Collaborative Modeling of Business Processes – A Comparative Case Study

Peter Rittgen
Vlerick Leuven Gent Management School
Reep 1
9000 Gent, Belgium
+32 9 210 9862
peter.rittgen@vlerick.be

ABSTRACT
We study collaborative modeling of business processes with respect to the impact of tool support on the modeling process. For this purpose we compared model quality and modeling costs in two cases. The first was carried out with the help of a collaborative modeling tool; in the second case we kept all other parameters as closely as possible to the first one but conducted the modeling session in the usual way without tool support. We observed a marked increase in modeling time in the second case and a reduction in model quality.

Categories and Subject Descriptors
I.6.5 [Model Development]: Modeling methodologies

General Terms
Management, Documentation, Design.

Keywords
Computer-mediated communication and collaboration, systems design and implementation, business process modeling.

1. INTRODUCTION
Enterprise modeling is a collaborative process that usually involves a number of participants. Conventional wisdom has it [1-5] that modeling sessions should be carried out in a chauffeured manner. This implies that the session is led by a facilitator who is served by a scribe and a number of domain experts. The facilitator elicits information about the business process from the domain experts and translates it into the modeling notation, typically using a flip chart to sketch parts of the model to get feedback from the audience. The scribe enters the sketches that are approved into the computer with the help of a modeling tool. This tool can be any modeling tool for the language that is used or even a simple drawing tool.

Chauffeured sessions have been criticized [6, 7] for a number of reasons. The fact that all interaction has to happen via the facilitator introduces a substantial bottleneck that severely limits the efficiency of the modeling team. [6] estimates this effect to account for session times that are up to 3 times longer as compared to sessions that are supported by the EMS-IDEF0 tool they used. This figure was never measured, though, but the potential benefit justifies a closer investigation of this issue, especially as today’s tools provide more comprehensive support that, for example, also covers negotiation aspects. The latter account for much of the complexity of group modeling as consensus is notoriously hard to achieve [8].

Another issue concerns model quality. By model quality we mean the degree to which it fulfils the purpose for which it was created. In business process modeling the purpose of the model is often to serve as a basis for a re-engineering or system development project. The success of such a project largely depends on stakeholder buy-in. As key stakeholders are often represented in modeling teams, the quality of models can be assessed as the quality perceived by the team members.

But in traditional chauffeured modeling the involvement of team members is rather passive. Participants often have only a limited understanding of the model which they see grow but cannot actively shape. They are not sure that it actually reflects the process as they see it. This can imply poor identification with the model and ultimately threaten stakeholder buy-in. A group modeling tool, on the other hand, offers all team members direct access to the model under development and therefore has the potential to give the participants a more active role in the process. We believe that this will strengthen their understanding of the model, their awareness of its implications and also their agreement with its consequences.

To take a closer look at the impact of tool support on modeling we have performed a comparative case study. The study was part of a larger project where we did a business process analysis of the customer distribution center of a manufacturer of mobile network components. The process covered the whole logistics chain from materials receipt via production to configuration, packaging and dispatch of goods.

2. RESEARCH METHODOLOGY
The study was performed as a comparative case study. We compared two cases where each case was one modeling session. Within a session a modeling task had to be performed that consisted of the modeling of a complex activity in the logistics
To exclude dependence on participant characteristics we used the same team in both cases. We also made sure that the complexity of the respective modeling tasks was approximately the same by choosing two tasks, Handle problem goods and Store in high-bay warehouse, which we deemed equally difficult. To verify that they were indeed of roughly the same complexity we evaluated them ex post with the help of two measures of complexity, i.e. model complexity and perceived modeling task complexity. For the former we used a simple measure, i.e. the sum of all vertices and edges in the final model. For the latter we used two indicators:

1. A question in the questionnaire that asked for a subjective judgment of the task complexity on a scale from 1 (very easy) to 10 (very hard). This indicator was averaged for all group members.

2. The number of modeling rounds (i.e. versions) that were required to reach the final model.

To exclude dependence on the participants’ learning curve we arranged the two sessions in the order: first with tool support then without. In the case of the reverse order a better result in the tool supported case could have been attributed to the modeling experience gained in the unsupported case (instead of the tool supported case) could have been attributed to the then without. In the case of the reverse order a better result in we arranged the two sessions in the order: first with tool support then without. In the case of the reverse order a better result in the tool supported case could have been attributed to the modeling experience gained in the unsupported case (instead of the tool supported case) could have been attributed to the then without.

The main dependent variables we measured were the perceived model quality and the modeling time. The former was measured with a questionnaire that was administered directly about the participants’ learning curve we arranged the two sessions in the order: first with tool support then without. In the case of the reverse order a better result in the tool supported case could have been attributed to the modeling experience gained in the unsupported case (instead of the tool supported case) could have been attributed to the then without. In the case of the reverse order a better result in the tool supported case could have been attributed to the modeling experience gained in the unsupported case (instead of the tool supported case) could have been attributed to the then without.

The pragmatic quality is related to the audience’s understanding of the model [11]. Model comprehensibility was measured by having the participants make up a list of things that they thought were missing in it. We averaged on the list count for a model as an inverse measure of its comprehensibility.

The pragmatic quality is related to the audience’s understanding of the model [11]. Model comprehensibility was measured by having the participants make up a list of things that they thought were missing in it. We averaged on the list count for a model as an inverse measure of its comprehensibility.

The validity test was performed with two groups of 17 and 14 students taking part in the courses Information Systems and Business Processes and Unified Modeling Language (UML), respectively. All students used the same case that concerned the admission, care, ward transfer and discharge processes of a hospital. The first group was required to create both EPC (Event-driven Process Chain) and UML models while the latter did only UML models leading to a total of 48 models. Models
were developed in teams of 4 where each member was responsible for a different core process. That person was not allowed to judge the model. The odd person in the first group modeled one process by himself. The two odd students in the second group did two of the processes.

Modeling was done with the help of a drawing tool to allow for syntactic errors which a modeling tool would have prevented. After the final modeling session we administered the questionnaire. This was done first to ensure that the detailed analysis of each model would not affect the participants' subjective perceptions. After that we counted the syntactic errors with the help of a modeling tool and the built-in syntax checker. During the next tutorial the students were asked to make up three lists containing the mistakes they saw in the model, the things that they thought were missing in it and the things they did not understand.

The length of the lists were counted and averaged for each model and all data was entered into SPSS to perform correlation analysis. The result is shown in Table 1.

<p>| Table 1. Correlations between the quality indicators and the perceived quality measurement |
|-----------------------------------------------|-----------------------------------------------|
| <strong>Quality indicator</strong> | <strong>Statistic</strong> | <strong>Perceived quality</strong> |
|-----------------------------------------------|-----------------------------------------------|
| Syntactic errors | Pearson Correlation | -0.298 |
| Semantic mistakes | Pearson Correlation | -0.321 |
| Missing things | Pearson Correlation | -0.415 |
| Things not understood | Pearson Correlation | -0.384 |
| Sig. (2-tailed) | Sig. (2-tailed) | Sig. (2-tailed) |</p>
<table>
<thead>
<tr>
<th>4.0 %</th>
<th>2.6 %</th>
<th>0.3 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7 %</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results show that the correlation between the quality indicators and the perceived quality measurement is significant for all indicators on the 5% level. We have thereby established convergent validity of the new measurement. The correlation factors range from -0.298 to -0.415. They are all negative because the indicators represent inverse measures of the syntactic, semantic and pragmatic qualities. They are also not very high which was to be expected because the perceived quality is an aggregate measure that covers a number of aspects, but each indicator just deals with a certain one and therefore only explains some part of the overall quality.

The advantage of the perceived quality measure lies primarily in the fact that it is easy to administer and needs just a few minutes as well as only basic modeling knowledge. The more common metric proposed in [11], on the other hand, requires substantial expertise in modeling to detect the different kinds of mistakes correctly and completely, and it is also quite time-consuming. In our experiment it took between 2 and 3 hours of thorough inspection for each model. The new measure does not indicate the cause of a poor quality, though, and does therefore not provide support for model improvement.

4. CHOICE OF TOOL

With validity being established we can return to the study of the impact of tool support on collaborative modeling. For this purpose we use the new measure to determine the benefit of modeling, which is related to model quality in a crucial way as we will argue later on. The study of tool impact requires the identification of a suitable collaborative tool. But the majority of the currently existing modeling tools are single-user tools. Strangely, this is even the fact for the ones that explicitly address group modeling (cf. e.g. Compendium [12]). Some notable exceptions do exist [6, 13, 14] but they do not explicitly support negotiation and consensus building, two important aspects of collaborative modeling.

A recent tool that claims to address these concerns is the COMA tool (COllaborative Modeling Architecture, see http://www.coma.nu). The COMA tool is a collaborative tool for modeling in UML. A client of COMA is installed on the computer of each group member including the facilitator. Work among the participants is coordinated via a server. Besides a conventional model editor COMA provides the negotiation functions: Propose, support, challenge and accept.

A proposal is a suggestion for the revision of the current version of the model. It implies that the modeler posts the content of the local model editor to the group. In building the local or personal version of the model the modeler can make use of bits and pieces of existing versions (i.e. the group model or other proposals), or even copy a whole version and apply changes to it.

A support is a positive assessment of a proposal. It can be logged by any team member after reviewing the respective proposal. It can be complemented by a comment that provides a rationale for the decision and perhaps includes suggestions for minor changes.

A challenge is a negative assessment of a proposal. It has to be complemented by a justification for the decision as well as constructive comments regarding improvements of the proposal.

COMA offers two rules to decide on the acceptance of proposals: A rule of majority and a rule of seniority. When a rule of majority is used, the team operates in an unfacilitated mode where each modeler has a vote of the same weight. Acceptance only depends on the number of supports and challenges. The rule specifies the minimum number of supports required, and the maximum number of challenges allowed for a proposal to be accepted. The required number of supports should be at least two to avoid that a modeler alone (e.g. the proponent) can make the decision. A maximum number of challenges of 0 would force a unanimous decision. When a rule of seniority is applied, the team has a facilitator that makes the decision. Other group members cannot directly influence the decision, but they can do so indirectly by making suitable comments (i.e., supports and challenges). The facilitator can and should consider the supports and challenges in the decision. Fig. 1 contains a screenshot of the COMA tool.
Figure 1. Screenshot of the COMA tool

It shows the point in time where the group is about to prepare the second version of the activity diagram for the treatment of problem goods (i.e., goods with an unclear recipient). The upper window contains the old version V001 as a point of reference. Here the participants had only collected all the relevant activities. A proposal from Peter (in the lower right window) suggests ordering these activities in a certain sequence but Jenny (the user from whom the screenshot was taken) refines it based on her more detailed knowledge of the process (lower left window). In the next step Jenny’s proposal was accepted by the group and became version V002 (not visible in the screenshot).

5. RESULTS

We used the research methodology described in section 2 on a group of 7 participants. The group was made up of 1 facilitator (a consultant), 3 active modelers from the software department and 3 domain experts from the logistics unit. All of them received training with the COMA tool in advance of the study. The training consisted in a presentation of ideas and concepts behind COMA (ca. 1 h), a demonstration of the tool itself (ca. 1.5 h) and a short trial session where the group modeled a simple fictitious example (ca. 2 h).

One week later we conducted the first modeling task, Store in high-bay warehouse, which required two sessions with 2 and 3 rounds each. The first session was carried out during the morning and lasted for 2 hours and the second session was done after lunch from 13:00 to 15:35. Both sessions were performed with tool support, i.e. participants had access to a networked laptop with the COMA client.

Yet another week later we conducted the second modeling task, Handle problem goods, which required 4 sessions, a morning and an afternoon session each on two consecutive days. Each session lasted for almost 3 hours and led to a new model version. These sessions were performed in the conventional way with the facilitator using a flip chart to develop the model based on participant input. Between sessions the model was drawn up with a conventional drawing tool, printed out and distributed to the participants as a starting point for the next session but no tool was used during the sessions. The results are shown in Table 2.

Table 2. Impact of tool support on model quality and modeling time

<table>
<thead>
<tr>
<th>Case</th>
<th>Model complexity</th>
<th></th>
<th>Modeling complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nodes</td>
<td>Edges</td>
<td>Total</td>
</tr>
<tr>
<td>Store in high-bay warehouse</td>
<td>14</td>
<td>18</td>
<td>32</td>
</tr>
<tr>
<td>Handle problem goods</td>
<td>10</td>
<td>17</td>
<td>27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tool support</th>
<th>Model quality</th>
<th>Modeling time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Store in high-bay warehouse</td>
<td>Yes</td>
<td>0.76</td>
</tr>
<tr>
<td>Handle problem goods</td>
<td>No</td>
<td>0.48</td>
</tr>
</tbody>
</table>

6. DISCUSSION

Our study is not the first one that has found a positive impact of tool support on group modeling. E.g. [6] reports on a better model quality that can be achieved by involving a larger number of participants that can contribute knowledge and provide quality assurance of the model. But larger groups cannot be handled in the conventional way anymore and therefore require new facilitation techniques and tool support. The same paper also gives a rough estimate of the potential time savings. But neither this study nor any other we are aware of actually
measures the impact of tool support. Our approach mainly aims in that direction and suggests a first step towards a research method that quantifies the important parameters of modeling: the process costs and the quality of the outcome.

We are aware of the fact that our method is restricted to one cost and one quality dimension: modeling time and perceived model quality, respectively. Although other dimensions could be considered we argue that those two are the most important ones. Modeling as a collaborative activity requires the involvement of many people but very little other infrastructure. The team members and facilitators are typically highly paid domain experts and consultants. Personnel costs are therefore the major cost driver. As the number of involved people cannot be reduced without sacrificing model quality and buy-in, as already mentioned, the only option left for cost saving is reducing the modeling time [6].

The situation regarding model quality is less obvious. Here many dimensions have been suggested such as correctness, completeness, clarity, adequacy etc. We do not challenge the importance of these but rather claim that perceived model quality subsumes the most relevant of them. The significant correlation between four of these dimensions and our measure as found in the student experiment shows this.

But we can also arrive at the same conclusion by way of a simple argument: Apart from the facilitator the participants of most modeling teams fall into two categories: domain experts and stakeholders [15]. The domain experts contribute the knowledge that is contained in the model, i.e. they are responsible for the content. The stakeholders are affected by the output, i.e. they try to make sure that the model fulfils its purpose (e.g. the model’s suitability as input for reengineering or systems development). Now, the perceived model quality with respect to the group is an aggregate measurement that averages over the quality perceptions of the individual members. If its value is high, then most of the members assess the model quality as high, i.e. domain experts as well as stakeholders. This implies that the content is correct and that the model is suitable for its purpose.

But why should we measure model quality and costs at all? From a practical point of view the answer is quite clear. Any manager having to justify a modeling budget will need hard facts on the costs-benefits situation. From a scientific perspective, it gives us an instrument that demonstrates the value of tool support in collaborative modeling. This issue has so far been fairly neglected although it is folklore in IT in general [16]. To give a few examples: It allows us to compare existing tools with respect to their effectiveness. It can also be used to assess the effect of a certain feature or function in a tool, which in turn can drive method development and improve our understanding of collaborative modeling. But tool support is only one factor we might study. The list of factors worth investigating can be augmented with the modeling method that is used, the way that the group is organized, the facilitation technique and so on.

In this paper we have applied the suggested research method to the determinant ‘tool support’ and a particular tool. In doing so, we have arrived at an outcome that agrees with that of similar studies. We have thereby established face validity of the research method.

7. CONCLUSIONS
The purpose of the study presented in this paper is two-fold: On the one hand we wanted to investigate the impact of tool support on the group modeling process in a more rigorous way than has been done before. But on the other hand we also wanted to test the new quality metric in practice. Modeling in itself is already a highly demanding task that is further complicated by the dynamics of group work. Effective support is therefore essential, especially if some of the group members are inexperienced as is typically the case in business process modeling, where a majority of the participants often does not have any modeling background. But it is precisely this latter type of participant that contributes most to the actual design of the model with his or her knowledge of the relevant business domain. Both the speed and quality of the models can therefore benefit tremendously if we can manage to involve these people directly as modelers instead of relying on the bottleneck of the facilitator for all communication within the group. The investigated tool support can accommodate this by giving the expert seniority (i.e., the right to make the final decision) and turning the domain experts into effective consultants that make proposals (thereby reversing the traditional roles in IT consulting).

A collaborative modeling tool is a distributed modeling support system and can also be seen as a special kind of group decision support system (GDSS, [17]) if we consider that the accept and reject decisions in the negotiation process are key to model design. There is significant empirical support for the claim that GDSS are beneficial [17-26], particularly for larger groups or complex tasks. Many of these benefits carry over to collaborative modeling tools, e.g., reduced meeting time, higher quality of decisions, broader involvement of participants, higher effectiveness of decisions, etc.

The comparative case study we have performed cannot claim to prove that tool support in general reduces modeling time and improves model quality. But it shows that the methodology works in principle and that it is worthwhile to invest in a larger field study that not only covers more cases but also a wider scope of different modeling teams and a variety of other factors. Beyond assessing the usefulness of tool support a broader study can help in better understanding the dynamics of group modeling.

8. REFERENCES


