of analytical dimensions in the presented approach
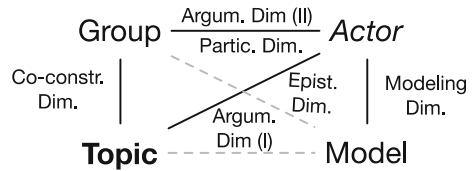
might benefit from an identification of model manipulations during which modifiers are added to already existing model elements).

3.3 Summary

Figure 3 shows the analytical dimensions of the proposed approach embedded in the structure used to review related work. The main objects of investigation are the statements on the topic of modeling made by the actors. Due to the extension of the original approach to also consider modeling activities, the actors have to be included as secondary objects of investigation mainly with respect to their manipulations of the conceptual model. The main contribution of the proposed approach is the explicit analysis of the quality of statements made by the actors about the topic (link between *actor* and *topic*), which is covered by the epistemic dimension. Statements on the topic are classified there regarding their point of reference. The same aspect is addressed in the first step of the analysis in the argumentative dimension, in which claims about the topic are classified regarding how they are embedded in the overall topic. In addition, the co-construction dimension enables a more detailed, fine-grain analysis of how the group works on the development of a common understanding about the topic than is enabled by using FoCons. The second part of the argumentative dimension as well as the participation dimension on the link between *group* and *actor* enables an analysis similar to that proposed in CoPrA. The modeling dimension is derived from the approach of Rittgen (2007) and consequently leads to similar results.

The contribution of the approach introduced in this article has now been outlined on a conceptual level. Its practical added value is demonstrated in the next section by conducting a comparative review of the present approach with those presented in related work based on a real-world case.

4 Comparative review of evaluation approaches

The aim of the comparative review is to contrast the evaluation results of the present methodology with those achievable with related approaches. It demonstrates that the analytical dimensions are complementary to those already proposed in related work and shows the potential value that can be generated by combining those dimensions.

Methodologically, a real world collaborative modeling session is used as a sample case. This case has been selected based on the exposed heterogeneity of interaction among the actors and the different model manipulation activities that are

contained. The case has been used consistently as the subject of analysis for the different evaluation approaches. This allows to compare the results and discuss the different qualities of the generated data. For reasons of diversity, one approach of each of the strains of research identified in the related work section has been selected for the following comparison. We start with the coding approach proposed by Rittgen (2007), where coding is structured along different semiotic levels. The next section describes the application of the CoPrA approach (Seeber et al. 2012). As a representative of model-centric evaluation approaches, the modeling phase diagram (Pinggera et al. 2012) is presented in the next section. We continue with describing the results of the methodology introduced here. Finally, we analyze the sample case along its conversational structure as proposed by Hoppenbrouwers and Rouwette (2012). In the subsequent section, we discuss the results of these codings regarding conceptual overlaps and potentially complementary dimensions. We finally give an account on the implications of adopting an articulation and alignment perspective when evaluating collaborative modeling sessions.

4.1 Sample case

The sample case used for comparing the different evaluation approaches has been selected from a series of modeling workshops aiming at the facilitation of creating a shared understanding about a work process for inexperienced modelers. The modeling methodology is inherently cooperative and relies on a combination of articulation, elicitation, and negotiation to ultimately reach common ground on the way the work process should be implemented. It is described in detail in (Oppl 2016a). The case was conducted in a school for vocational training in healthcare in Germany and dealt with the process of organizing one's practical placement to gain professional experiences. The modeling session reviewed here is an excerpt of the whole workshop and lasted 31 min and 9 s. It was preceded by an introduction to the modeling methodology and a phase of individual reflection on each actor's role within this process. It was followed by a plenary discussion on the implications of the generated models. During the collaborative modeling session, a total of 9 actors actively contributed to the modeling process. 8 of them were female, 1 of them was male. Their age ranged between 18 and 42. None of them had prior experiences in any form of conceptual modeling. The process was facilitated by a teacher of the vocational training school, who had participated in a facilitation training offered by the developers of the modeling methodology.

The modeling result of the evaluated modeling session is shown in Fig. 4 to provide context to the following discussions. Modeling follows an interaction-based paradigm of describing collaborative work processes. Blue elements represent the involved process participants, red elements represent the activities of these participants (with causality indicated by the vertical order of the elements), and yellow elements represent acts of interaction or communication between the connected participant lanes. The modeling language is described in more detail in Oppl (2016b).
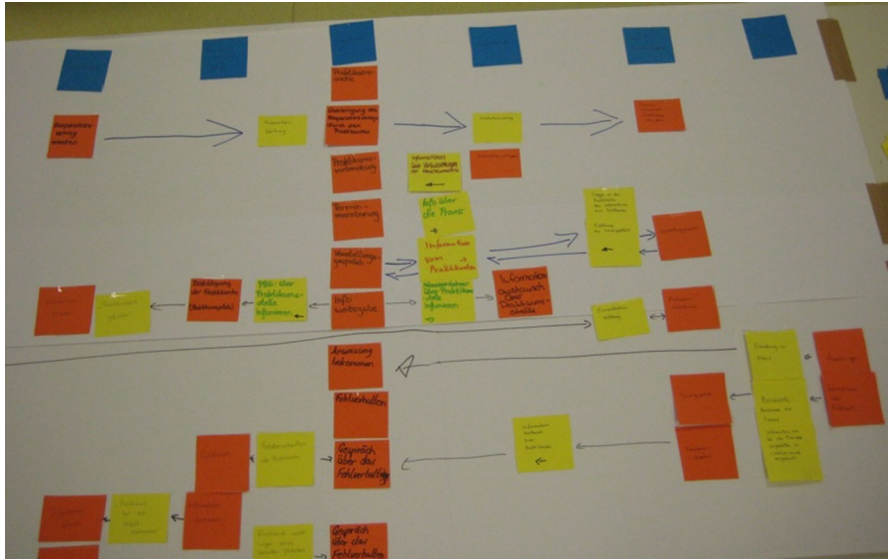
Fig. 4 Modeling result of sample case

4.2 Identification of units of analysis

In order to analyze the modeling process, the units of analysis need to be identified as described above. The modeling session was recorded on video in order to enable the creation of a concurrent verbal protocol (Trickett and Trafton 2009). Instead of using think-aloud to further enrich collected data [as proposed by Trickett and Trafton (2009) and also adopted by Rittgen (2007)], the sample case was recorded without intervention, only collecting statements uttered in the context of collaborative modeling, but including the modeling activities of the actors. Segmentation was performed by two independent raters and consolidated by the coordinator of the study. Overall, 117 distinct units of analysis were identified. Evaluation was conducted independently by two raters to avoid rater's bias following the procedure proposed by Trickett and Trafton (2009). Raters have been trained using a five-minute video sequence of the same modeling method. They were provided with a code-book based on the descriptions of the evaluation methods. The inter-rater reliability was assessed using Cohen's Kappa. The test coding conducted after the first training led to a value for Cohen's Kappa of 0.506. After a revision and clarification of the code-book and another training session, another sample video was coded, reaching a value of 0.932 for Cohen's Kappa.

In the following, the different evaluation approaches are presented in a diagrammatical format that allows to compare their results (cf. Figs. 5, 10). The coding results for each result are presented in a time-proportionally scaled visualization of the units of analysis on the x-axis. Coding categories are stacked on the y-axis clustered along the addressed dimensions. In addition, derived visualizations of the data are provided, if they are described in the source articles.

Fig. 5 Coding results for semiotic levels (*left*) and CoPrA (*right*)
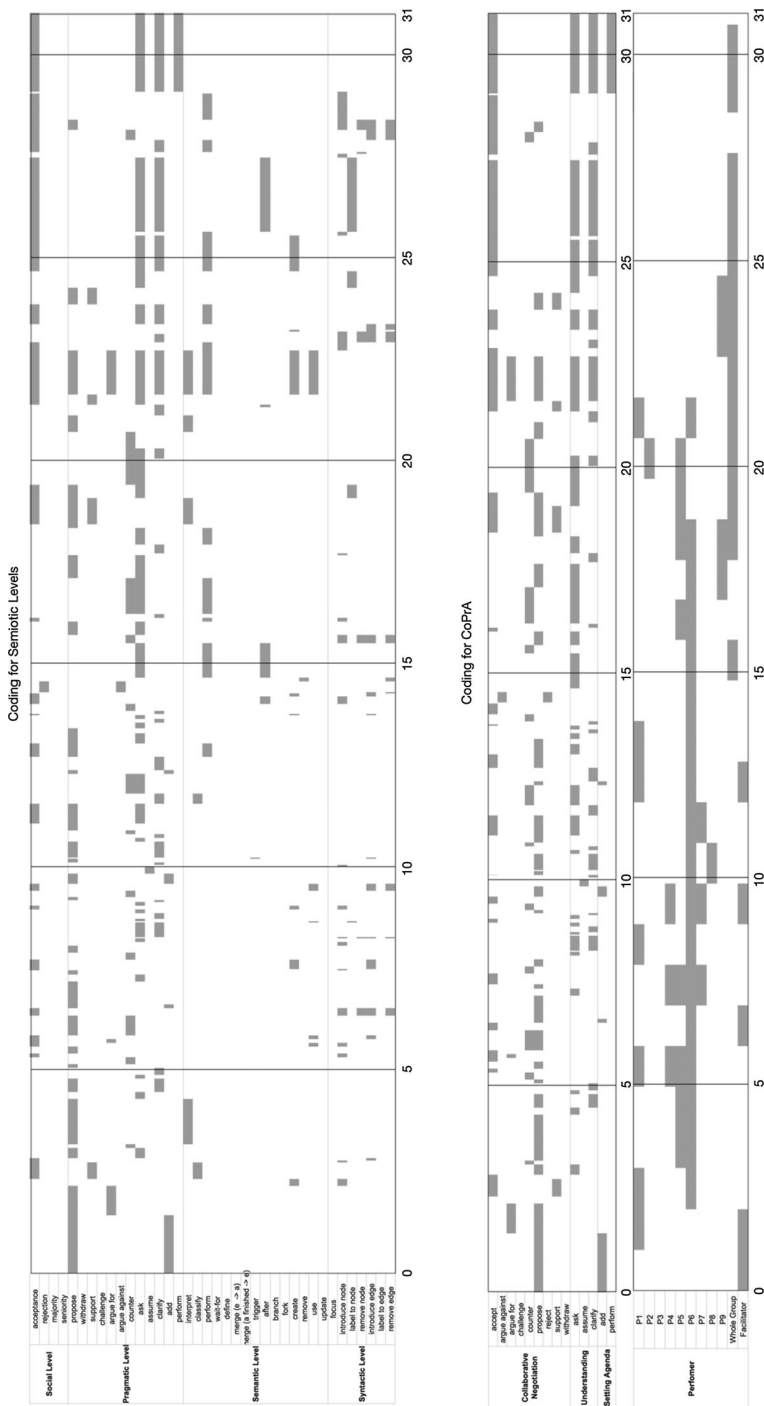
Based on these visualizations, we discuss potential conclusions about the sample case that can be drawn from the coding results. In addition, we provide a description of the limitations encountered during coding.

4.3 Coding structured along semiotic levels

Coding was carried out based on the method description provided by Rittgen (2007).

4.3.1 Analytic dimensions

Collaborative modeling session are analyzed along four different dimensions. The *syntactic level* refers to manipulations of the model, in particular adding, removing, or altering nodes or edges. The *semantic level* describes statements referring to the content of model, i.e. statements that are used to describe the subject that should be depicted in the model. The *pragmatic level* describes the interaction of the modelers during the process of model manipulation. This includes proposals and questions as well as negotiation-related activities such as supporting statements or objecting them. The *social level* refers to observations on the behavior of the actors when making decisions about proposals. While decision making can take different forms, Rittgen (2007) explicitly mentions rules of majority and rules of seniority as prominent examples for this dimension.

4.3.2 Coding result

Applying these dimensions to the sample case leads to the results visualized in Fig. 5 (left). Coding on syntactic and semantic level has been affected by the modeling methodology used in the sample case. The paper-based modeling methodology does not separate the process of adding an element to the model from labeling it. The activity type "labeling of nodes and edges" thus is only used, when labels of already exiting nodes or connections are changed. The limited expressiveness of the modeling language used in the sample case renders some of the categories on semantic level redundant. As the language is tailored to describing case-based models, forks or branches cannot be observed.

4.3.3 Observations regarding the sample case

The visualization in Fig. 5 (left) makes visible four prominent phases of the modeling process. Until approximately minute 5, low amounts of modeling activities are observable. The focus of interaction clearly is put on making proposals about the content of the model. This is followed by a second phase of intense and dynamic interaction that lasts from minute 5 to minute 16. This phase is characterized by sequences of asking questions and providing clarifications about the modeling subject, with frequent additions of nodes and edges to the model. From minute 16 to 23, hardly any manipulations of the model can be observed. This is accompanied by a less dynamic sequence of questioning the content of the model in the pragmatic dimension. Starting from minute 23 until the end of the observed

session, the coding shows changes being made to existing elements in the model. These changes are accompanied by a large amount of clarifications (visible in the pragmatic dimension), which are accepted by the group without formally entering a negotiation process (visible in the social dimension).

4.3.4 Limitations on coding and interpretation

Analysis of the semantic level requires to have a detailed view on both, the content of the model and how it is changes as well as the interaction of the modelers. Co-located modeling sessions without technical support, such as the sample case, often suffer from restrictions on data collection (e.g. the opportunity to use multiple cameras). Analysis of this dimension thus can hard to be carried out unambiguously.

The syntactic level does not account for changes to the model layout. While layout changes in general do not change the syntax of a model, the sample case shows that the spatial arrangement of modeling elements is used to reflect pending proposals that still need to be agreed upon. Considering layout changes thus could add data relevant for overall analysis of the modeling process.

The method also does not allow to code off-topic interventions such as explanations about the modeling methodology provided by a facilitator. As such interventions can have significant impact on the modeling and negotiation process, their identification could be of interest for the analysis.

The method does not allow to describe simultaneous interaction of multiple actors. The co-located modeling setup, however, facilitates such behavior, which cannot be appropriately represented in the coding scheme.

4.4 CoPrA

CoPrA (Seeber et al. 2012), in its original implementation, is implemented in a tool, which supports coding of pre-segmented modeling sessions and allows to carry out higher level analytics of the modeling process automatically (e.g. in terms of time distributions for different modeling activities, etc.). In order to allow for consistent comparison, the tool has not been used here.

4.4.1 Analytic dimensions

CoPrA does not explicitly focus on collaborative modeling processes, but aims at analyzing collaboration processes in general. CoPrA adopts the work of Rittgen (2007) for analysis of collaboration. CoPrA clusters the categories of the pragmatic level in (a) those aiming at setting the agenda of the collaborative process, (b) those aiming at understanding the topic at hand, and (c) those aiming at collaborative negotiation processes. In contrast to Rittgen (2007), CoPrA analyzes the individual contributions of the actors during the collaboration process. Participation of an individual actor is a binary category for each unit of analysis and is not categorized any further.

Higher level analytics is generated based on this coding. Analytics include distribution of action types [based on the categories specified by Rittgen (2007)] and

distribution of contributions by actor. Both distributions are proposed to be visualized as pie charts, showing relative amounts, and as time-line based diagrams, showing the absolute amount of observed occurrences in a given time-interval.

4.4.2 Coding result

Applying CoPrA to the sample case leads to the observations visualized in Fig. 5 (right). The upper coding table resembles the pragmatic dimension of the coding presented for the methodology of Rittgen (2007). The lower coding table indicates participation for each of the nine actors. As an addition to the original methodology, participation of the facilitator has been added as well as participation of the whole group. The latter has been used when the dynamics of interaction did not allow to identify single active actor.

Based on this coding, higher level analytics were generated. Figure 6 shows the distribution of activity types overall and separated by the categories identified above as pie charts.

Figure 7 gives an overview about the relative amount of participation for all actors as a pie chart.
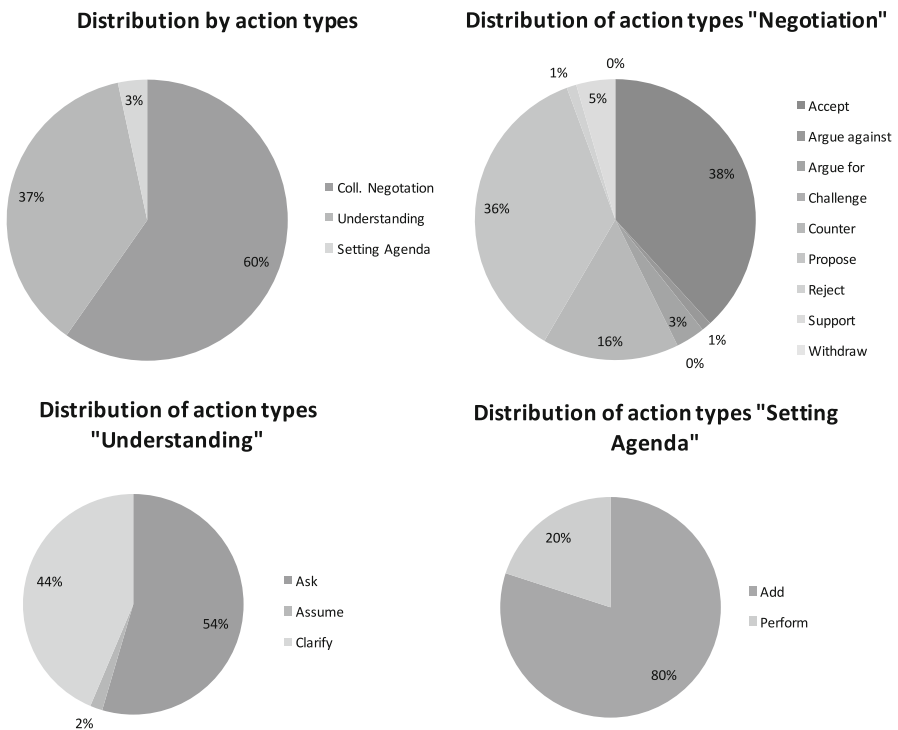

Fig. 6 Distribution of activity types in sample case

Distribution of activity by performer

Pie chart segments: 11% (P1), 1%, 2%, 9%, 10%, 33%, 7%, 5%, 9%, 9%, 4%

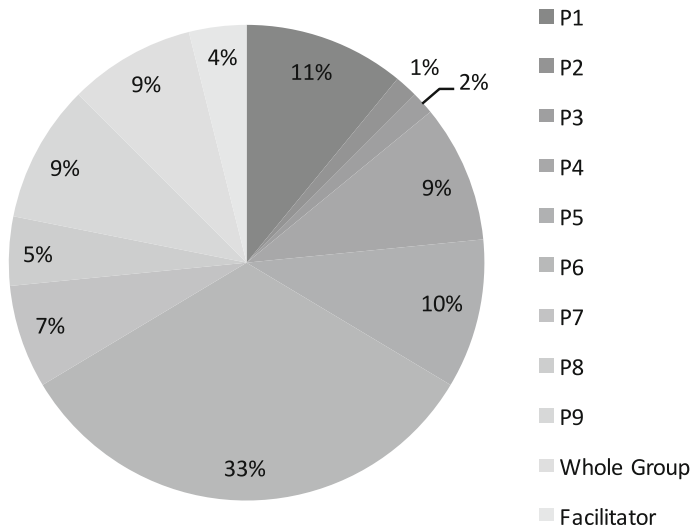Legend: P1, P2, P3, P4, P5, P6, P7, P8, P9, Whole Group, Facilitator

Fig. 7 Relative amount of participation per actor in sample case

Figure 8 shows examples for the timeline-based analysis of the distribution of action types (above) and amount of participation (below). The interval chosen for visualization has been 1 min. The stacked column chart on the left is a form of visualization chosen in the present article, whereas the line chart resembles the originally proposed form of visualization. These different forms have been chosen to discuss differences in potential observations below.

4.4.3 Observations regarding the sample case

As the fundamental coding scheme closely resembles the coding following the pragmatic level of Rittgen's (2007) scheme, the findings described above also apply to the methodology discussed in this section. CoPrA is not tailored to the analysis of collaborative modeling processes in particular, and thus lacks analytical dimensions reflecting model manipulation.

The coding of individual actors' contributions gives a more comprehensive picture of the interactions among the actors. Specifically, it shows a very heterogeneous amount of participation of the involved actors. P6 is responsible for one-third of the activities observed in the modeling process, whereas P2 and P3 only have contributed to <2 % of the observed activities. 9 % of the activities could not be assigned to specific persons and are considered activities of the whole group. When reviewing the timeline-based visualization of participation, the phase of dynamic interaction between minutes 5 and 15 identified in the former section are clearly visible here, too. They show that most of these activities can be attributed to
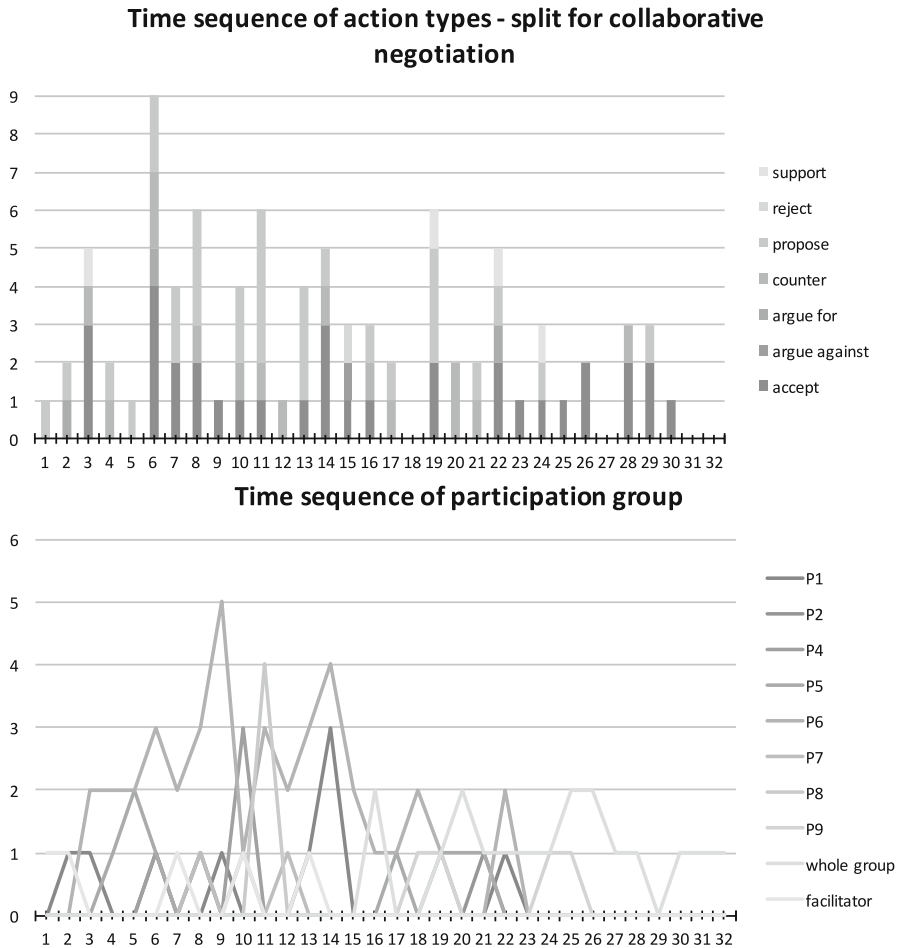
Fig. 8 Timeline-based visualization of action types (*above*) and participation (*below*)

P6 interacting with P4 and P8. The final phase of modeling shows a large amount of interaction in the whole group that could not be assigned to a specific person.

The visualization of the distribution of activity types show that 60 % of the observed activities were dedicated to negotiation activities and 37 % were concerned with activities dedicated to creating mutual understanding about the topic. Only 26 % of the negotiation activities were concerned with actual discussion activities, whereas 36 % were proposals, which often were accepted without further discussion (38 %). The timeline-based visualization shows that the largest share of discussion-oriented negotiation activities is again located in the timespan between minutes 6 and 15.

4.4.4 Limitations on coding and interpretation

CoPrA focusses on analyzing interaction in collaboration processes and does not consider model manipulation activities. Data capturing consequently can focus on

the interaction of the actors. The data needs to allow the identification of the contributor of each statement. The challenge of handling dynamic interaction situations with inseparable individual contributions has been addressed by explicitly introducing a coding category for the whole group.

The higher-level analytics proposed in CoPrA can be derived solely from the coding results as described above and do not require additional metadata. More complex analytics, such as the HeuristicNet proposed by Seeber et al. (2012) to visualize the probability of activity sequences, rely on advanced data mining techniques and thus require dedicated tool support, as available in the CoPrA toolset.

4.5 Modeling phase diagrams

Pinggera et al. (2012) introduce an analytical method for the process of process modeling. This process not necessarily is collaborative. The method explicitly also can be applied for individually created models. While the authors demonstrate the applicability of their approach for the area of process modeling, it can be generalized for analyzing conceptual modeling processes in general. They consider the process of modeling to be an iterative process which exposes different phases of activities.

4.5.1 Analytic dimensions

Pinggera et al. (2012) propose to split the modeling process in distinct phases based on the observable types of activities. They identify three fundamental types of activities: During *comprehension*, the group tries to understand the topic to be modeled and/or the content already represented in the model. During *modeling*, this understanding is used to manipulate the model, i.e. adding or removing nodes and edges or otherwise changing the structure of the model. *Reconciliation* phases are identified, when the model is reorganized to enhance its understandability. This includes renaming of nodes and edges as well as altering the layout of the model.

These phases are identified by interpreting the observable modifications of the model. Similarly to the syntactic dimension proposed by Rittgen (2007), the authors propose to identify each act of model manipulation, distinguishing between structural changes and reorganization of an existing model. Sequences of structural model changes are classified as *modeling* phases, whereas sequences of model reorganization activities are classified as *reconciliations* phases. Sequences with no observable modeling activities are classified as *comprehension* phases.

The classification results are visualized using a timeline-based diagram. The number of model elements is displayed over time, showing the evolution of the model size during the process of modeling. Each line segment (i.e. each phase identified above) is displayed with a different stroke type to enable the distinction of the phases.

4.5.2 Coding result

Figure 10 (left) shows the coding performed according to the criteria identified above. Modeler interactions have been identified based on the observable model
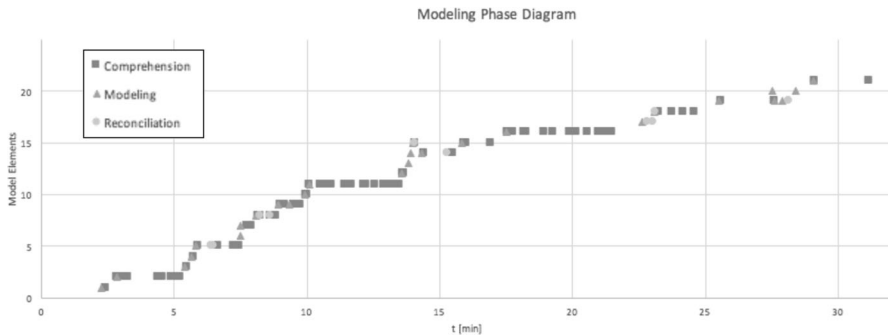
Fig. 9 Modeling phase diagram

manipulations. The modeling phase diagram shown in Fig. 9 has been derived from this coding. In contrast to the originally proposed form of visualization, the diagram has been augmented with information on turn-taking, which is relevant for collaborative modeling processes. Visualization is based on the units of analysis identified in the source data. The start of each segment has been represented with a data point in the diagram, where its shape indicates the type of phase it belongs to. Regions with a high density of dots consequently indicate a large amount of turn taking, whereas gaps in the diagram indicate longer lasting activities of a single contributor.

4.5.3 Observations regarding the sample case

Both visualizations (Figs. 9, 10, left) indicate that the largest share of modeling time is dedicated to comprehension activities, which are not directly reflected in the model. Model manipulations occur in batches, which are interrupted by longer phases of comprehension. The most dynamic phase of the modeling process can be found between minutes 5 and 10, which is indicated by a quickly rising modeling size and a heterogeneous sequence of rather short modeling, reconciliation and comprehension activities. Reductions in model size only occur twice and in both cases can be accounted to the removal of edges in the model (as identifiable in the coding table in Fig. 10, left). Other than that, the model is constantly increasing in size. The later phases of modeling are characterized by longer comprehension phases without any observable turn taking.

4.5.4 Limitations on coding and interpretation

In contrast to the former methods, the application of this method only requires the availability of data about model manipulations. As the method does not take into account whether modeling is an individual or a collaborative act, data about the individual actors are not required. The algorithm specified by Pinggera et al. (2012) allows to identify phases solely based on model manipulations.

Fig. 10 Coding result for modeling phase diagram (*left*) and co-construction of knowledge (*right*)
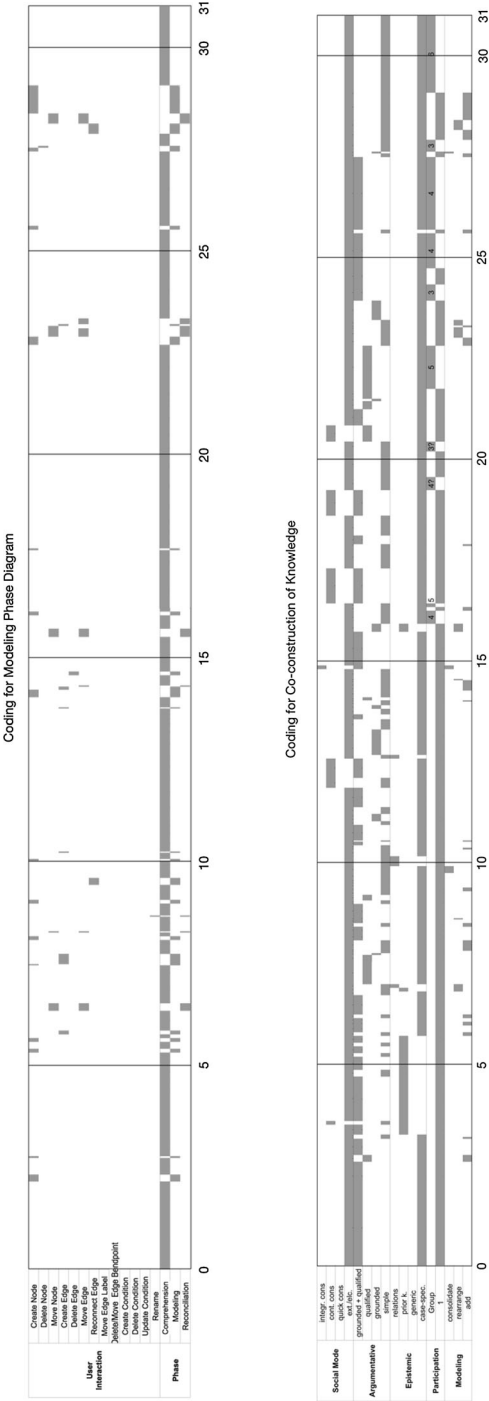
The chosen object of investigation also appears to lead to a major limitation of the approach, especially when applied in collaborative settings. No observable modeling activity not necessarily means that comprehension activities take place. Simple breaks or off-topic interactions between actors cannot be identified as such, when solely relying on data generated from observed model manipulations.

4.6 Coding structured along articulation and alignment of knowledge

The different dimensions of the method introduced here are informed by concepts to describe discourse happening in the context of knowledge articulation and alignment, in particular the concept of creating common ground among a group of actors (Clark and Brennan 1991).

4.6.1 Analytic dimensions

The analytic dimensions of the framework have already been described in Sect. 3. Summarizing, collaborative modeling sessions are analyzed regarding the amount of *participation* by the involved individuals, the *epistemic* quality of the statements contributed by the actors, the *argumentative* quality of claims uttered in the course of collaboration, and the observable *social modes of co-construction*. The original framework of Weinberger and Fischer (2006) is augmented with an additional dimensions reflecting *model manipulations*.

4.6.2 Coding result

Figure 10 (right) shows the coding result for the sample case based on the dimensions introduced above. The participation dimension has been augmented with information on how many actors were active in units of analysis showing simultaneous activities.

4.6.3 Observations regarding the sample case

Reviewing the patterns in Fig. 10 (right) reveals 4 phases in the modeling process: Until minute 6, hardly any modeling activities took place. This phase is largely dedicated to articulation of individual views on the work process to be modeled. At minute 3, a brief sequence of controversial statements (cf. social mode) led to a switch from discussing the case at hand to referring to prior knowledge (cf. epistemic dimension). The switch back to discussing the actual work process at minute 6 also marks the end of the first phase. It is followed by a phase of intense modeling activities, which lasts until minute 11. Here, individuals contribute their views without being fundamentally questioned. From minute 12 to 21, a phase characterized by controversial consensus building activities can be recognized (cf. social mode). The participation dimension indicates dynamic interactions, as group-level interactions start to occur more often. Starting from minute 22 until the end of the modeling session, a large amount of group-level interactions is observable (cf. participation dimension), which is accompanied by frequent modeling activities, in

particular layout changes (cf. modeling dimension). Claims are better grounded than in the preceding phases, indicating a more careful line of argumentation for the final changes to the model (cf. argumentative dimension).

Overall, the "social modes of co-construction" show a large amount of the modeling time being dedicated to elicitation and externalization activities. Explicit consensus building activities could hardly be observed. The epistemic dimension shows that most of the discussions referred to the concrete work case being discussed. No attempts to generalize the findings to a more abstract level were made. This, however, is in line with the aims of the modeling technique, which explicitly facilitates the construction of case-based models.

4.6.4 Limitations on coding and interpretation

As the proposed evaluation methodology requires to observe statements of the actors regarding the topic of modeling and model manipulations at the same time, it faces similar challenges as the coding based on semiotic levels as proposed by Rittgen (2007). Still, as it does not describe activities related to model content in detail, data collection not necessarily needs to include details on the created model.

4.7 FoCon-based analysis

For FoCon-based analysis (Hoppenbrouwers and Rouwette 2012), source data needs to be split along discernible topics of discourse in a modeling process. Each FoCon is described with respect to different aspects of both, model building and interaction, which are described in the following.

4.7.1 Analytic dimensions

The identification of FoCons is based on changing topics of discourse. Hoppenbrouwers and Wilmont (2010) state that the granularity of segmentation of FoCons is dependent on the level of abstraction one takes when analyzing a modeling session. A single FoCon could be used to describe a modeling session as a whole on an abstract level, whereas FoCons also can be used to decompose a modeling session into different segments with a more constrained focus of discourse.

Each FoCon is described using the following information categories (Hoppenbrouwers and Wilmont 2010): *input* ("what may or must go in") describes required information and sources thereof; *outcome* ("what should come out") describes the desired syntactic, semantic and pragmatic results of the modeling session; *abstraction activity* describes whether model creation has been performed via generation, classification, selection, or any combination thereof; *focus questions* describe how modeling was focused pragmatically or semantically using specific questions; *involved participants* indicate which persons or types of persons with specific skills or knowledge were involved in the FoCon; *instructions* describe the explicit or implicit instructions, guidance measures, procedures or conventions that have been provided to the involved participants; and *context* describes further

situational aspects and constraints influencing the FoCon, such as used media, availability of resources, social issues, etc.

4.7.2 Coding result

For the sample case, five FoCons have been identified. Each of these FoCons deals with a different topic of organizing and implementing a student's practical training in a healthcare institution:

1. Supervision procedure and cooperation with school
2. Reflection on mistakes during therapy
3. Documentation of and feedback on training experiences
4. Formal requirements on confirmation of training by hosting institution
5. Requirements on patient information process

The aims of the FoCons in general have been to facilitate common ground about the topics of discourse, and to create model representations of the involved people, their tasks and their interactions. The textual description of FoCons is too extensive to be comprehensively described here for the whole case. Two FoCons thus have been selected for demonstrating their qualities in the following:

FoCon 2: Reflection on mistakes during therapy

- Timespan (based on video timestamp): 03:50–10:00.
- Input: Knowledge about permitted therapeutic activities, awareness of mistakes that can happen because of negligence or misconduct.
- Outcome: Awareness of the need to openly communicate mistakes during therapy and how to document them. The process of documenting mistakes and the circumstances that led to their occurrence.
- Abstraction Activity: Generation (of proposals for communicating and documenting therapeutic mistakes), selection (of actually implementable ways of communication and documentation)
- Focus Questions: Why is an open communication about therapeutic mistakes important in professional practice? How can mistakes be communicated in an open and constructive way?
- Involved Participants: Students (having no practical experiences in therapeutics and collaborative care taking), teacher (striving to convey the importance of openly communicating mistakes that might happen during therapeutic activities).
- Instructions: Identify the relevant communication and documentation tasks in case of a mistake. Represent those using the modeling language constructs and layout guidelines introduced earlier in the workshop.
- Context: This FoCon has been characterized by a dynamic discussion in its second half on in which form mistakes could and should be communicated to colleagues and team leaders.

FoCon 5: Requirements on patient information process

- Timespan (based on video timestamp): 20:00–31:09.
- Input: Assumptions and knowledge about legal and social requirements on a patient information process distributed across the involved participants.
- Outcome: The desired outcome was to generate a fundamental understanding about the requirements on a patient information process. This process was represented as an interaction-oriented model of the work process.
- Abstraction Activity: Generation (of proposals for potentially relevant tasks and information items), selection (of actually necessary tasks and information items).
- Focus Questions: How should a patient information process be implemented? What are the legally required information to be passed on to the patient? Which social factors should be taken into consideration when interacting the patient?
- Involved Participants: Students (having no knowledge about legal requirements and regulations and no experiences in handling patients), teacher (aware of all legal requirements and regulations), supervising professional (experienced in conducting patient information processes).
- Instructions: Identify the relevant actors, tasks and information to be passed to the student. Represent those using the modeling language constructs and layout guidelines introduced earlier in the workshop.
- Context: Due to the FoCon covering the last in a series of topics discussed in the modeling session, concentration of the participants already was vanishing and discipline to adhere to the modeling guidelines was below optimum.

4.7.3 Observations regarding the sample case

The documented FoCons provide a rich picture of the content of the collaborative modeling session. The FoCons show that five distinct topics have been discussed during the modeling sessions. They also expose that the actors involved in the modeling session had different roles with respect to their knowledge about the topic to be modeled. In particular, the group comprised students of a healthcare profession, a teacher and an experienced professional. The instruction category shows that the process was supported procedurally by a facilitator and structurally by guidelines encoded in the modeling artifacts. The categories for input, outcome and focus questions differ between the FoCons due to their different semantic foci. In general, they show that the FoCons were specified targeting issues that might arise during the practical training of the students. Abstraction activities largely could be classified as generation and selection activities, whereas classification was rarely observed. This indicates that the actors were largely concerned with articulating proposals (generation) and discussing their validity (selection), whereas striving towards a more generic level of description of tasks during a practical training (classification) was not observed. The context dimension shows that the modeling process was affected by social effects among the actors (in particular in

phases of selection) and the large number of topics covered in the session (which led to concentration issues in the final FoCons).

4.7.4 Limitations on coding and interpretation

FoCon-based coding requires to be able to exactly follow the content of the discourse among the actors. Model manipulation itself is not explicitly considered in the method. It is thus sufficient to be able to identify and content-wise understand the contributions of the actors involved in the modeling process.

4.8 Analysis based on complementary dimensions

Reviewing the coding results and interpretations shows that they partially offer complementary perspectives for analysis. The complementary aspects raise the question, whether value also can be generated from combining the perspectives of the different approaches. This section discusses findings that emerge from such combinations for the sample case on two examples.

The FoCon-based coding is the only approach that makes visible the different topics of discourse during the observed modeling session. Using the FoCons as an overlay for the timeline-based coding of the other approaches allows to identify the topics that were controversially discussed or have shown dynamic interaction patterns.

In combining FoCons with other approaches that consider the model creation dimension, it becomes obvious that not all topics of discourse were equally represented in the model. The modeling phase diagram makes visible five phases of substantial model manipulation. While those are basically aligned with the FoCons, only FoCons 2 (from 3:00 to 10:00) and FoCon 5 (from 20:00 to 31:09) show modeling activities that are spread over a longer duration in time. In addition, the CoPrA-based coding shows heterogenous participation of a large number of individuals or the whole group during these FoCons. The social modes of co-construction, however, do not show controversial consensus-building in these FoCons but rather elicitation and externalization activities. This indicates that a large number of actors have contributed their individual views during these FoCons not only in discourse but also by manipulating the model. Finding a common understanding, however, does not appear to have been difficult, but rather has been a unanimous combination of different viewpoints.

The combination of the pragmatic level coding used by Rittgen (2007) and CoPrA with the argumentative dimension introduced in the present article allows to identify, whether different kinds of contributions to the negotiation process show different qualities in terms of to which amount the contributions are grounded and qualified in the context of the modeling topic. Figure 11 shows the argumentative quality of "proposals" made during the observed modeling session. It shows that proposals were rarely made without any grounding or without specifying the scope of validity of the proposal (i.e. without "qualifying the proposal"). This indicates that contributions were well argued for, which in turn did not lead to major
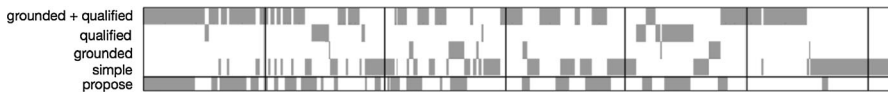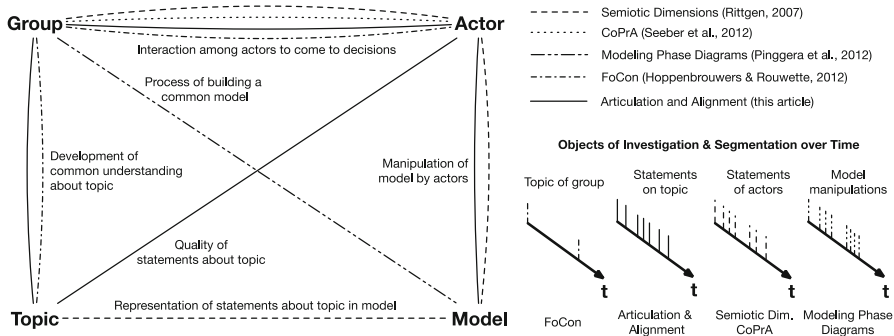
Fig. 11 Example of combined analytical dimensions

Fig. 12 Complementarity of approaches for evaluation of collaborative modelling processes

objections. Even during controversies (as visualized in the social modes of co-construction), the argumentative quality of proposal remained high (Fig. 11).

Overall, the combination of the topic-oriented coding using FoCons with the activity-oriented segment-based coding adopted in the other approaches allows to draw much information from the raw coding results. Also, the combination of different perspectives used by the other approaches shows potential for deriving conclusions previously not identifiable when applying a single approach.

5 Implications

The implications of the evaluation results in the light of the article's objectives are discussed in the following. In showing the complementarity of the approaches, we outline the implications of considering articulation and alignment of actors' views about the topic of modeling as an aspect relevant for evaluation. Finally, we argue to work towards a combination of the of the approaches to allow for more comprehensive evaluation designs.

As has been shown in the study of related work, the five approaches reviewed above address different aspects of the modeling process. The evaluation of the sample case shows that each of the approach provides insights into features of a collaborative modeling process, which are not obtainable by any of the other approaches. In addition, the different granularities of analysis enable to describe the process of modeling on either micro- or macro-level. Figure 12 summarizes the different characteristics of the approaches using the structure developed in Sect. 2.

For each of the examined aspects (i.e. the links between *actor, group, model,* and *topic* in Fig. 12), the analytical dimensions of the reviewed approaches and their complementarity are discussed in the following:

Model Manipulation by actors is described by Rittgen (2007) on a detailed level by distinguishing manipulations affecting model nodes and edges as well as the type of observable manipulation (e.g., add, remove, change). The approach presented in this article analyses model manipulations on a more aggregated level, only distinguishing the type of manipulation. The latter appears to be appropriate in cases, where a focused overview about the evaluation of the model is required for analysis. Still, this level of abstraction can be derived from the results of Rittgen (2007).

The *interaction among actors to come to a decision* is the aspect addressed by the largest number of approaches in different dimensions. The participation of actors in group interaction is analyzed in CoPrA by considering the distribution of *individual contributions* of each actor. The active actors are identified, generating a detailed view on their interactions. The approach presented here uses an aggregated view and distinguishes activities by a single actor from those of a group. For analysis, CoPrA relies on splitting the modeling session in time segments of equal length, rather than the activity-based segmentation used here. Thus, both approaches show different aspects of individual contribution to modeling and augment each other. FoCon-based analysis also uses a similar observational focus, but only describes participation on FoCon-granularity. In contrast to the other approaches, it does not focus on the individual activities, but describes the roles, skills and knowledge of the active participants. Rittgen (2007) and CoPrA go beyond participation, and also analyze the statements actors utter to negotiate the models' content and its creation process. CoPrA here builds upon the data generated by Rittgen (2007) and further derives aggregated views and visualizations of activity distributions.

The *development of a common understanding about the topic* is addressed on different levels of granularity by two approaches. FoCon-based analysis content-wise captures the required input and desired outcome of each FoCon and also identifies the focus questions used to shape and direct collaborative discourse about the topic. It furthermore describes the type of abstraction activity used by the actors to find a common understanding and provides context to the analysis of the modeling session by describing the instructions, guidelines and constraints influencing the social interaction processes. The approach presented here focusses on how the group interacts on a fine grain level to align the understanding of the modeling topic. It distinguishes phases of elicitation and externalization from those concerned with different forms of consensus building.

The three remaining aspects are all only addressed by a single approach. The *representation of statements about the topic in the model* is covered by Rittgen's (2007) semantic level. It allows to map statements on the topic of modeling to be mapped to conceptual model semantics (such as fork, join, trigger, etc.). The *process of building a common model* is only addressed by Modeling Phase Diagrams. While this approach distinguishes different types of activities derived from observable model manipulations, it basically is agnostic about whether the model is created by an individual or a group. In the context of collaborative

modeling, the identified activities thus always refer to the group as a whole. Still, this aspect might offer further analytical potential when examined with an approach that explicitly tries to identify group model building activities. This potential, however, has not been further explored in the present article.

The *quality of actors' statements about the topic of modeling* is only addresses in the approach introduced here and should be considered the main contribution of the paper. This aspect is addressed in two dimensions. First, the epistemological dimension allows to analyze the point of reference for statement made on the topic. This allows to identify whether the model building activities are derived from a concrete, case-based level or from a more abstract perspective, which considers a generic business process with all its variants. This is valuable to determine, if a collaborative modeling process has achieved a sufficient level of abstraction for the respective goal of modeling. Second, it analyzes the quality the observed arguments with respect to whether they are grounded and qualified, assuming that the ability to ground and qualify claims about the modeling topic allows to draw conclusions about the maturity of the respective actor's understanding of the topic of modeling and its context.

While evaluating the quality of actors' statements about the topic of modeling provides value in itself, the results described in Sect. 4.8 and the complementarity of the approaches described in this section hint at a potential for further combining the different analytical dimensions and collecting data on different levels of granularity (c.f. Fig. 12, bottom right). This could allow to analyze a single aspect of investigation more thoroughly. As an example, combining FoCons and the social modes of co-construction as described in this article to analyze development of a common understanding about the topic could show, which actors set elicitation activities in a certain role in the modeling process to facilitate the development of a common understanding. Furthermore, as indicated in Sect. 4.8, combining analytical dimensions of different aspects of investigation (such as the pragmatic dimension of Rittgen (2007) and the argumentative dimension proposed here) could allow to identify in-depth insights of how actors negotiate a common understanding and refine their own views about the modeling topic in the course of a collaborative modeling session.

Still, for deployment in practice, the relevant aspects of evaluation would need to be selected based on the objectives of the planned study. This calls for a structured approach for selecting appropriate dimensions and informing the data collection process. Such an approach is beyond the scope of the present article and will be subject of future research. The structure shown in Fig. 12 together with the insights in planning evaluation processes described by Ssebuggwawo et al. (2013) could provide a starting point for the development of such an approach.

6 Conclusions

The field of the evaluation of collaborative modeling processes has shown significant progress over the last years. Current approaches focus on examining observable modeling activities or negotiation processes that show how actors agree

on how to build a model. One aspect that has hardly been examined, is how actors come to a common understanding about their topic by means of modeling. However, literature on conceptual modeling quality argues for the importance of this aspect (Gemino and Wand 2003; Krogstie et al. 2006; Mayer 1989). The article has set out to fill this gap by introducing an evaluation approach that explicitly addresses articulation and alignment activities to develop a common understanding about the topic of modeling during the collaboration of the involved actors.

The major contribution of this article is to introduce an articulation- and alignment-perspective on the evaluation of collaborative modeling processes. This concept has been implemented in an evaluation scheme based on an existing approach adapted from the field of analyzing collaborative learning processes. As a side contribution, the article provides a comprehensive overview about the state-of-the-art in collaborative modeling process evaluation by comparing the proposed scheme with other approaches on an operative level and demonstrates their complementarity and combinability.

The research presented in this article has several limitations that need to be addressed in future research. On a conceptual level, the analytical dimensions of the proposed coding scheme require further validation from an articulation- and alignment-perspective. It is currently motivated from the concept of argumentative co-construction of knowledge (Noroozi et al. 2012), but should also be reflected against other approaches in this field, such as group model building (Zagonel 2002) or cognitive apprenticeship (Dennen and Burner 2008). The conceptualization of successful articulation and alignment being reached when a shared understanding about the topic emerges is challenged in recent research (Kolikant and Pollack 2015), which emphasizes the value of non-convergent collaboration for developing the actors' understanding of the topic. These results could inform a revision of the analytical dimension representing the observable social modes of co-construction. Second, the comparison with other evaluation approaches proposed in related work hints at the potential for a more comprehensive theoretical framework for evaluating collaborative modeling processes. These hints, however, are too immature and require more extensive empirical validation in terms of the applicability of the identified observational aspects on different types of modeling processes. Finally, the data analysis for the introduced evaluation approach is currently cumbersome to conduct, in particular when based on video recordings of collaborative modeling sessions. The approach would largely benefit from tool support, which should cover all phases of analysis, from the segmentation of source material and coding, to higher level analytics such as correlation analysis across the results in different dimensions.

The next research steps will focus on generating further empirical evidence on the applicability of the proposed methodology in further practical settings. These results will be used to determine analytical gaps and shortcomings of the current approach and allow for refinement of the evaluation scheme.

References

Antunes P, Simões D, Carriço L, Pino JA (2013) An end-user approach to business process modeling. J Netw Comput Appl 36(6):1466–1479. doi:10.1016/j.jnca.2013.03.014

Claes J, Vanderfeesten I, Pinggera J, Reijers HA, Weber B, Poels G (2013) A visual analysis of the process of process modeling. Inf Syst E-Bus Manag 13(1):147–190

Claes J, Vanderfeesten I, Gailly F, Grefen P, Poels G (2015) The structured process modeling theory (SPMT) a cognitive view on why and how modelers benefit from structuring the process of process modeling. Inf Syst Front 17(6):1401–1425

Clark HH, Brennan SE (1991) Grounding in communication. Perspect Soc Shar Cognit 13(1991):127–149

Dennen VP, Burner KJ (2008) The cognitive apprenticeship model in educational practice. Handb Res Educ Commun Technol 3:425–439

Forster S, Pinggera J, Weber B (2013) Toward an understanding of the collaborative process of process modeling. Proc CAiSE Forum 13:98–105

Gemino A, Wand Y (2003) Evaluating modeling techniques based on models of learning. Commun ACM 46(10):79–84. doi:10.1145/944217.944243

Gemino A, Wand Y (2004) A framework for empirical evaluation of conceptual modeling techniques. Requir Eng 9:248–260

Herrmann T, Hoffmann M, Loser KU, Moysich K (2000) Semistructured models are surprisingly useful for user-centered design. Des Cooper Syst Proc COOP 2000:159–174

Hjalmarsson A, Recker JC, Rosemann M, Lind M (2015) Understanding the behavior of workshop facilitators in systems analysis and design projects: developing theory from process modeling projects. Commun AIS 36(22):421–447

Hoppenbrouwers S, Rouwette E (2012) A dialogue game for analysing group model building: framing collaborative modelling and its facilitation. Int J Organ Des Eng 2(1):19–40

Hoppenbrouwers S, Wilmont I (2010) Focused conceptualisation: framing questioning and answering in model-oriented dialogue games. In: van Bommel P, Hoppenbrouwers S, Overbeek S, Proper E, Barjis J (eds) Proceedings of IFIP working conference on the practice of enterprise modeling. Springer, Berlin, Heidelberg, pp 190–204

Hoppenbrouwers S, Proper HA, van der Weide TP (2005) A fundamental view on the process of conceptual modeling. Er 3716(Chapter 9):128–143. doi:10.1007/11568322_9

Kolikant YB-D, Pollack S (2015) The dynamics of non-convergent learning with a conflicting other: internally persuasive discourse as a framework for articulating successful collaborative learning. Cogn Instr 33(4):322–356

Konaté J, Sahraoui AEK, Kolfschoten GL (2013) Collaborative requirements elicitation: a process-centred approach. Group Decis Negot 23(4):847–877. doi:10.1007/s10726-013-9350-x

Krogstie J, Sindre G, Jørgensen HD (2006) Process models representing knowledge for action: a revised quality framework. Eur J Inf Syst 15(1):91–102. doi:10.1057/palgrave.ejis.3000598

Lesh R, Harel G (2003) Problem solving, modeling, and local conceptual development. Math Think Learn 5(2–3):157–189

Mayer RJ (1989) Models for understanding. Rev Educ Res 59(1):43–64

Mendling J, Hahn C, Recker JC (2011) An exploratory study of IT-enabled collaborative process modeling. Workshop Proc BPM 2010(66):61

Mullery GP (1979) CORE-a method for controlled requirement specification. In: ICSE '79 proceedings of the 4th international conference on software engineering, pp 126–135

Nelson HJ, Poels G, Genero M, Piattini M (2011) A conceptual modeling quality framework. Softw Qual J 20(1):201–228. doi:10.1007/s11219-011-9136-9

Noroozi O, Weinberger A, Biemans HJA, Mulder M, Chizari M (2012) Argumentation-based computer supported collaborative learning (ABCSCL): a synthesis of 15 years of research. Educ Res Rev 7(2):79–106. doi:10.1016/j.edurev.2011.11.006

Oppl S (2016a) Supporting the collaborative construction of a shared understanding about work with a guided conceptual modeling technique. Group Decis Negot. doi:10.1007/s10726-016-9485-7

Oppl S (2016b) Articulation of work process models for organizational alignment and informed information system design. Inform Manag 53(5):591–608. doi:10.1016/j.im.2016.01.004

Pinggera J, Zugal S, Weidlich M, Fahland D, Weber B, Mendling J, Reijers HA (2012) Tracing the process of process modeling with modeling phase diagrams. In: Proceedings of business process management workshops 2012, pp 370–382

Pinggera J, Furtner M, Martini M, Sachse P, Reiter K, Zugal S, Weber B (2013) Investigating the process of process modeling with eye movement analysis. Proc Bus Process Manag Workshops 2013:438–450

Recker JC, Mendling J, Hahn C (2013) How collaborative technology supports cognitive processes in collaborative process modeling: a capabilities-gains-outcome model. Inf Syst 38(8):1031–1045

Rittgen P (2007) Negotiating models. In: Krogstie J, Opdahl A (eds) Advanced information systems engineering, vol 4495. Springer, Berlin, pp 561–573. doi:10.1007/978-3-540-72988-4_39

Rouwette EAJA, Vennix J, van Mullekom T (2002) Group model building effectiveness: a review of assessment studies. Syst Dyn Rev 18(1):5–45. doi:10.1002/sdr.229

Saaty TL (1990) How to make a decision: the analytic hierarchy process. Eur J Oper Res 48(1):9–26

Sedrakyan G, Snoeck M, De Weerdt J (2014) Process mining analysis of conceptual modeling behavior of novices–empirical study using JMermaid modeling and experimental logging environment. Comput Hum Behav 41:486–503

Seeber I, Maier R, Weber B (2012) CoPrA: a process analysis technique to investigate collaboration in groups. Proc HICSS 2012:363–372

Sindre G, Lindland OI, Solvberg A (1994) Understanding quality in conceptual modeling. IEEE Softw 11(2):42–49

Soffer P, Kaner M, Wand Y (2012) Towards understanding the process of process modeling: theoretical and empirical considerations. Proc Bus Process Manag Workshops 2012:357–369

Ssebuggwawo D (2012) Analysis and evaluation of collaborative modeling processes. BOXPress, 's-Hertogenbosch

Ssebuggwawo D, Hoppenbrouwers S, Proper HA (2013) Applying AHP for collaborative modeling evaluation. Int J Inf Syst Model Des 4(1):1–24. doi:10.4018/jismd.2013010101

Strauss A (1993) Continual permutations of action. Aldine de Gruyter, New York

Trickett SB, Trafton JG (2009) A primer on verbal protocol analysis. In: Schmorrow D, Cohn J, Nicholson D (eds) The PSI handbook of virtual environments for training and education, vol 1. Praeger Publishers Westport, CT, USA, pp 332–346

Van Someren MW, Barnard YF, Sandberg JAC (1994) The think aloud method: a practical guide to modelling cognitive processes. Academic Press, Cambridge

Vennix J, Akkermans HA, Rouwette EAJA (1996) Group model-building to facilitate organizational change: an exploratory study. Syst Dyn Rev 12(1):39–58

Weinberger A, Fischer F (2006) A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. Comput Educ 46(1):71–95. doi:10.1016/j.compedu.2005.04.003

Weinberger A, Stegmann K, Fischer F (2007) Knowledge convergence in collaborative learning: concepts and assessment. Learn Instr 17(4):416–426. doi:10.1016/j.learninstruc.2007.03.007

Wilmont I, Hengeveld S, Barendsen E, Hoppenbrouwers S (2013) Cognitive mechanisms of conceptual modelling. In: Wilfred NG, Storey VC, Trujillo JC (eds) Conceptual modeling. Proceedings of ER 2013. Lecture notes in computer science, vol 8217. Springer, pp 74–87

Zagonel AA (2002) Model conceptualization in group model building: a review of the literature exploring the tension between representing reality and negotiating a social order. In: Proceedings of the international conference of the system dynamics society