

EVALUATION OF BPMN MODELS QUALITY

A Family of Experiments

Elvira Rolón

Autonomous University of Tamaulipas, Centro Universitario Tampico-Madero, 89336, México

Félix García, Francisco Ruiz, Mario Piattini

*Department of Information Technologies and Systems, Indra-UCLM Research and Development Institute
University of Castilla-La Mancha, Paseo de la Universidad 4, 13071, Ciudad Real, Spain*

Corrado Aaron Visaggio, Gerardo Canfora

*RCOST – Research Centre on Software Technology, University of Sannio
Palazzo ex Poste, Viale Traiano, 82100 Benevento, Italy*

Keywords: Business process models, BPMN, Empirical validation, Family of experiments, Measures.

Abstract: The design phase is of special importance in the development of a business process. This phase refers to the modeling, handling and redesigning of processes, but when maintenance tasks have to be performed, this stage may be rather complicated. It implies a heavy investment of time and resources, since it involves both technical developers and business analysts. Moreover, process modeling should permit not only the production of models which are understandable to the users, but also the early detection and correction of errors. All of this adds to the overall quality of the model. We therefore propose a set of measures with which to assess the structural complexity of conceptual business process models. Our aim is to obtain useful indicators to be used when carrying out maintenance tasks on these models, thus obtaining higher quality models by means of an early evaluation of the model's given quality properties. With the development of a family of experiments, it has been possible to discover a set of measures which may be useful in assessing the usability and maintainability of conceptual business process models.

1 INTRODUCTION

The objectives of a business process are basically (Multamäki, 2002): a) To improve the understanding of a situation so that it can then be communicated to and among the different *stakeholders* and b) to use that process as a tool to attain the goals of a process development project. Nevertheless, for business processes to be able to comply with their objective, they are constantly exposed to changes. These come about as a result of organizations' continuous improvement programmes.

Business process modeling consists of the description and visualization of processes by means of a model which represents them in formal or informal ways or in the form of a graph or diagram. Likewise, the manipulation and redesign process is

carried out in the design phase (Smith and Fingar, 2003). Business processes modeling is therefore one of the first steps towards achieving organizational goals. It is an activity which has gained great importance due to the fact that today's organizations are ever-more focused on their business process (Andersson *et al.*, 2005).

The importance that business process modeling represents has been the springboard for a variety of studies such as that of Bandara (Bandara *et al.*, 2005) in which the authors attempt to identify process modeling success factors and measures. Their empirical evidence comes from the case studies of nine process modeling projects.

Furthermore, business process modeling is of interest in a number of different fields such as that of business and software engineering. This is because its importance lies not only in the description of the

process, but in that it also usually represents a preparatory phase for activities such as (Succi *et al.*, 2000): business process improvement, business process reengineering, technology transfer and process standardization.

High-quality conceptual-modeling plays an important role in carrying out business process reengineering in particular, making it possible to detect errors at an early stage and thus correct them (Wand and Weber, 2002). In addition to this, the analysis of the level of process maturity (Bider, 2005; Francis, 2005), also forces us to have bases which facilitate modeling in the design phase. This is also true as regards the work of future maintenance.

Bearing these factors in mind, and considering the lack of studies on the possible difficulty that business process models may represent in maintenance tasks, our work takes as its main focus of attention the assessment of the structural complexity of business process models (BPMs) at a conceptual level. Our aim is to give support to business process management, allowing an early evaluation of certain quality properties of the models. It also makes the evolution of process models possible, providing as it does so objective information about maintainability, especially in those organizations which have given themselves over to ongoing improvement.

In this work we present the motivation of our research, basing it on how important it is to evaluate business process conceptual models if we are to aid their maintainability (Section 3). In addition, we present the results obtained in five experiments within the context of a family of experiments (Section 4 and 5). With these results we have attempted to obtain a set of measures which will serve as useful indicators towards the usability and maintainability of the BPMs. Finally, in Section 6 some of the conclusions drawn from this work will be put forward.

2 RELATED WORKS

The importance that the subject of business process and its modeling has acquired in the last few years has also generated great interest in the scientific community with respect to its study, analysis and measurement. However, very little can be found in literature as regards the measurement and assessment of business processes, at least at a conceptual level, which is the main topic of our research.

A recent work on measures of complexity for business process models is that presented by (Gruhn and Laue, 2006), in which the authors discuss how ideas that are already a familiar part of research into software complexity might be used to analyse the complexity of business process models.

On the other hand, the reference by (Cardoso, 2005) describes a measurement for analyzing the control-flow complexity of Web Processes and Workflows. This measurement is used at the time of design to evaluate the complexity of the design of a process before its implementation.

Having taken into consideration the studies made in the field of software engineering in (Rolón *et al.*, 2006b), we have defined a set of measures for the evaluation of conceptual business process models on the basis of the adaptation and extension of a framework defined for the modeling and measurement of the software process. In (Cardoso *et al.*, 2006) a similar type of compilation of insights from software engineering cognitive science and graph theory is provided, and the authors discuss to what extent analogous metrics of these areas can be defined for business. Finally, in (Latva-Koivisto, 2001) a collection of complexity measures for business process models found in the relevant literature was compared to a set of given criteria.

3 MEASURES FOR BUSINESS PROCESS MODELS

Our interest lies in evaluating the complexity of business processes by starting from the model which represents them at a conceptual level, and in order to do this we have used BPMN (OMG, 2006) as a modeling language. One of the reasons for the use of BPMN in our proposal is that it is one of the most widely recognized standard notations for the modeling of business processes and it is that which is most often used by both business analysts and systems analysts.

Moreover, a variety of business process modeling tools already use the BPMN metamodel and certain studies, such as that of (Mendling *et al.*, 2005), show how BPMN, in comparison with another 14 specifications includes the 15 high-level metamodel concepts defined by the author, almost in their entirety. (Wohead *et al.*, 2006) also provide a comprehensive evaluation of the capabilities of BPMN and its strengths and weaknesses when used for business process modeling. These studies, as is the case of other similar ones found in literature

experimental material in the MPNs, with the intention of confirming whether the measures validated in both first experiments might or might not be useful in evaluating the usability and maintainability of the MPNs. A detailed description of the experimental design can be consulted in (Rolón *et al.*, 2006a).

Of the five experiments conducted, it is important to emphasize that the third was carried out with the Masters students of the University of Sannio in Italy, because this implied an additional effort, which was mainly that of considering the language as a barrier since the training session was given to them in English.

In addition, it took longer to carry out this experiment, since the training was more thorough with regard to business process modeling and BPMN modeling notation. Also, subjects such as Business Process Management and tools such as the BPMS (Business Process Management Systems) were mentioned.

4.1 Subjects

The participant subjects in all the experiments had similar knowledge as far as process modeling was concerned. All the groups were nevertheless given a training session to ensure that they were conscious of the aspects that we were attempting to evaluate. A summary of the groups who participated in each experiment can be seen in Table 2.

Table 2: Participant groups in the experiments.

Exp.	Group	N° Sub.	Profiles
1	UCLM (Spain)	27	PhD students, Students, Research assistants and Lecturers in computer engineering.
2	UAT (Mexico)	31	Students of Master in Information Systems.
3	University of Sannio (Italy)	37	Masters Students in: <ul style="list-style-type: none"> • Software Technology • Software Management and Technology • Computer Science Technology for Organizational Management and Knowledge.
4	HGCR	6	Administrative Staff and health professionals.
5	UCLM (Spain)	8	PhD Students

4.2 Material

In all cases the material consisted of ten BPMs represented with BPMN, whose structural characteristics and dimensions were different from each other; that is to say, models with different degrees of complexity were selected. These were obtained by varying the value of the measures in each model. Our intention, upon choosing models with different dimensions, was to determine the

influence of the complexity of the model upon different subjects such as business analysts and software engineers, who are the main focus of our study.

Moreover, two questionnaires were formulated for each of the aforementioned models. The first consisted of a series of six questions related to the model's understandability, and the second proposed a series of modifications to be carried out on the model, such as evaluating the complexity of the process models presented. In addition, at the end of each questionnaire a question was included, whereby the subjects were asked to assess the complexity of the models presented in a subjective manner. The material also included an example of a solution which showed how the exercises should be done. An example of the material used can be found in (Rolón *et al.*, 2007).

4.3 Objective

Using the GQM template (Goal Question Metric) the goal in all the experiments is defined as being:

- *To Analyse* measures of BPM structural complexity
- *To evaluate* them *as regards* their capability of being used as indicators of business process model understandability and modifiability
- The researchers' point of view
- *The context* of PhD students, research assistants and lecturers in software engineering (Exp. 1); Students of the Masters degree in Information Systems (Exp. 2); Post graduate students (Exp. 3); Administrative staff and health professionals (Exp. 4) and PhD Students (Exp. 5).

4.4 Variables and Hypothesis

Within the context of the family of experiments the same variables have been considered. These are:

- Independent variables: those which correspond with the proposed measures, which is to say the base measures and derived measures already described.
- Dependent variables: those which relate to the understandability and modifiability of the BPMs, which will be measured according to the subjects' efficiency when performing the tasks, which is calculated as the ratio between the number of right answers and the time.

The dependent variables were measured through the subjects' response times when carrying out the required tasks, the success rate in the questions relating to the understandability and modifiability

tasks of the model, the subjective evaluation with respect to the complexity of the models, and also the efficiency of the successes in relation to the times.

The hypotheses proposed with respect to the objective of our investigation are the following:

- Null hypothesis, H_{0u} : There is no significant correlation between the structural complexity measures and the understandability.
- Alternative hypothesis, H_{1u} : There is a significant correlation between the structural complexity measures and the understandability.
- Null hypothesis, H_{0m} : There is no significant correlation between the structural complexity measures and the modifiability.
- Alternative hypothesis, H_{1m} : There is a significant correlation between the structural complexity measures and the modifiability.

5 ANALYSIS OF RESULTS

Once the individual experiments had been carried out, a global analysis of the results took place within the context of the family of experiments in order to determine whether or not they had attained the general objective of the empirical evaluation. In order to do this, a descriptive analysis and a statistical analysis of the data collected in all five experiments were carried out. A description of both analyses is presented in the following sections.

5.1 Descriptive Analysis

Having taken into account that the dependant variables are those which are relative to the understandability and modifiability of the model, a summary of the data obtained from the results of each experiment was carried out.

Each variable was measured according to the response times, the success rate in the required tasks, the subjective evaluation that the subjects made, and the efficiency of the successes in relation to the time. We shall now present the results which were obtained after having analyzed each of these aspects.

Table 3 shows a summary of the results obtained from the experiments which were carried out, with regard to the time (in minutes) that the subjects needed to respond to the tasks relating to understandability and modifiability.

Upon analyzing the time taken by the subjects to carry out the required tasks and upon obtaining the mean times of the five experiments, it can be observed in Table 3 that, in the case of the tasks

relating to the model’s understandability, the subjects took more time with models 5, 7 and 10, whilst they took more time to carry out the requested modifications with models 3, 4 and 7.

Table 3: Answer times.

BPM	Understandability Times					Modifiability Times				
	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
1	121	181	230	178	132	327	323	325	316	247
2	166	159	218	134	148	401	454	450	305	581
3	185	182	228	174	189	291	384	418	348	773
4	149	175	214	164	362	306	2546	1509	420	272
5	280	248	295	337	293	375	438	384	519	407
6	279	220	270	142	205	345	409	383	196	540
7	221	230	307	145	284	416	473	419	453	405
8	211	193	225	143	218	305	392	416	284	379
9	187	240	225	101	241	392	362	343	306	527
10	238	247	277	243	187	319	454	461	319	364

These results can be better appreciated in Figure 2, in which the results that appear in Table 3 are grouped according to the average of the results obtained in each experiment in order to discover both which models the subjects took most time to respond to, and the models’ understandability and modifiability.

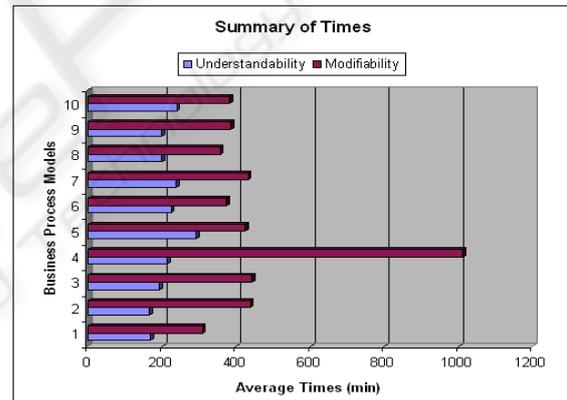


Figure 2: Summary of the average times.

The descriptive analysis relating to success, subjective evaluation and efficiency was carried out in a similar manner. With regard to success in the required tasks, the results of the five experiments show that models 3, 4 and 7 were those which led to the subjects producing the greatest amount of errors in the responses related to understandability, whilst in the tasks related to modifiability the majority of mistakes were made with models 3, 7 and 10 (Fig.3).

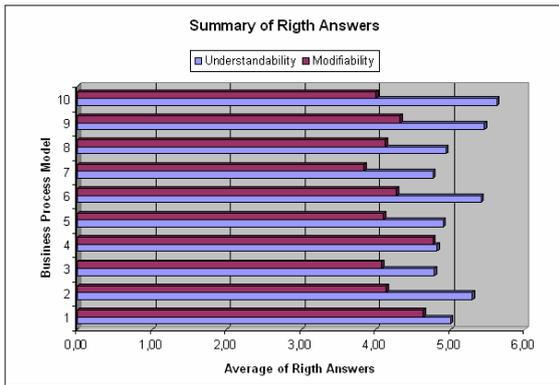


Figure 3: Summary of right answers.

As regards the subjects' subjective evaluation of the complexity of understandability of the models presented, models 5, 6, 9 and 10 were evaluated as being the most complex, whilst in the case of modifiability the most complex models were 5, 7 and 10 (Figure 4).

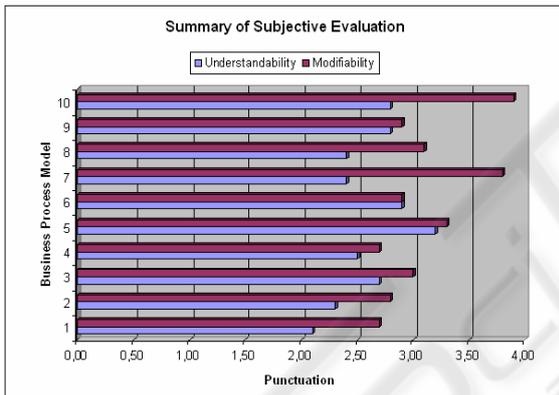


Figure 4: Subjective evaluation chart.

In this case, models 5 and 10 coincide in both tasks as being the most complex models according to the subjects' criterium, and these results coincide with the values of the measurements of each of the MPNs in which the models of greatest structural complexity were 7, 9 and 10.

Finally, the efficiency of the successful responses to the tasks in relation to the time taken to carry them out was obtained from the statistical analysis of the dependant variables.

Figure 5 shows the mean results of the five experiments and, as can be seen, the models which have the lowest level of efficiency as regards understandability were 5, 7 and 10. Those which had the lowest level as regards modifiability were 2, 5 and 7.

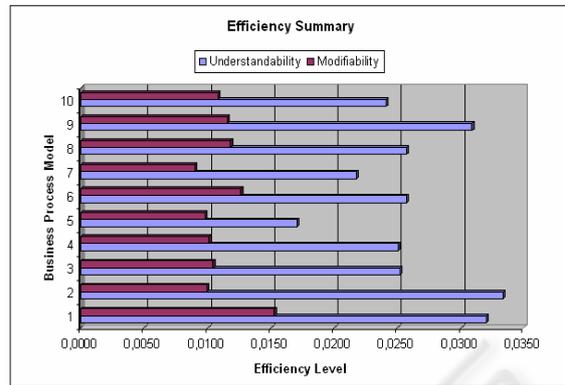


Figure 5: Efficiency chart.

5.2 Statistical Analysis

The summary of the mean times, successes, subjective evaluation and efficiency, both for the understandability and for the modifiability tasks, along with the values of the measurements of the business process models were used to carry out a statistical analysis.

Initially, a correlation analysis of the values of the measurements as regards the response times and the number of successful responses from the results obtained in the five experiments was carried out by following the suggestions of (Perry *et al.*, 2000), (Wohlin *et al.*, 2000), (Juristo and Moreno, 2001), (Ciolkowski *et al.*, 2002) and (Briand *et al.*, 1995).

In order to prove whether the distribution of the data obtained was normal, the Kolmogorov-Smirnov test was applied. As a result of this it was obtained that the distribution was not normal, and for this reason we decided to use a non-parametrical statistical test such as the Spearman correlation coefficient with a level of significance of $\alpha = 0.05$ which indicates the probability of rejecting the null hypothesis when it is certain (type I error). That is to say, a confidence level of 95% exists.

The Spearman correlation coefficient was used to separately correlate each of the measurements with the dependant variables as regards each of the aspects evaluated in the descriptive analysis.

The results of the correlation analysis of the five experiments to discover understandability and modifiability times obtained that the measures which correlated with the response times for the tasks relating to understandability, and which were validated in at least two of the five experiments were: NIMsE (Number of Intermediate Message Events), NEDDB (Number of Exclusive Decision Data-Based), TNIE (Total Number of Intermediate Events of the model), NSFE (Number of Sequence

Flows from Events) and TNE (Total Number of Events of the model).

With regard to modifiability, the NEDEB (Number of Exclusive Decision Event-Based) and CLA (Connectivity Level between Activities) measures were validated in experiments 2 and 3.

The analysis of the correlations with regard to successes, subjective evaluation and efficiency was carried out in a similar manner. With regard to the correlations of the measures as regards successes in the required tasks, only the TNSE (Total Number of Start Events of the model) was validated in two of the five experiments as far as successes in understandability were concerned. In the case of successes in the modifiability tasks, of the various correlation measurements the NDOIn (Number of Data Object-In of the process) and TNEE (Total Number of End Events of the model) measures were only validated in two experiments.

In the efficiency analysis (Table 4), the measures validated in at least two of the five experiments with regard to understandability were: NIMsE, NEMsE (Number of End Message Events), NEDDB, NSFE, TNE y NSFL. In relation to modifiability, the following measures were validated: NCS (Number of Collapsed Sub-processes), TNCS (Total Number of Collapsed Sub-processes of the model), NEDEB and CLA.

Table 4: Efficiency correlations.

Measure	Efficiency									
	Understandability					Modifiability				
	Exp-1	Exp-2	Exp-3	Exp-4	Exp-5	Exp-1	Exp-2	Exp-3	Exp-4	Exp-5
TNIE	✓									
NIMsE	✓	✓				✓				
NEMsE	✓	✓				✓				
NEDDB	✓	✓								
NSFE	✓	✓								
TNE	✓	✓								
TNEE		✓								
TNA		✓								
NSFL			✓	✓						
NPF				✓						
NDOOut				✓						
TNDO				✓						
PDOOut				✓						
NCS					✓					
TNCS					✓					
NSFG						✓				
TNG						✓				
NEDEB						✓		✓		
CLA						✓		✓		
NITE								✓		
PDOIn										✓
PDOOut										✓

Finally, upon analyzing the correlations as regards the subjective evaluation that the subjects made of the complexity of the models, we obtained the result that measures which were validated in at least two of the five experiments only existed in the case of modifiability.

These measures were: TNE, TNA (Total Number of Activities), NENE (Number of End None Events), NT (Number of Tasks), NSFL, TNEE, TNT (Total Number of Tasks in the model), NEDDB, NSFG (Number of Sequence Flows from Gateways), and TNG (Total Number of Gateways of the model).

6 CONCLUSIONS AND FURTHER WORK

This work shows the results obtained from carrying out a family of experiments. This was done with the objective of analyzing and evaluating the structural complexity of business process models. The analysis took place at a conceptual level of the models and used a set of measures which were defined on the basis of the BPMN standard notation.

As a result of this family of experiments we have obtained a significant set of measures that could serve as indicators towards the maintainability of the business processes models expressed in BPMN. From the 60 defined measures, 22 have been correlated with BPMN model understandability or modifiability in at least one of the experiments. With these 22 measures, which have been validated as results in order to measure the efficiency in the accomplishment of the tasks (dependent variable), it is possible to reject the formulated null hypotheses.

Future experimentation will focus upon evaluating this set of measures which we consider relevant from the results obtained in this first family of experiments. To achieve this, new material will be designed in which the validated metrics will be grouped into three categories (participants and roles, activities and control flows and decision nodes). Specific material will be designed for each group. This will allow us to obtain a higher variation of the complexity in each subgroup of models which may confirm the usefulness of the measures validated in the first family. Other aspects to be tackled in the future are:

- We shall conduct new experiments with the aim of analyzing two further sub-characteristics of the quality of the model, namely analyzability and ease of learning, which are respectively related to usability and maintainability.
- Furthermore, the development of business process models within a company in the health sector is being carried out, which will allow us to use experimental models of real cases.

ACKNOWLEDGEMENTS

This work has been partially financed by the INGENIO Project (Junta de Comunidades de Castilla-La Mancha, Consejería de Educación y Ciencia, PAC08-0154-9262), MECENAS (Junta de Comunidades de Castilla-La Mancha, Consejería de Educación y Ciencia, PBI06-0024) and ESFINGE Project (Ministerio de Educación y Ciencia, Dirección General de Investigación/Fondos Europeos de Desarrollo Regional (FEDER), TIN2006-15175-C05-05).

REFERENCES

- Andersson, B., Bider, I., et al., 2005. "Towards a Formal Definition of Goal-Oriented Business Process Patterns." *Business Process Management Journal* 11(6): 650-662.
- Bandara, W., Gable, G. G., et al., 2005. "Factors and measures of business process modelling: model building through a multiple case study." *European Journal of Information Systems* 14: 347-360.
- Bider, I., 2005. "Choosing Approach to Business Process Modeling Practical Perspective." *Journal of Conceptual Modeling*(34).
- Briand, L., El Emam, K., et al., 1995. "Theoretical and Empirical Validation of Software Product Measures." *International Software Engineering Research Network Technical Report ISERN-95-03*.
- Cardoso, J., 2005. *How to Measure the Control-flow Complexity of Web Processes and Workflows*. Workflow Handbook. WfMC. Lighthouse Point, FL, USA: 199-212.
- Cardoso, J., Mendling, J., et al., 2006. *A Discourse on Complexity of Process Models*. BPM 2006 Workshops, Workshop on Business Process Design, Viena, Austria, LNCS. J. Eder and Dustdar, S. e. a. (Eds). pp. 115-126.
- Ciolkowski, M., Shull, F., et al., 2002. *A Family of Experiments to Investigate the Influence of Context on the Effect of Inspection Techniques*. Proceedings of the 6th International Conference on Empirical Assessment in Software Engineering (EASE), Keele (UK). (Eds). 48-60.
- Francis, J., 2005. *Managing BPM, BP Trends*.
- Gruhn, V. and Laue, R., 2006. *Complexity Metrics for Business Process Models*. Proceedings of 9th International Conference on Business Information Systems (BIS'06), Klagenfurt, Austria. W. Abramowicz and Mayr, H. C. (Eds). 1-12.
- Havey, M., 2005. *BPMI Standards: BPMN and BPML. Essential Business Process Modeling*. A. Odewahn and O'Brien, M. USA, O'Reilly: 143-173.
- Juristo, N. and Moreno, A., 2001. *Basics of Software Engineering Experimentation*, Kluwer Academic Pub.
- Latva-Koivisto, A. M., 2001. *Finding a complexity measure for business process models*. Systems Analysis Laboratory, Helsinki University of Technology.
- Mendling, J., Neumann, G., et al., 2005. *A Comparison of XML Interchange Formats for Business Process Modeling*. Workflow Handbook 2005. L. Fischer. Lighthouse Point, Florida USA: 185-198.
- Multamäki, M., 2002. *Objective-driven planning of business process modeling*. Department of Industrial Engineering and Management, Helsinki University of Technology.
- OMG, 2006. *Business Process Modeling Notation*. Object Management Group. <http://www.bpmn.org/Documents/OMG%20Final%20Adopted%20BPMN%201-0%20Spec%2006-02-01.pdf>.
- Perry, D., Porte, A., et al., 2000. "Empirical Studies of Software Engineering: A Roadmap." *Future of Software Engineering*: pp 345-355.
- Rolón, E., Garcia, F., et al., 2006a. *Métricas para la Evaluación de Modelos de Procesos de Negocio*. 9º Workshop Iberoamericano de Ingeniería de Requisitos y Ambientes de Software (IDEAS'06), La Plata, Argentina. (Eds). 419-432.
- Rolón, E., Garcia, F., et al., 2007. *Experimento Exploratorio para la Validación de Medidas para Modelos de Procesos de Negocio*. VI Jornadas Iberoamericanas de Ingeniería del Software e Ingeniería del Conocimiento (JIISIC'07), Lima, Perú. (Eds). 283-292.
- Rolón, E., Ruiz, F., et al., 2006b. "Applying Software Metrics to evaluate Business Process Models." *CLEI-Electronic Journal* Vol. 9(1, Paper 5).
- Smith, H. and Fingar, P., 2003. *Business Process Management: The Third Wave*. USA, Meghan-Kiffer Press.
- Succi, G., Predonzani, P., et al., 2000. *Business Process Modeling with Objects, Costs and Human Resources*. Systems Modeling for Business Process Improvement. D. Bustard et al, Artech House: 47-60.
- Wand, Y. and Weber, R., 2002. "Research Commentary: Information Systems and Conceptual Modeling - A Research Agenda." *Information Systems Research* 13(No. 4): 363-376.
- Wohed, P., van der Aalst, W. M. P., et al., 2006. *On the Suitability of BPMN for Business Process Modelling*. Business Process Management (BPM'06), Vienna, Austria, Springer-Verlag Berlin Heidelberg. S. Dustdar et al (Eds). 161-176.
- Wohlin, C., Runeson, P., et al., 2000. *Experimentation in Software Engineering: An Introduction*. Kluwer Academic Publishers.