Analysis and Validation of Control-Flow Complexity Measures with BPMN Process Models

Elvira Rolón[1], Jorge Cardoso[2], Félix García[1], Francisco Ruiz[1], Mario Piattini[1]

[1]Alarcos Research Group, University of Castilla La Mancha
Paseo de la Universidad No. 4, 13071 Ciudad Real, Spain
+34 926 295300 Ext. 3747
{elvira.rolon, felix.garcia, francisco.ruizg, mario.piattini}@uclm.es

[2]University of Madeira, Department of Mathematics and Engineering
9050-390 Funchal, Portugal
jcardoso@uma.pt

Abstract. Evaluating the complexity of business processes during the early stages of their development, primarily during the process modelling phase, provides organizations and stakeholder with process models which are easy to understand and easy to maintain. This presents advantages when carrying out evolution tasks in the process models - key activities, given the current competitive market. In this work, we present the use and validation of the CFC metric to evaluate the complexity of business process models developed with BPMN. The complexity is evaluated from a control-flow perspective. An empirical evaluation has been carried out, in order to demonstrate that the CFC metric can be useful when applied to BPMN models, providing information about their ease of maintenance.

Keywords. Business process models, BPMN, Measurement, Validation.

1 Introduction

Business process modeling is the first step towards the achievement of organizational goals, because its importance resides not only in the description of the process, but in that it also usually represents a preparatory phase for activities such as business process improvement, business process reengineering, technology transfer and process standardization [1].

But in all these activities the business process models are managed by different stakeholders (business process analysts, domain experts, technical analysts, software developers, among others). Therefore, one of their main purposes is support communication between the stakeholders and to fulfill this purpose, the business process models should be easy to understand and easy to maintain. High complexity in a process has several undesirable drawbacks: it may result in bad understandability, as well as in errors, defects, and exceptions, thus leading to the need for more time to develop, test and maintain the processes. That being so, the first step towards reducing complexity is to recognise its existence, and then measure it.

In this context, Cardoso [2] defines process complexity as the degree to which processes are difficult to analyze, understand or explain. He defines the Control-Flow Complexity (CFC) metric for analyzing the degree of complexity of business process.

On the other hand, in previous works [3, 4, 5] a set of measures for the evaluation of models developed with BPMN (Business Process Modeling Notation) [6] have been defined and empirically validated. They are based on the measurement of structural properties of these models. As a result of the empirical validation, several measures were correlated with usability and maintainability. However, we believe that since the measures proposed in both works are based on the analysis of the complexity of the models of business processes, it would be important to analyze the influence of CFC metric in the complexity of BPMN models from a control-flow perspective.

On this basis in this paper the analysis and empirical validation of the influence of CFC metric on usability and maintainability of BPMN models is presented. This is done by using the data obtained from two families of experiments which were carried out earlier to validate measures of structural complexity of BPMN models [5].

This paper is organized as follows. Section 2 provides an overview of related works and Section 3 introduces the CFC metric, presenting an example of computation in a business process model with BPMN. Section 4 provides an overview of the two families of experiments carried out to empirically validate measures for BPMN process models. Section 5 presents the analysis of results in the validation of the CFC metric, using the data obtained from the experiments with the BPMN models. Finally, conclusions are outlined in Section 6.

2   Related Work

The complexity and other characteristics and aspects of business processes models (BPMs) such as size, density, cohesion, coupling (among others) have been analyzed and measured by some researchers who agree that, as with software processes, business processes should minimize their complexity, so that they provide adequate support to the various stakeholders. The majority of measures proposed for analyzing the complexity of BPMs, have their origin in, or are adaptations of, measures previously defined for the evaluation of software. For instance, in works such as [7, 8, 9, 10] this topic is analyzed and software complexity metrics (or others characteristics of software) are analyzed and compared with corresponding metrics for BPMs.

However, it is important to highlight the different perspectives from which the complexity of business process has been evaluated. For instance, Gruhn and Laue [11] has adopted complexity measures based on cognitive weights, assuming that these are good ways of measuring the difficulty of understanding the BPMs elements. Mendling, in [12], investigates how the complexity of models influences errors observed in a wide range of existing BPMs, developing a set of metrics to measure the probability of error and testing 28 business process metrics as error predictors on a set of over 2000 process models from different samples [13, 14].

In [7] , some ideas from McCabe's cyclomatic complexity are borrows and CFC metric are defined, which can be used to analyze the complexity of business processes

from a work-flow perspective (see section 3). Moreover, there is not much published work about empirical validation of the measures proposed. In a recent study, the use of BPMN elements in practice and their implications were analyzed [15].

In this context we use the CFC metric defined by Cardoso to evaluate the control-flow complexity of several BPMs developed with BPMN standard notation [6]. In this context, both works presented in [2, 3] coincide in the study of the metrics defined for software processes and their extension and adaptation to business processes, on the basis of the similarity that is present at the time of evaluating their complexity. In addition, both share the idea that when information with regard to process model complexity is obtained, the model is easier to understand and modify in order to perform maintenance tasks and process quality improvement is more likely to occur.

3   Control-Flow Complexity Measure

An important aspect to consider in the quest to achieve effective process management is the complexity analysis of processes. This is the aim of the CFC metric, whose definition is based on the hypotheses that the control-flow behaviour of a process is affected by constructs such as splits and joins.

As a result, the formula developed captures the complexity of the XOR-split, OR-split, and AND-split constructs as follows:

- XOR-split Control-flow Complexity. Determined by the number of mental states that are introduced with this type of split. The function $CFC_{XOR\text{-}split}(a)$, where a is an activity, computes the control-flow complexity of the XOR-split a. For XOR-splits, the control-flow complexity is simply the fan-out of the split.

$$CFC_{XOR\text{-}split}(a) = \text{fan-out}(a) \qquad (1)$$

- OR-split Control-flow Complexity. Determined by the number of mental states that are introduced with the split. For OR-splits, the control-flow complexity is $2^{(n-1)}$, where n is the fan-out of the split.

$$CFC_{OR\text{-}split}(a) = 2^{\text{fan-out}(a)-1} \qquad (2)$$

- AND-split Control-flow Complexity. For an AND-split, the complexity is simply 1. The process designer needs only to consider and analyze one state that may arise from the execution of an AND-split construct, since it is assumed that all the outgoing transitions are selected and followed.

$$CFC_{AND\text{-}split}(a) = 1 \qquad (3)$$

Mathematically, the Control-Flow Complexity metric is additive. This is done by simply adding the CFC of all the split constructs and is calculated as follows:

$$CFC = \sum CFC_{XOR\text{-}split}(a) + \sum CFC_{OR\text{-}split}(a) + \sum CFC_{AND\text{-}split}(a) \qquad (4)$$

The greater the value of the CFC, the greater the overall structural complexity of a process will be. CFC analysis seeks to evaluate complexity without direct execution of processes.

3.1 Example of CFC Calculation

Figure 1 shows the process of online ticket purchase modelled with BPMN. This process states that the customer has to choose between different outgoing paths once the process is initiated. Basically, it consists of selecting the type of tickets that is being sought on the Web, and for each option there are diverse outgoing paths. The process finishes when the purchase is carried out satisfactorily or when the customer cancels the purchase process. As example, the results of the Control-Flow Complexity calculation carried out in the process of Figure 1 are shown in Table 1.
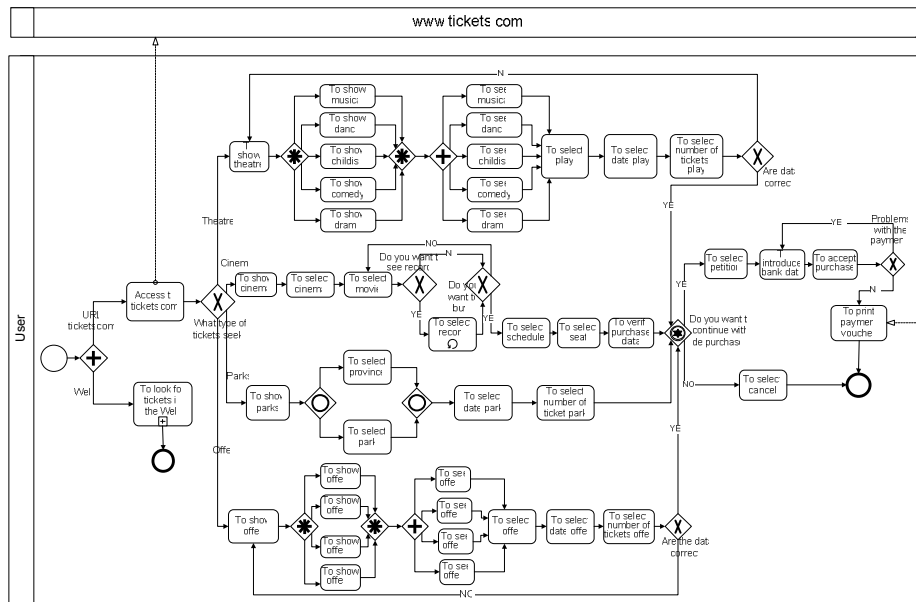

Fig. 1. Online Ticket purchase process

The calculation of the overall CFC value basically consists of adding the individual CFC of each split. The value obtained gives an indication of the complexity of the ticket purchase process. It would, however, be important to have a parameter or rank of values with which we could decide whether or not the resulting value represents a greater degree of the complexity of the model.

Table 1. Values of CFC metrics for the process from Fig. 1

CFC Metric	Value	CFC Metric	Value
$CFC_{XOR\text{-}split}$(tickets type?)	4	$CFC_{XOR\text{-}split}$(shows offer type?)	4
$CFC_{XOR\text{-}split}$(to select theatre type)	5	$CFC_{XOR\text{-}split}$(offer data correct?)	2
$CFC_{XOR\text{-}split}$(theatre data correct?)	2	$CFC_{OR\text{-}split}$(to select province/park)	2^2-1
$CFC_{XOR\text{-}split}$(wants to see record?)	2	$CFC_{AND\text{-}split}$(to access the web)	1
$CFC_{XOR\text{-}split}$(wants to buy?)	2	$CFC_{AND\text{-}split}$(to select theatre type)	1
$CFC_{XOR\text{-}split}$(wants continue the purchase?)	2	$CFC_{AND\text{-}split}$(to select offer type)	1
$CFC_{XOR\text{-}split}$(payment' problems?)	2		
CFC (Purchase on-line)			**31**

With this example, it has been possible to verify that CFC metrics can be used to measure the complexity of BPMN models, fulfilling their objective of analyzing the control-flow complexity of business processes.

4 Measures for BPMN Models

Our work consists of analyzing and empirically validating the CFC metric on the basis of previous work carried out for the evaluation of models developed with BPMN. Hence, in this section a summary of the previous works is included to place the results presented in this paper into context.

With the aim to evaluate the complexity of business processes by starting from the model which represents them at a conceptual level, in previous work a set of measures were defined in which can be grouped into two categories: Base Measures and Derived Measures. Table 2 shows an example of some derived measures (the complete list of measures can be found in [3]).

Table 2. Derived measures for BPMN models

Measure	Definition	Formula
TNE	Total Number of Events of the Model	TNE = NTSE + NTIE + TNEE
TNG	Total Number of Gateways of the Model	TNG=NEDDB+NEDEB+NID+NCD+NPF
TNDO	Total Number of Data Objects	TNDO = NDOIn + NDOOut
CLA	Connectivity Level between Activities	$CLA = \dfrac{TNT}{NSF}$
PDOPOut	Proportion of Data Object as Outgoing Product and the total of Data Objects	$PDOPOut = \dfrac{NDOOut}{TNDO}$
PDOTOut	Proportion of Data Object as Outgoing Product of Activities of the Model	$PDOTOut = \dfrac{NDOOut}{TNT}$

The following subsections present the research context and an overview of the two families of experiments which were conducted to empirically validate the relationship between the proposed measures and usability and maintainability of BPMN models.

4.1 Research Context

The objective of carrying out families of experiments to empirically validate the measures presented in [3] was to discover which of the measures defined could provide useful and objective information about the external quality of business process models. They focused mainly on two characteristics of the ISO 9126 external quality: Usability (understandability) and Maintainability (modifiability). The results obtained in the empirical validation of the first family are presented in [5].

Initially, the measures were theoretically validated according to the Briand *et al.* theoretical framework [16]. As a result, it was possible to group them in relation to the different properties of structural complexity (size, coupling and complexity) they evaluate (Fig. 2).
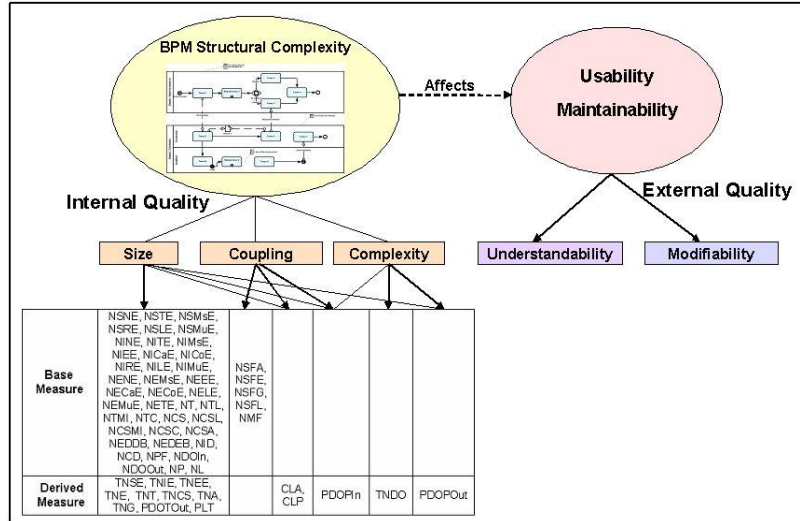
Fig. 2. Relationship between structural complexity and quality attributes.

For the empirical validation of the measures defined, a set of experiments was planned and designed. Using the GQM template (Goal Question Metric) [17], the research objectives were defined as *analyse* measures of the structural complexity BPM's, *with the purpose of* evaluating them *as regards* their capability of being used as indicators of understandability and modifiability of BPMs, *in the context of* PhD students, research assistants and others.

The hypotheses proposed with respect to the objective of investigation were to ascertain if there is a significant correlation between the measurements of structural complexity and the understandability and modifiability subcharacteristics. The independent variables were the measures defined for BPMN models and the dependent variables were those relating to the understandability and modifiability of BPMs. These latter were measured according to answer times, number of correct answers, subjective evaluation and the efficiency in the accomplishment of the tasks.

4.2 First Family

The first family of experiments was composed of five experiments. The experimental design used was the same for all five experiments. They were thus therefore carried out in similar circumstances and in the same context, that is to say, applying the same research objective, hypotheses and variables. In the experimental design a within-subjects design was carried out, in which all the subjects had to do all the tests.

Material composed of ten randomly ordered BPMN models was handed out to each subject. These BPMN models had different structural characteristics; that is to say, different degrees of complexity; they included two questionnaires formulated for each process models. The first one was related to the understandability, and the second to the modifiability. A subjective question about the complexity of the model was also included. A more detailed description of the material can be found in [4].

The subjects (Table 3) were chosen as suited our purposes due to all of them had enough knowledge about modelling to carry out the experimental tasks. To leverage their knowledge about process modelling a training lesson was carried out before the experiments run. This session consisted of an introduction to business processes and training about the BPMN standard notation.

Table 3. Participant groups in the first family of experiments.

Exp	Group	Nº Sub.	Profiles
1	UCLM (Spain)	27	PhD students, Research assistants and Lecturers in computer engineering.
2	UAT (Mexico)	31	Master's Students in Information Systems.
3	University of Sannio (Italy)	37	Master's Students in: • Software Technology • Software Management and Technology • Computer Science Technology for Organizational Management and Knowledge.
4	HGCR (Spain)	6	Health professionals.
5	UCLM (Spain)	8	PhD Students

4.3 Second Family

The second family of experiments included the development of five experiments. In the experiments of the second family, the understandability and modifiability aspects were also analyzed the difference being that in this instance separate experiments were designed to analyze each aspect. In this way, of the five experiments included in the second family, the first three were carried out to analyze the understandability of the models and in both last experiments the modifiability was evaluated.

The experimental material to analyze the understandability consisted of fifteen BPMN models with different structural characteristics and degrees of complexity. For each model a questionnaire with three questions related to the understandability of the process model was elaborated. In order to analyze the modifiability, the experimental material consisted of twelve BPMN models and a questionnaire with two requirements of modifications for each model. Moreover, for all cases the subjects answered a subjective question regarding the complexity of the process model [5].

As with the first family, the participant subjects in the second family (Table 4) received a training session about BPMN.

Table 4. Participant groups in second family of experiments.

Exp.	Group	Nº Sub.	Profiles
1 (U)	UCLM (Spain)	22	PhD students and students in computer engineering.
2 (U)	UCLM (Spain)	40	Students of 4th year in Computer engineering.
3 (U)	UCLM (Spain)	9	PhD students and students in computer engineering.
4 (M)	University of Bari (Italy)	29	Students in computer engineering
5 (M)	UAT – (Mexico)	15	Master's Students in Information Systems.

In the context of this paper, now we have described the families of experiments, the following step is to present the descriptive and statistical analysis that was carried out to validate the CFC metric. This was done by taking the data obtained concerning the dependent variables, to determine the feasibility of using the CFC metric to measure the structural complexity of business process models developed with BPMN.

As both the CFC metric and the measures proposed in [3] evaluate the structural complexity of BPMs, the same experimental design, hypothesis and variables in the two families of experiments can be stated. Consequently, the data obtained in the two former empirical studies can be used to analyze whether a correlation between the CFC metric and the maintainability of the BPMN models exists. The results of the CFC validation are shown next.

5   Analysis and Validation of CFC Metric

The CFC metric presented in section 3 has been validated previously, by analyzing its values in different process models represented with the METEOR workflow management system and with respect to the subjective evaluation of such models by process designers [18]. As a result, the authors conclude that the CFC metric is highly correlated with the control-flow complexity of processes, and therefore can be used by business process analysts and process designers to analyze the complexity of processes and, to develop simpler processes, if possible.

In the current work, the aim is to corroborate if the CFC metric can be used to analyze the complexity of business process developed with a standard notation such as BPMN. Our goal is also to provide some insight based on objective data about the metric's influence on the easiness of understanding and modifying of BPMN models. With this objective in mind the stated research hypotheses are:
- Null hypothesis, H_{0u}: There is no significant correlation between the CFC metric and the understandability.
- Alternative hypothesis, H_{1u}: There is a significant correlation between the CFC metric and the understandability.
- Null hypothesis, H_{0m}: There is no significant correlation between the CFC metric and the modifiability.
- Alternative hypothesis, H_{1m}: There is a significant correlation between the CFC metric and the modifiability.

5.1   Descriptive Analysis

Initially, in order to carry out the descriptive analysis, the values of the CFC metrics of the models used in all the experiments were obtained (Table 5). The global values of the CFC metrics reflect the degree of complexity of control-flows between process models. For example, process models 7 and 10 of the first family have the highest values of CFC. It is therefore possible to state that they have greater structural complexity than process model number 1. In the second family, the highest CFC values were obtained with the first five models, as these models contained more gateways and sequence flows.

In both families of experiments, the dependent variables were measured 1) based on the times that subjects needed to carry out the required tasks, 2) the percentage of correct answers, 3) the subjective evaluation with regard to the complexity of the models, and 4) the efficiency of the answers (this is calculated as the ratio between the number of correct answers and the time needed to respond).

Table 5. Values of CFC metric in experimental material.

Process Model	1st Family			2nd Family	
	Exp. 1, 2 and 5	Exp. 3	Exp. 4	Exp. 1, 2 and 3	Exp. 4 and 5
1	2	2	2	25	25
2	2	2	2	25	25
3	6	6	6	33	33
4	8	8	8	31	
5	7	7	7	2	2
6	6	6	6	7	
7	11	11	8	9	9
8	2	2	3	5	5
9	2	2	8	8	8
10	14	15	15	0	0
11				2	
12				4	4
13				8	8
14				4	4
15				0	0

Table 6 shows a summary of the results obtained from all 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.

Table 6. Values of answer times - First Family

Process Model	First Family										Process Model	Second Family				
	Understandability Times					Modifiability Times						Underst. - Times			Mod. - Times	
	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	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	1	135	137	178	308	137
2	166	159	218	134	148	401	454	450	305	581	2	137	124	137	331	124
3	185	182	228	174	189	291	384	418	348	773	3	238	245	331	253	245
4	149	175	214	164	362	306	2546	1509	420	272	4	135	137	205	~	~
5	280	248	295	337	293	375	438	384	519	407	5	52	53	63	181	53
6	279	220	270	142	205	345	409	383	196	540	6	120	122	163	~	~
7	221	230	307	145	284	416	473	419	453	405	7	102	114	142	242	114
8	211	193	225	143	218	305	392	416	284	379	8	101	96	108	180	96
9	187	240	225	101	241	392	362	343	306	527	9	92	97	159	294	97
10	238	247	277	243	187	319	454	461	319	364	10	56	53	57	171	53
											11	123	126	178	~	~
											12	94	97	122	144	97
											13	174	161	262	312	161
											14	111	112	192	184	112
											15	49	53	116	162	53

By analyzing the time taken by the subjects to carry out the required tasks, it is possible to identify the process models in which more time was needed. For instance, for the understandability tasks in the first family, the subjects took more time to analyse process models 5, 7 and 10, whilst they took more time to carry out the modifications requested with process models 3, 4 and 7. On the other hand, the time taken by the subjects in the second family of experiments to carry out the tasks relating to the model's understandability is greater for process models 1, 2, 3, 4 and 13. For the modifiability tasks, the models 1, 2 and 13 had more spend time.

The results in both families reflect, in the first instance, the relationship of the understandability times - degree of model complexity, when comparing tables 5 and 6, since process models 7 and 10 in the first family and process models 1 to 4 in the second family coincide as being those of greater complexity. The descriptive analysis relating to correct answers, subjective evaluation and efficiency was carried out in a similar manner. Once the descriptive analysis of the data had been completed, the statistical correlation analysis was carried out, which is presented next.

5.2 Correlation Analysis

The first step in the correlation analysis was to ascertain whether the distribution of the data was normal, so the Kolmogorov-Smirnov test was applied. As we obtained that the distribution was not normal, we decided to use a non-parametrical statistical test, namely 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 measures with the dependant variables as regards each of the aspects evaluated in the descriptive analysis (answer times, correct answers, subjective evaluation and efficiency). The following subsections show the results obtained in the two families of experiments.

5.2.1 Results of the First Family

Table 7 shows the results of the correlations of the CFC metrics with regard to the measures of the dependent variables. With regard to understandability, only the $CFC_{AND\text{-}split}$ metric was validated in the fourth experiment in correlation with the answer times and subjective evaluation. In this case, we can assume that the number of AND-split construct affects the understandability of the model, which is reflected in the answer time.

Table 7. Correlations of the CFC metrics and understandability - First Family

Measure	Understandability		Modifiability				
	Times	Sub. Eval.	Times	Subj. Eval.		Efficiency	
	Exp-4	Exp-4	Exp-2	Exp-2	Exp-3	Exp-2	Exp-3
CFC (XOR)			X	X	X	X	X
CFC (OR)							
CFC (AND)	X	X			X		
CFC			X	X	X	X	

With regard to modifiability, the correlation analysis shows that the $CFC_{XOR\text{-}split}$ and CFC metrics were validated in experiments 2 and 3 in relation to the answer times, subjective evaluation and efficiency. On the other hand, only the $CFC_{AND\text{-}split}$ was validated in the third experiment in relation to the subjective evaluation.

From the results of the correlations analysis obtained in the first family of experiments, we can observe that the relationship of CFC metrics to process complexity is greater with respect to the modifiability aspect in particular.

Specifically, these results show that the XOR-split construct affects above all the modifiability of the model. In addition, the validation of the CFC metric (which adds all the split constructors) gives us an indication that the structural complexity of a process, from the point of view of control flows, affects modifiability.

5.2.2 Results of the Second Family

In the second family of experiments also, the understandability and modifiability aspects were evaluated but in separate experiments, designed to analyse each aspect. By following the same procedure as that which was carried out in the first experimental family, once we had obtained the summary of data for each of the dependent variables measures (answer times, correct answers, subjective evaluation and efficiency) we carried out the analysis of correlations.

Table 8 shows that the CFC metrics were, on the whole, validated in relation to the variables analyzed. With regard to the understandability the correlations with the answer time the CFC metrics were validated in all experiments. This same correlation exists with respect to the variables of subjective evaluation and efficiency. With regard to the correct answers, these were only validated in the second experiment.

Table 8. Correlations of the CFC metrics and Understandability - Second Family

Measure	UNDERSTANDABILITY												MODIFIABILITY							
	Times			C. Answer			Sub. Eval.			Efficiency			Times		C. Answer		Sub. Eval.		Efficiency	
	E-1	E-2	E-3	E-1	E-2	E-3	E-1	E-2	E-3	E-1	E-2	E-3	E-4	E-5	E-4	E-5	E-4	E-5	E-4	E-5
CFC (XOR)	X	X	X		X		X	X	X	X	X	X	X	X			X	X	X	
CFC (OR)	X	X	X		X		X	X	X	X	X	X	X	X			X		X	
CFC (AND)	X	X	X		X		X	X	X	X	X	X	X				X		X	
CFC	X	X	X		X		X	X	X	X	X	X	X	X			X	X	X	

The correlation analysis results with regard to the modifiability also indicate that all the CFC metrics are highly correlated with the modifiability of the process models. The influence of the control-flow complexity on the modifiability of the BPMN models is reflected principally in the answer time in the required tasks, as well as in the subjective evaluation and efficiency in the accomplishment of the tasks.

There are significant differences between the results obtained from the correlation analysis in the experiments of the second family as compared to the first one. These differences were also observed when the validation of measures for BPMN models was carried out. A reason for this was that the experimental material used in the second family of experiments was an improved version of the first one (which had not much variability in the structural complexity), according to the feedback obtained. This way, the accomplishment of the second family was based on two main characteristics: a) the selection of a subset of structural complexity measures which included only the most significant measures (29 from the 60 initially defined) according to empirical results and an analysis of principal components and b) the increase in the variability of the structural complexity of the models. All this being so, we can consider the results obtained in the second family to be more conclusive.

The results obtained indicate that XOR-split, OR-split, and AND-split constructors affect the understandability and modifiability of the model. Therefore, based on the results and as regards the hypothesis proposed, it is possible to reject the null

hypotheses and conclude that there is a significant correlation between the CFC metric and the understandability and modifiability of BPMN models.

Finally, as result of this empirical study, we consider that the CFC metric is a suitable complement in measuring the structural complexity of business processes models with BPMN alongside the measures proposed in [3]. With the use and validation of the CFC metrics it is possible to obtain additional information with regard to the structural complexity of BPMs, in this case from a control-flow perspective. This allows designers to determine at the time of building process models (given more than one possible and equivalent modelling alternative) which of those models is more usable and maintainable.

6   Conclusions

In this work we have presented the evaluation and empirical validation of the CFC metric for measuring BPMN business process complexity from the point of view of their control-flows. The empirical validation was made possible by using the results obtained from two families of experiments which included the carrying out of a total of ten experiments. Initially, these experiments were carried out with the aim of evaluating the structural complexity of BPMs, as a means towards obtaining useful information concerning their understandability and modifiability.

The CFC is a design-time measurement. It can be used to evaluate the difficulty of producing a BPMN process design before implementation. When control-flow complexity analysis becomes part of the process development cycle, it has a considerable influence on the design phase, leading to further optimized processes. It is a well-known fact in software engineering that it is cost-effective to fix a defect earlier in the design lifecycle rather than later. To enable this to be done we introduce the first steps with which to carry out process complexity analysis.

As a result of applying the CFC metric, we were able to obtain additional information regarding the structural complexity of business processes. It was also possible to validate the CFC metric and to prove that it is highly correlated with the control-flow complexity of a business process and therefore with its understandability and modifiability. These results, along with the results on the validation of BPMN measures previously obtained, provide useful information when carrying out improvements or maintenance tasks in the process model. A better understanding of the process facilitates its later modelling and evolution.

We believe that evaluation and measurement of business process complexity in early phases of development (such as design and modeling phases) can help to identify problems in a process model and therefore assist us to design or choose process models that are easy to understand for all stakeholders. Understandable models also facilitate maintenance tasks, reducing implicit costs. On the other hand, models that are easy to understand and maintain can provide support on development of tasks, such as process reengineering, redesign of business process on a large-scale and refactoring, amongst others.

References

1. Succi, G., Predonzani, P., and Vernazza, T., Business Process Modeling with Objects, Costs and Human Resources, in Systems Modeling for Business Process Improvement, D. Bustard, P. Kawalek, and M. Norris, Editors. Artech House. pp. 47-60. (2000)
2. Cardoso, J., How to Measure the Control-flow Complexity of Web Processes and Workflows, Workflow Handbook, WfMC Ed. Lighthouse Point, FL. pp. 199-212. (2005)
3. Rolón, E., Ruiz, F., Garcia, F., and Piattini, M., Applying Software Metrics to evaluate Business Process Models. CLEI-Electronic Journal, Vol. 9(1, Paper 5). (2006)
4. Rolón, E., Garcia, F., Ruiz, F., and Piattini, M. An Exploratory Experiment to Validate Measures for Business Process Models. First IEEE International Conference on Research Challenges in Information Science (RCIS´07). Ouarzazate, Marruecos: IEEE. (2007)
5. Rolón, E., Garcia, F., Ruiz, F., Piattini, M., et al. Evaluation of BPMN Models Quality: a Family of Experiments. 3rd International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE´08). Funchal, Madeira. (2008)
6. OMG, Business Process Modeling Notation (BPMN) Specification, (2006)
7. Cardoso, J., Mendling, J., Neumann, G., and Reijers, H.A. A Discourse on Complexity of Process Models. BPM 2006, Workshop on Business Process Design. LNCS. (2006)
8. Ghani, A.A.A., Wei, K.T., Muketha, G.M., and Wen, W.P., Complexity Metrics for Measuring the Understandability and Maintainability of Business Process Models using Goal-Question-Metric (GQM). International Journal of Computer Science and Network Security (IJCSNS), 8(5): pp. 219-225. (2008)
9. Gruhn, V. and Laue, R. Complexity Metrics for Business Process Models. 9th Int. Conference on Business Information Systems (BIS´06). Klagenfurt, Austria. (2006)
10. Cardoso, J. Control-flow Complexity Measurement of Processes and Weyuker´s Properties. 6th International Enformatika Conference. Transactions on Enformatika, Systems Sciences and Engineering. Budapest, Hungary. (2005)
11. Gruhn, V. and Laue, R. Adopting the Cognitive Complexity Measure for Business Process Models. 5th IEEE International Conference on Cognitive Informatics (ICCI´06). Beijing, China. (2006)
12. Mendling, J. and Neumann, G. Error Metrics for Business Process Models. 19th International Conference on Advanced Information Systems Engineering (CAISE´07). Trondheim, Norway. (2007)
13. Mendling, J., Detection and Prediction of Errors in EPC Business Process Models, Vienna University of Economics and Business Administration: Vienna, Austria. (2007)
14. Mendling, J., Neumann, G., and van der Aalst, W.M.P. Understanding the Occurrence of Errors in Process Models based on Metrics. OTM Conference: LNCS, Springer. (2007)
15. Muehlen, M.z. and Recker, J. How Much Language is Enough? Theoretical and Practivcal Use of the Business Process Modeling Notation. 20th International Conference on Advanced Information Systems Engineering (CAISE'08). Montpellier, France. (2008)
16. Briand, L., Morasca, S., and Basili, V., Property-Based Software Engineering Measurement. IEEE Transactions on Software Engineering, 22(1): pp. 68-86. (1996)
17. Basili, V. and Rombach, H., The TAME Project: Towards Improvement-Oriented Software Environments. IEEE Transactions on Software Engineering, 14(6): pp. 728-738. (1988)
18. Cardoso, J. Process control-flow complexity metric: An empirical validation. IEEE International Conference on Service Computing (SCC´06). Chicago, USA. (2006)