Spring 6-10-2017

ON THE EFFECT OF MIXING TEXT AND DIAGRAMS ON BUSINESS PROCESS MODEL USE

Toomas Saarsen
University of Tartu, Estonia, toomas.saarsen@ut.ee

Marlon Dumas
University of Tartu, Estonia, marlon.dumas@ut.ee

ON THE EFFECT OF MIXING TEXT AND DIAGRAMS ON BUSINESS PROCESS MODEL USE

Research paper

Saarsen, Toomas, University of Tartu, Estonia, toomas.saarsen@ut.ee

Dumas, Marlon, University of Tartu, Estonia, marlon.dumas@ut.ee

Abstract

A picture is worth a thousand words, but a few words can greatly enhance a picture. It is common to find textual and diagrammatic components complement each other in enterprise models in general, and business process models in particular. Previous work has considered the question of the relative understandability of diagrammatic versus textual representations of process models for different types of users. However, the effect of combining textual and diagrammatic components on the actual use of process models has to the best of our knowledge not been considered. This paper addresses the question of how the mix of diagrammatic and textual components in business process models affects their sustained use. This question is approached via a case study in a telecommunications company where models with different mixtures of text and diagrams have been collected over time. The study shows that models, in which the ordering relations between tasks are captured in diagrammatic form, while the details of each task are captured in textual form, are more likely to be used on a sustained basis.

Keywords: Business Process Modelling, Process Model Use.

1 Introduction

Business process models are a widely employed vehicle for preserving and communicating critical knowledge about business operations. A key tradeoff that business process models need to strike is that they need to be simple enough to be understood by a wide range of stakeholders, yet precise and detailed, so that these stakeholders can extract from them actionable insights for the performance and improvement of daily business operations (Bateman et al., 2001)

A common approach to strike this tradeoff is to combine diagrammatic and textual components in a business process model. For example, the diagrammatic component may be used to highlight important relations between the elements of the business process, e.g. to show temporal relations between tasks (Eppler and Burkhard, 2007). Meanwhile, textual descriptions are used to provide detailed documentation about each element, e.g. steps involved in a task and business rules or guidelines relevant for its performance.

If the intended users of a process model are process managers and analysts, a wide range of details can be captured in diagrammatic form by exploiting advanced process modelling constructs, such as those available in the Business Process Model and Notation (BPMN). If, however, the intended users are process workers, i.e. employees who work on the process on a daily basis, and if the model is part of their operational knowledge base (Nonaka, 2008), it might be counter-productive to capture too many details in diagrammatic form. Indeed, process workers do not usually have the required fluency in process modelling notations to understand subtle process modelling constructs. In addition, many of the details they are looking after are very fine-grained and might affect only one task locally and hence do not affect the flow of control across tasks. These observations raise the following question: in the context of process models intended to be used by process workers, how much and what information should be presented in diagrammatic form, and how much and what information should be captured in textual form?

The aim of this study is to identify relations between the use of process models in an operational knowledge base, and their representation format (text, diagrams and combinations thereof). Specifically, the study seeks to identify combinations of text and diagrams in such process models are associated with their sustained use. In this context, sustained use is defined as the regular use of a process model by its intended users, past the project or initiative where the model was initially produced.

Importantly, the study focuses on process models that capture operational knowledge (i.e. how the process should be performed at the lowest level of detail) and are hence intended for consumption by process workers, as opposed to process models that capture tactical knowledge and are intended to be used by process managers or analysts for process improvement. The study also excludes executable process models, i.e. machine-readable process models that are intended to be used to configure a process-aware information system, such as a Business Process Management System (BPMS).

In order to address the above question, we analysed the use of process models in a large organization that maintains an operational knowledge base consisting of process models with different styles, ranging from models consisting mainly of text and tables, to models with a predominantly diagrammatic style. We analysed the mix of text and diagrams in the process models of the organization, and related this mix to their sustained use.

The rest of the paper is structured as follows. Section 2 lays down the theoretical background of the study by analysing previous work and deriving from it terminology. Section 3 presents hypotheses and variables used to test these hypotheses. Section 4 then presents the case study setting and data collection, while Section 5 discusses the findings. Finally, Section 6 summarizes the contribution of the paper.

2 Related work

Nolte et al. (2013) has identified factors that contribute to the use of process models, covering both organizational factors and usability factors (e.g. perceived ease of interpretation, perceived semantic quality and satisfaction with use). Sanches-Gonzalez et al. (2010) and Mendling et al. (2007) have empirically analysed a number of factors that determine the understandability of business process diagrams. These and other studies have proposed and evaluated several complexity measures of process diagrams (Sánchez et al., 2010) such as the number of nodes, the average gateway degree and the density of the diagram. However, these studies focus on purely diagrammatic process models without taking into account textual components.

In order to improve the understandability of process models, different practical recommendations and guidelines have been assembled and validated in a number of studies (Mendling et al., 2007), (Dumas et al., 2007), (Mendling and Strembeck, 2008). Mendling et al. (2010) outline and empirically validate seven modeling guidelines aimed at increasing the understandability of process models. The first five guidelines are specific to diagrammatic modeling notations. The sixth guideline ("use verb-object activity labels") refers to the labeling of tasks in diagrammatic process models. This guideline ties up diagram and text but only in the narrow setting of task labeling. The last guideline ("decompose a model with more than 50 elements into smaller models") refers to process-subprocess decomposition and is applicable to both diagrammatic and textual process descriptions. This latter guideline is related to more general guidelines (Mesarovic, 2000) for decomposing complex models and documents. Variants of this latter guideline can be found in a variety of fields, e.g. management (McMillan, 2002), software design (Yourdon, 1989), document management (de Waard et al., 2012). Accordingly, we retain the number of elements in a model as one of the main parameters in our study.

Ottensooser et al. (2012) analyse the relative understandability of purely textual vs. purely graphical process descriptions. Their results show that process diagrams are associated with higher understandability. However, their study does not consider combinations of diagrams and text in the same model. A common point between the study of Ottensooser et al. and ours is that we focus on on process models that are intended to be used by process workers during the performance of the process.

In other studies with broader scope (not focused on process models), the supporting role of diagrams in understanding textual descriptions has been highlighted: Eppler and Burkhard (2007) analyse the visual representation of information in the context of knowledge management; Carney and Levin (2002) study focus is on the learning aspect; the study by Larkin and Simoni (1987) brings out the context where diagrams are efficient to use. These and other studies assert that interleaving text and diagrams generally enhances understandability (Bateman et al., 2001). These articles provide general recommendations for enhancing knowledge reuse, but no concrete guidelines that would be specifically applicable to business process models.

Links have also been established between various quality dimensions and usability of process models (Krogstie, 2016). In this respect, it has been established that both semantic quality (the fact that the model reflects reality) (Batini et al. 2009), and syntactic quality (correct use of the modeling notation) (Moody, 2005) contribute to process model usability. In our study, we concentrate on assessing the balance between diagrams and text in a process model and its relation to sustained use.

In summary, this paper differs from previous ones in that it studies how the mixture of diagrammatic and textual components in a process model relates to its sustained use. Other studies have either studied the understandability of general-purpose documents that combine visual and textual components, or the understandability and usability of diagrammatic process models taken in isolation, or compared to purely textual process descriptions as in (Ottensooser et al., 2012). Another distinguishing feature of the present study is that it focuses on process models that are intended for consumption by process workers. Previous studies have studied the understandability and use of process models in a broader setting, without distinguishing between process models intended for use by analysts and managers only (e.g. for process analysis, improvement or implementation) versus models that are intended to be used as a reference during the performance of a process.

3 Hypotheses and variables

In this section, we discuss the hypotheses of the study in terms of relations between independent variables capturing different characteristics of a process model, and the dependent variable, namely (sustained) process model usage.

3.1 Hypotheses

We are interested in establishing links between variables characterizing the mixture of textual and diagrammatic components in process models, and the sustained use of these models. Accordingly, the general null hypothesis is that there is no connection between the variables characterizing the balance between text and diagrams in a model, and its sustained use.

> H0. The variables of models that are used on a sustained basis do not differ from the variables of a model with a narrower scope of use.

This null hypothesis will be instantiated for each of the characteristics discussed below, each of which late gives rise to an independent variable.

3.1.1 Visual presentation of the tasks

Since we are interested in finding a suitable balance between diagrams and text in a process model, and given that the tasks are arguably the main elements of a process, a natural question is how many tasks should be presented graphically in a process model vs. how many should be described only as text? The aim here is to determine whether or not the presentation of more tasks in diagrammatic form increases the sustained use of a model.

H1. Process models where more tasks are visually presented (i.e. more tasks are represented as diagrammatic shapes) are more likely to be used on a sustained basis.

3.1.2 Visual presentation of the process hierarchy

In order to present the context of the tasks in a process, a structured decomposition is generally used (Mesarovic et al., 2000); this enables modellers to decompose complex objects (in our case, tasks) into smaller and simpler sub-objects. Such decomposition is carried out until objects are reached at a level of detail that is sufficient to comprehend their relationships. If the reader is given a visual representation of the decomposition (Browning, 2009), this may help him/her with a way of navigating in order to locate specific objects. The importance of structure in process models has been emphasized in several studies, e.g. Laue and Mendling (2010). The structure of a collection of processes is called a process architecture.

H2. Models that include a visual presentation of the process architecture are more likely to be used on a sustained basis.

3.1.3 Visual presentation of the ordering relation

In addition to being used to capture hierarchical (part-of) relations, process models are used to capture ordering relations between tasks. There is a tradeoff here between capturing these ordering relations in diagrammatic versus in textual form. Capturing ordering relations via diagrammatic constructs with clear execution semantics can enhance the understandability and precision of process models (Ottensooser, 2012). On the other hand, if all ordering relations are captured diagrammatically (including those between very fine-grained tasks, also known as steps), the diagrams may become overly complicated (Sánchez-González, 2010). Hence we are interested in testing the following hypothesis.

H3. Models that include a diagrammatic presentation of the ordering relations between tasks are more likely to be used on a sustained basis.

3.1.4 Size of the model

If a model is to be used by process workers on a daily basis, it needs to include enough details so that process workers cannot learn them all by habit and thus find value in consulting the model constantly. Hence, one can hypothesize that a model that is used on a sustained basis is likely to be larger than models that are used on an ad hoc basis. We can find a similar claim in Nolte et al. (2016) where one of the factors that is found to promote reuse of process models is their (total) size (including the size of all subprocess models if any).

H4. Larger process models are more likely to be used on a sustained basis.

3.2 Variables and Scales

For the independent variables, we chose general variables that directly map to the four characteristics of process models discussed above. We have deliberately chosen coarse-grained scales for these variables because gathering more detailed information is prone to errors in the case of models with textual descriptions – for example, the size of the model (number of tasks) cannot always be ascertained with high accuracy for textual models as the notion of task can be subjective. Also, it is unlikely that a finer granularity would add accuracy to data analysis and to the testing of the hypotheses. Likewise, in the implementation of the insights obtained from the study, coarse-grained results are more significant.

The variables and their hypothesized relations are summarized in Figure 1.

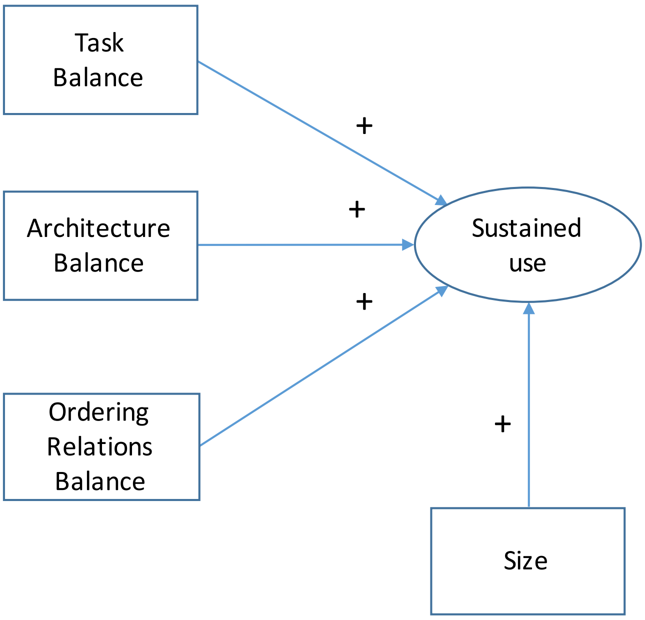

Figure 1. Hypothesized relations between variables

3.2.1 Dependent variable - Process Model Usage

We include one dependent variable (Process Model Usage) in the study. The variable in question captures whether the process model has been used on a sustained basis or not. The value range of the dependent variable is:

- 0 – No sustained use – the model is not used regularly; it is potentially used by managers or analysts on an ad hoc basis, e.g. a couple of times a year.
- 1 – Sustained use – the model is used at least once per week by at least one process worker in their performance of the process, for a period of at least one year after initial creation of the model.

3.2.2 Variable - Task Balance

A task in a process model may appear in diagrammatic form (i.e. as a task node in a process diagram) or in purely textual form (e.g. as a step in a textual description or as an item in a checklist). To capture this dichotomy, we define a variable task balance as the ratio between the number of diagrammatically presented tasks and the total number of tasks in a model (incl. those in textual form). To keep a coarse granularity (cf. discussion above), we present the variable as a factor characteristic similar to the scale by Schindler and Cooper (2005):

- 1 – all tasks are presented textual form;
- 2 – up to a third of the tasks are presented in diagrammatic form (most of the tasks have been described in the form of text, important tasks are presented by process diagrams – diagrams support the reading of the text);
- 3 – up to 66% of the tasks are presented in diagrammatic form (most of the tasks have been described in the form of text – diagrams provide the basis for documentation, text supplements the diagram);
- 4 – all or close to all tasks are presented in diagrammatic form (process model is depicted in the form of a diagram – the user receives most of the information from a diagram, and the descriptions of the elements of this diagram has been added as text).

3.2.3 Variable - Architecture Balance

We define the architecture balance of a process model as a characteristic factor derived from the percentage of task decomposition relations (e.g. process-subprocess relations) that are captured in diagrammatic form (Saarsen and Dumas, 2012). In the definition of the scale of this variable, we further distinguish the case where decomposition relations are captured in free-text form versus the case where they are captured in textual but structured form (tables and lists).

- 1 – all or most task decomposition relations are in free-text form;
- 2 – all or most task decomposition relations are in textual form: some in free-text form and others in structured text form (table of contents has been added to the descriptions of tasks, which brings out a structure consisting of up to two levels);
- 3 – all task decomposition relations are in structured text form (a detailed table of contents has been added to the descriptions of the tasks, which brings out a structure consisting of more than two levels);
- 4 – decomposition relations are partly in diagrammatic form and partly in textual form (in addition to a detailed table of contents, a visual diagram has been added to the descriptions of the tasks, which provides a visual overview of the hierarchy of tasks, simplifies the understanding of the structure of the table of contents in the use of the process model);
- 5 – all or most decomposition relations are captured in diagrammatic form (in the use of the process model, one relies on the process hierarchy presented in the graphic form).

3.2.4 Variable - Ordering Relations Balance

Similarly, we define the ordering relations balance as a characteristic factor derived from the percentage of task ordering relations captured in diagrammatic form, following existing definitions such as the one in (Zur Muehlen and Recker, 2008):

- 1 – all ordering relations are captured in textual form;
- 2 – all ordering relations are captured spatially, i.e. tasks in diagrams are arranged from left to right or top-down, but no arcs are used to denote ordering relations;
- 3 – all or most ordering relations are captured via arcs;
- 4 – in addition to the above, the start and end points of the process are explicitly captured in diagrammatic form;
- 5 – in addition to the above, alternative and parallel process branches are captured, e.g. using gateways in BPMN (Chinosi and Trombetta, 2012)

3.2.5 Moderating variable - Size

Finally, we define a factor variable by discretizing the size of the model, where size refers to the number of tasks, including (sub-)processes, tasks and steps.

- 1 – 10-50 – small number of tasks, typically high-level models where, for example, a general list or sequence of tasks is presented;
- 2 – 50-100 – small-to-medium-sized models typically used to for the purpose of analysis;
- 3 – 100-500 – medium-sized models with an average level of detail;
- 4 – 500-1000 – detailed models covering a significant portion of a value chain;
- 5 – 1000+ – detailed models of end-to-end processes.

We note that there may be a relation between size and architecture balance, as larger process models might require deeper hierarchical decompositions to remain manageable. However, this relation is not in the scope of this study.

4 Case Study

This section introduces the case study conducted to validate the hypotheses. First, an overview of the context of the study is provided. Second, the data collection methodology is introduced. Finally, the findings and validity issues related to the study are brought out.

4.1 Context

The case study was conducted at Telia, a large European telecommunications company with business units in 17 countries. The study focused on the Estonian branch, which has around 2000 employees. Telia Estonia has implemented process management practices for over a decade and self-assesses itself at level 4 on the CMM scale (Rosemann et al., 2004). It maintains process models covering all core and some support processes of the organization, that form the entire knowledge base of the organization These models are used by managers and analysts as well as by process workers. Although BPMN is the most widely used graphical process modelling notation, the company has enforced the use of BPMN in the creation of process diagrams, at the same time, there is no direct requirement to present all (especially more detailed) descriptions of processes only in diagrammatic form. Hence, models are maintained in a variety of formats, including free-text, structured text (tables, listings and checklists), free-form diagrams, BPMN diagrams and combinations thereof. This latter characteristic makes this organization suitable to test the formulated hypotheses.

The organization has been developed and managed in the process view already as of the year 2000, and process diagrams have been used as an important part of daily documentation from that time. Employees have been attending different process modelling workshops, and training, and done relevant tasks; for this reason, employees are familiar with the common elements of BPMN notation.

Process models in the case organization can be clearly divided into those that have been created for a one-off purpose (e.g. implementation of an IT solution in a particular unit) and those that have been created for wider use and indexed accordingly in their Intranet portal – the knowledge base of the organization. We excluded the first category of process models to avoid biasing the results – regardless of their size and characteristics, these one-off models are not indexed for reuse and thus they are unlikely to be used on a sustained basis.

Altogether, we gathered 48 indexed process models meant for use by process workers including: work instructions (33), models used for analysis (process models are created in the context of different BPM projects and integrated afterward into the knowledge base) (11) and high-level models intended for communication (4). We involved in the study all process models (48) that the knowledge base of the organization contains. These models are designed for usage by process workers and available to all employees of the organization in the Intranet. Models are defined at levels 4-6 with respect to the eTOM reference model (Kelly, 2007), which is common in this industry. In the context of our research, each model falls under one area of the eTOM reference model (which covers 15 areas altogether) like for example Service Development & Management or Customer Relationship Management. Models have been created and updated during a period of about 15 years. Models describing the same area may partially overlap each other.

We did not involve those process models in the study which had been directed at a smaller circle of users in the context of specific project, such as for IT-development or process analysis; during one year, approximately 20 of such models are created. In most cases, these models are created with the help of information from the knowledge base; also, these models could serve as triggers for implementing changes in the knowledge base (for example, if errors are discovered or changes are implemented in processes in the course of a project).

A significant proportion of process models has been captured using Enterprise Architect (20). These models can be accessed using Enterprise Architect directly or via the Web view exposed by this tool, where the user of the output is able to move on the model by using links determined in the model (for instance, to move between linked diagrams in the model or to move from the diagram to the relevant textual description). 15 process models are captured using combinations of diagrams and text, where primary information is provided to the reader via text, process diagrams are included to illustrate and visualize process flow (there was about one process diagram in the context of approximately two (1-3) pages page of text). Remaining process models (13) are mainly in textual form. The size of these 28 process models was about 50 pages; documents were managed and used by employees using document management system (Livelink). All models in the study were accessible to every employee of the organization.

All models in the study have been composed by employees of the organization, primarily by process managers and business analysts. Process workers are generally consulted during the creation of most models, but they do not directly edit them. Process diagrams in the study contained the basic set of BPMN notation: task, sub-process, event (start, end), gateway, data store, data object, sequence flow, message flow, pool, lane. Process diagrams visualize the sequence of tasks, data flows and actor involvement in general; BPMN notation is not always strictly applied.

Data of the logs included:

- the user name who looked up the model;
- code of the model;
- time stamp (date and time) of entering the page.

If the same user entered into the model many times during the day, then we considered it in the variable Use as one contact. Use by process managers was excluded from the dataset (their tasks are mainly concerned with amendment and update of models); hence, only the process model usage by process workers (employees) on their own initiative is accounted in variable Use.

4.2 Data Collection

In collecting data, we tried to aggregate information as complete as possible on all 48 process models. In order to improve the quality of the data, we collected data on every single model from at least two people. First, we interviewed the process managers (11) who participated in the creation of the models in question and who had a stake in the respective processes. Second, we interviewed project managers (5) who have been involved in process modelling and analysis projects. Finally, we received data from the document manager who is responsible for all systems and databases related to different models. During the interviews process managers provided values for each dependent variable and for each model they had been involved with.

In cases where different interviewees gave different values for a specific variable of the same model, we assess the variable directly on the specific model (but we only three such discrepancy between the assessments occurred). In addition, we directly assess the variables of five randomly chosen models (10% of the sample) in order to test the validity of the assessments given by the respondents. Procedurally, we completed the following steps for data collection:

- organised interviews with each process manager, project manager and the document manager in the organization. During this interview, we catalogued the process models that the specific manager has come in contact with; this resulted in a list of models. We also gathered data about the independent variables defined in the Section 3.2 from each interviewee who acknowledged being aware of a given model. At this stage, interviewees were not aware of the hypotheses to be tested;
- after the interviews, we added up the information gathered, and highlighted the missing information and those variables of models that received different answers from different respondents;

- we organised an additional review and examination in the form of a seminar where we went through the gathered information together with the involved employees: we corrected inconsistencies and added missing pieces of information;
- we asked the document manager to provide a table indicating which models fulfill the definition of sustained, and which do not. For confidentiality reasons, we did not get access to the full logs; instead we relied on the responses given by the document manager for each model based on the definition of sustained use and minor additional clarifications.

4.3 Findings

We first performed a descriptive analysis of the independent and dependent variables. The distribution of the independent variables – plotted in Figure 2 – shows that all values are represented in the sample. Furthermore, 26 of the 48 analysed models were used on a sustained basis, entailing that the population is well balanced with respect to the dependent variable. In order to verify the hypotheses, we applied logistic regression analysis (Hosmer and Lemeshow, 2004). The results of this analysis are given in the Table 1.

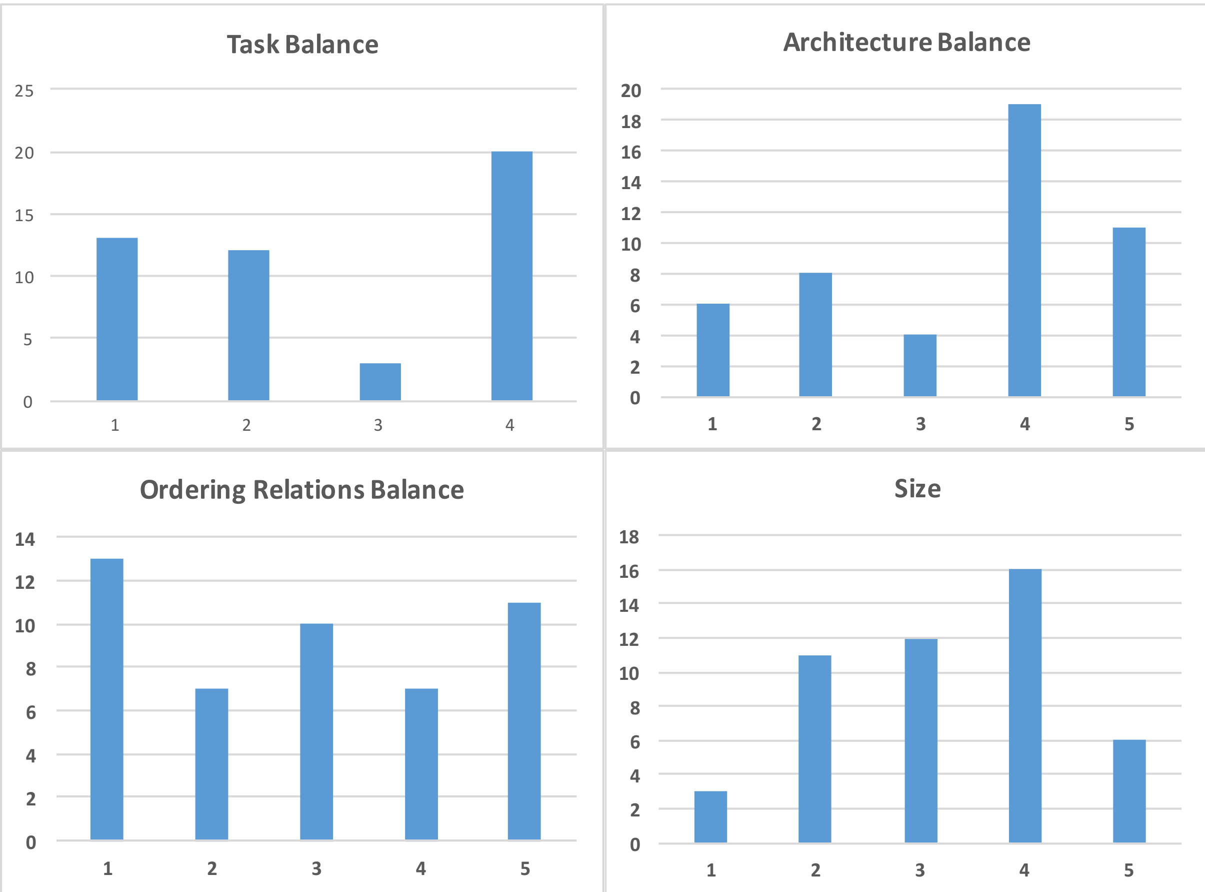

Figure 2. Distribution of variables

Looking at the Task Balance, a negative correlation with the sustained use is observed ($t=-2.451$, $p=0.0199$) at value 4 (more than 2/3 of tasks have been reflected on the process diagram); thus, we may conclude that if most of the processes are presented in a purely diagrammatic form, they are less likely to be used on a sustained basis.

With respect to the variable Architecture Balance, analysis did not highlight positive correlation with larger values of the variable (process hierarchy is presented in graphical form); at the same time the analysis shows a weak negative correlation (t=-1.739, p=0.0916) as regards sustained use of the model at value 1 (structure has been presented in the form of text) which indicates that the lack of visualized structure has a negative impact on the sustained use of the process model.

The link between the variable Ordering Relations Balance and the sustained use of the model becomes evident (p<0.05) especially at smaller values of the variable (2-4) – the ordering relations has been presented on the basis of simple diagrams. At the same time, it can be stated that as regards more complex diagrams (value of the factor characteristic 5 – the ordering relations described in more detail, using decision points), the analysis does not show a link.

The correlation between the size of the model and its sustained use (t=2.646, p=0.0125) becomes evident at value 4 (scale of tasks 500-1000). This indicates that larger models tend to be more likely to be used in a sustained manner. At the same time, the analysis did not show this association in case of very large models (more than 1000 tasks).

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.21669	0.24237	0.894	0.3780
factor(TaskBalance)2	-0.01399	0.25389	-0.055	0.9564
factor(TaskBalance)3	0.30108	0.34885	0.863	0.3945
factor(TaskBalance)4	-0.63269	0.25814	-2.451	0.0199 *
factor(ArchitectureBalance)1	-0.39854	0.22913	-1.739	0.0916 .
factor(ArchitectureBalance)2	-0.18128	0.28419	-0.638	0.5281
factor(ArchitectureBalance)3	-0.30432	0.26944	-1.129	0.2671
factor(ArchitectureBalance)4	0.02142	0.40869	0.052	0.9585
factor(OrderingRelationsBalance)2	0.68902	0.26473	2.603	0.0139 *
factor(OrderingRelationsBalance)3	0.53492	0.25490	2.099	0.0438 *
factor(OrderingRelationsBalance)4	0.61398	0.29809	2.060	0.0476 *
factor(OrderingRelationsBalance)5	0.32102	0.37166	0.864	0.3942
factor(Size)2	0.39064	0.23198	1.684	0.1019
factor(Size)3	0.17392	0.25019	0.695	0.4920
factor(Size)4	0.68143	0.25750	2.646	0.0125 *
factor(Size)5	0.28146	0.28354	0.993	0.3283
'*' means significant correlation with p-value < 0.05				

Table 1. Coefficients of the analysis.

4.4 Limitations and threats to validity

The findings of this study should be construed in the light of typical limitations and threats to validity of a case study research. A key threat to internal validity of the study is that the number of models was relatively small. To mitigate this threat to validity, we selected a large organization with models with different characteristics according to the defined variables, as reflected in histograms of the distribution of the variables. Conducting the study in other organizations with similar or larger amounts of models would naturally enhance the validity of the results.

In the interpretation of the study results, it must be taken into account that only those process models that had been integrated into the knowledge base were included in the study – process models developed for daily use by process workers and available to all employees in the organization. In the generalization of the results, the context of the study must be taken into account (experience, size, type of the organization). Further studies in organizations of different sizes are required to enhance the generalizability.

The choice of variables and the choice of discretization of these variables is a limitation of the study. These choices were driven by our objective to identify relations between the way text and diagrams are combined in a process model, and its sustained use. We acknowledge however that many other factors play a role in the sustained use of process models. Previous studies have investigated related questions, such as the relation between organizational and usability factors and use of process models (Nolte et al. 2013), the relation between internal characteristics of diagrammatic process models and understandability (Sánchez-González, 2010) and the relative understandability of purely diagrammatic and purely textual process models (Ottensooser et al. 2012). Combining these various models into a single overall model that explains sustained use from multiple perspectives is a possible direction for future work.

5 Discussion

The visual presentation of the ordering relations in diagrammatic form appears to be instrumental to sustained model use (H3: Models that include a diagrammatic presentation of the ordering relations between tasks are more likely to be used on a sustained basis). Here, the complexity of the process diagram plays an adjustment role – process diagrams should not be too complex, otherwise the diagrammatic representation of ordering relations has a lesser influence on sustained use.

The analysis demonstrated that larger models, where approximately 500-1000 tasks have been described, tend to be used more actively (H4: Larger process models are more likely to be used on a sustained basis). All process models cover one area in the respect of eTOM model and these areas are almost with the same size; for this reason, the differences in the variable Size comes due granularity of the model. With respect to size and granularity, users prefer to use models where information is presented at a more detailed level (levels 4-6 in the respect of eTOM model). At the same time, very detailed models of a technical nature (levels 6-7) are not used on a sustained manner.

The results also indicate that if most of the tasks of a model have been presented on diagrams, the model is less used on a sustained basis (H1: Process models where more tasks are visually presented (i.e. more tasks are represented as diagrammatic shapes) are more likely to be used on a sustained basis)). On the other hand, the analysis did not bring out a clear correlation between the sustained use of a process model and the existence of a diagrammatic representation of the process architecture (H2: Models that include a visual presentation of the process architecture are more likely to be used on a sustained basis); however, a lack of architecture (for example free-text format in the context of our study) shows a negative influence on the sustained use of the process model.

A summary of the above observations is given in Figure 3.

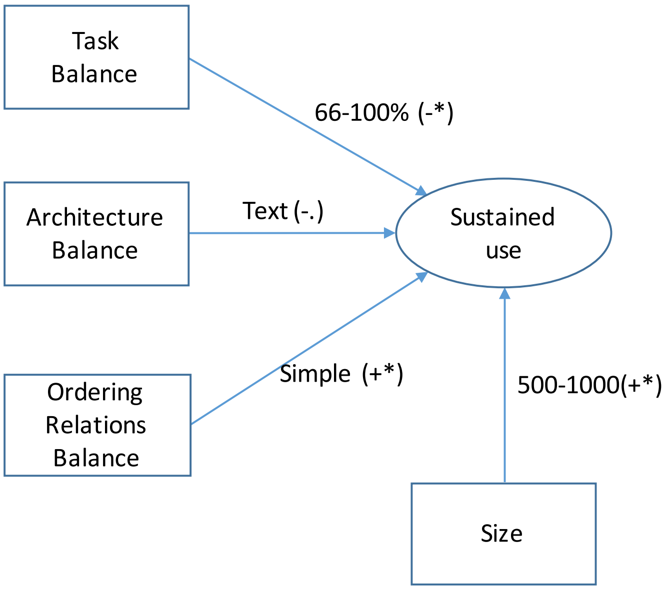

Figure 3. Results of the study

6 Conclusion

In conclusion, we can state that a typical process model that is used on a sustained basis as a knowledge base in the organization is one where the key tasks and their ordering relations are captured in diagrammatic form, while further details are left in explanatory (possibly structured) text. It is important that diagrams would provide the user who is reading the information with a process logic on a general level to which descriptions of details in the form of text will be given. The key observation here is that for smaller models, the diagrammatic representation of ordering relations between tasks is associated with more sustained use, but this does not necessarily hold when the models become larger.

A second insight is that when it comes to capturing the process architecture, the use of text to complement diagrams does not seem to play a role in the sustained use of the process model. Process hierarchy plays a vital role during process modelling, where graphical representation of the structure facilitates the decomposition of tasks; process workers are looking for a general and simple table of contents to understand the general structure of the process model.

In future work, we plan to extend the study to cover other organizations. This should enable us to extend the number of process models in the study and thus to enhance the validity and scope of the findings. Also, as stated in the limitations of the study, there are several other factors that potentially affect the (sustained) use of process models, including factors related to the type of process being captured (e.g. customer-facing versus backend processes), as well as organizational and usability factors. A direction for future work is to combine the findings of the present study with those from other studies referenced above, in order to build a broader model of (sustained) process model use. Validating such a broader model would require larger datasets, and hence this second direction of future work should go hand-in-hand with the first one.

Acknowledgments

We would like to thank the participants of the case study for their forthcoming attitude and their valuable input. This research is funded by the Estonian Research Council.

References

Bateman, J., T. Kamps, J. Kleinz, K. Reichenberger (2001). "Towards Constructive Text, Diagram, and Layout Generation for Information Presentation." *Computational Linguistics,* 27(3), 409-449.

Batini, C., C. Cappiello, C. Francalanci, A. Maurino (2009). *"Methodologies for Data Quality Assessment and Improvement." ACM Computing Surveys (CSUR)* 41(3).

Browning, T.R. (2009). "The Many Views of a Process: Toward a Process Architecture Framework for Product Development Processes." *Systems Engineering* 12(1), 69-90.

Carney, R.N. and J.R. Levin (2002). "Pictorial Illustrations Still Improve Students' Learning from Text." *Educational Psychology Review* 14(1), 5-26.

Chinosi, M. and A. Trombetta (2012). "BPMN: An Introduction to the Standard." *Computer Standards & Interfaces* 34(1), 124-134.

Dalkir, K. and J. Liebowitz (2011). *Knowledge Management in Theory and Practice.* MIT press.

Dumas, M., M. La Rosa, J. Mendling, R. Mäesalu, H. A. Reijers, N. Semenenko (2012). "Understanding business process models: the costs and benefits of structuredness." *In International Conference on Advanced Information Systems Engineering*, 31-46, Springer Berlin.

Eppler, M.J. and R.A. Burkhard (2007). "Visual Representations in Knowledge Management: Framework and Cases." *Journal of Knowledge Management* 11(4), 112-122.

Hosmer Jr, D.W. and S. Lemeshow (2004). *Applied Logistic Regression.* John Wiley & Sons

Kelly, M. (2007). "Enhanced Telecom Operations Map (eTOM) - The Business Process Framework." *TeleManagement Forum.*

Krogstie, J. (2016). *Quality in Business Process Modeling*. Springer.

Larkin, J.H. and H.A. Simon (1987). "Why a Diagram is (Sometimes) Worth Ten Thousand Words*." Cognitive Science* 11(1), 65-100.

Laue, R. and J. Mendling (2010). "Structuredness and its Significance for Correctness of Process Models." *Information Systems and E-Business Management* 8(3), 287-307.

Mendling, J., G. Neumann, W. M. van der Aalst (2007). "On the Correlation Between Process Model Metrics and Errors." *In International conference on Conceptual Modeling* 83, 173-178, Australian Computer Society, Inc.

Mendling, J., H. A. Reijers, J. Cardoso (2007). "What makes process models understandable?" *In International Conference on Business Process Management*, 48-63. Springer Berlin Heidelberg.

Mendling, J., H.A. Reijers, W. M. van der Aalst (2010). "Seven Process Modeling Guidelines (7PMG)." *Information and Software Technology* 52(2), 127-36.

Mendling, J., and M. Strembeck (2008). *"Influence factors of understanding business process models." In International Conference on Business Information Systems*, 142-153. Springer Berlin Heidelberg.

Mesarovic, M. D., D. Macko, Y. Takahara (2000). *Theory of hierarchical, multilevel, systems* (Vol. 68). Elsevier.

McMillan, E. (2002). "Considering Organisation Structure and Design from a Complexity Paradigm Perspective." *Tackling Industrial Complexity: the Ideas that Make a Difference*, 123-136.

Moody, D.L. (2005). "Theoretical and Practical Issues in Evaluating the Quality of Conceptual Models: Current State and Future Directions." *Data & Knowledge Engineering* 55(3), 243-276.

Nolte, A., E. Bernhard, J. Recker (2013). ""You've modelled and now what?" Exploring Determinants of Process Model Re-Use." *In 24th Australasian Conference on Information Systems (ACIS),* 1-11. RMIT University.

Nolte, A., E. Bernhard, J. Recker, F. Pittke, J. Mendling (2016). "Repeated Use of Process Models: The Impact of Artifact, Technological and Individual Factors." *Decision Support Systems*

Nonaka, I. (2008). *The knowledge-creating company*. Harvard Business Review Press.

Ottensooser, A., A. Fekete, H. A. Reijers, J. Mendling, C. Menictas (2012). "Making Sense of Business Process Descriptions: An Experimental Comparison of Graphical and Textual Notations." *Systems and Software* 85(3), 596-606.

Sánchez, L., F. García, F. Ruiz, M. Piattini (2010). "Measurement in Business Processes: A Systematic Review." *Business Process Management Journal* 16(1), 114-134.

Rosemann, M., T. De Bruin, T. Hueffner (2004). "A Model for Business Process Management Maturity." *In Proceedings of Australian Conference on Information Systems.*

Saarsen, T., M. Dumas (2012). "The Process Documentation Cube: A Model for Process Documentation Assessment." *International Conference on Business Process Management*, 501-512. Springer Berlin Heidelberg.

Sánchez-González, L., F. García, J. Mendling, F. Ruiz, M. Piattini (2010). "Prediction of Business Process Model Quality Based on Structural Metrics." *International Conference on Conceptual Modeling,* 458-463. Springer Berlin Heidelberg.

Schindler, P.S. and P.S. Cooper (2005). *Business Research Methods.* Mcgraw-hill.

Zur Muehlen, M. and J. Recker, (2008). "How Much Language is Enough? Theoretical and Practical Use of the Business Process Modeling Notation." *International Conference on Advanced Information Systems Engineering* 16, 465-479. Springer Berlin Heidelberg.

de Waard, A., S. Ananiadou, M. Martone, Á. Sándor, H. Shatkay (2012). "Detecting and Using Document Structure in Scientific Text." *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*, 883-884.

Yourdon, E. (1989). *Modern Structured Analysis.* Vol. 191. Englewood Cliffs, NJ: Yourdon Press.