Analyzing the Influence of Certain Factors on the Acceptance of a Model-based Measurement Procedure in Practice: An Empirical Study

Nelly Condori-Fernández, Oscar Pastor

Centro de Investigación en Métodos de Producción de Software
Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia.
{nelly, opastor}@pros.upv.es

Abstract. Full automatic software measurement from conceptual models is now accepted by academics, although take-up of these model-based measurement procedures in practice by software practitioners has been slow. To encourage acceptance in industry, an acceptance model for measurement procedures is proposed, identifying a set of factors that influence perceived usefulness and perceived ease of use when a user employs a measurement procedure. Analyzing the results of an empirical study carried out with software engineering academics, we find which factors have an influence on other factors. Using regression analysis, certain factors are identified that affect perceived usefulness and ease of use, and which in turn will affect intention to use.

1. Introduction

Although software measurement is recognized as a key element of engineering science, it has not yet been widely accepted in practice by software practitioners. The Software Engineering Measurement and Analysis (SEMA) group at the Software Engineering Institute (SEI) concluded from a series of explorative studies carried out from 2004-2005 [1] that there is still a significant gap between the current and desired state of software measurement. One of reasons for this is that there are no programs that use measures and empirical evidence to assess the practical relevance of such programs.

Nowadays, with the appearance of the model-driven development process, several approaches have arisen which allow for full-automatic software measurement of specific artifacts developed at early stages and in particular contexts [2][3][4][5][6]. However, the question is whether these model-based measurement procedures would be accepted in practice.

According to Cooper and Zmud [7], acceptance is one of the stages in the diffusion of technological innovations, and is defined from an employee perspective; an organization’s personnel are induced to commit to Information Technology application usage. Acceptance must not be confused with adoption; which is defined as a stage where negotiations are started in relation to the decision to adopt the innovation and mobilizing of organizational and financial resources for doing so [7].

The acceptance of technology has been investigated in a number of different fields [7][8][9]; however, in the software measurement field there are few papers on this subject in the literature.

Umarji and Emurian [10] focus on the evaluation of the likelihood of acceptance of a metrics program. Their model takes as input organizational culture, and the nature of the metrics program. Gopal et al. [11] researched the influence of institutional factors on the assimilation of metrics in software organizations. They also identified a set of determinants for metrics program success [12]. These determinants are divided into organizational and technical variables.

Our proposal focuses on a model-based measurement procedure relating to acceptance from a software practitioner's perspective. A number of models exist for evaluating the acceptance of new techniques and technology, in particular the Technology Acceptance Model (TAM) [14]. The Method Evaluation Model (MEM) [21], which uses the same TAM constructs, was the first to be applied in the context of Functional Size Measurement (FSM) procedures ([3], [17]). From preliminary results obtained with MEM, a theoretical model was defined, which includes a set of factors that affect practitioners' perceptions, perceptions that will determine the user's intention to use the model-based measurement procedures [13].

The aim of this paper is to analyze the influence of these factors on acceptance of RmFFP in practice, using the regression analysis technique. RmFFP is a measurement procedure designed to automatically estimate the functional size of object-oriented applications generated in an MDA environment

This paper is structured as follows: Section 2 introduces an acceptance model for model-based measurement procedures, Section 3 shows how an initial empirical study is carried out to analyze the causal relationships of the model, and finally, our conclusions are given and further work is suggested.

2. Evaluating the acceptance of measurement procedures

In order to define our model for evaluating acceptance of model-based measurement procedures; we use the same TAM constructs, but which have been redefined in the following way [13]:

- **Perceived Ease of Use:** the extent to which a person believes that using a particular measurement procedure would be free of effort.
- **Perceived Usefulness:** the extent to which a person believes that a particular measurement procedure will be effective in achieving intended objectives.
- **Intention to Use:** the extent to which a person intends to use a particular measurement procedure.

In addition, we identified the following factor types:

- **Intrinsic Factors** related to the intrinsic nature of software measurement procedure; these correspond to quality and tangibility of results, and the minimum number of actions required for calculating the measure using a measurement procedure.

- *Quality of results:* extent to which a person believes that the results of using a measurement procedure are accurate and convertible.
- *Tangibility of results:* extent to which a person believes that the results of using a measurement procedure are observable and understandable.
- *Minimum actions:* extent to which a person believes that using a particular measurement procedure would obtain results with the minimum number of actions required.

- **Extrinsic Factors** that do not depend on the measurement procedure in itself; these correspond to the experience and job relevance of the software practitioner.
  - *Job relevance:* extent to which an individual believes that a measurement procedure is applicable and relevant to his or her job.
  - *Experience:* knowledge or skill gained in the use of measurement procedures over a period of time.
- **External factors** that depend on the organization as a whole. These factors include where the business follows trends in the market based on advertising and marketing or peer company use, or the maturity level of an organization, or has business priorities giving rise to time or cost constraints.

The causal relationships hypothesized between the TAM constructs and factors of the model are shown in Figure 1. In the next section, we present an empirical study to analyze these causality relationships.

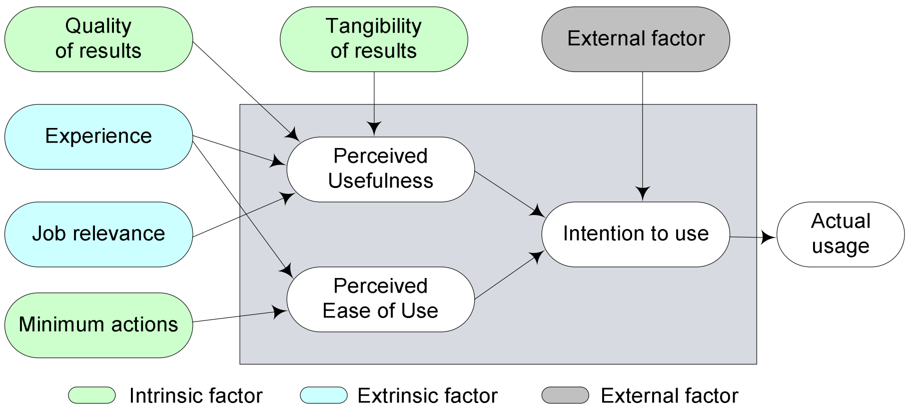

Figure 1. Acceptance model for model-based measurement procedures

3. Analyzing causality relationships in the Acceptance of RmFFP

RmFFP is a functional size measurement procedure designed on the basis of the COSMIC standard method, which has been approved by ISO/IEC 19761 [20]. RmFFP was proposed in order to automatically estimate the functional size of object-oriented systems generated in an MDA environment [5]. The object to be measured is the functional requirements specification obtained using the OO-Method requirements model [18].

This procedure starts with the definition of the *measurement strategy*, which includes the purpose, the scope, and the measurement viewpoint. The scope of RmFFP comprises the functionality to be included in a particular measurement. The measurement viewpoint corresponds to the 'analyst' viewpoint, which will focus on a requirements specification (object of interest).

Then, RmFFP starts a *mapping phase* to identify the significant primitives of the Requirements Model that contribute to the system's functional size according to the concepts of the COSMIC [20]. We defined sixteen mapping rules whose principal purpose is to reduce misinterpretation about the COSMIC generic concepts and to facilitate the automation of the RmFFP procedure. For instance, each use case is identified as a functional process; each message of the sequence diagram is identified as a data movement type, etc. The main outcome of this phase is the identification of data movements that are fundamental components of COSMIC.

Once the data movements have been correctly identified, we proceed with the *measurement phase*, whose purpose is to produce a quantitative value that represents the software functional size of a requirements specification. To do this, we apply the measurement function, which consists of assigning a numerical value of 1 Cfsu (Cosmic Functional Size Unit) to each data movement. We defined four rules to add together these quantified data movements. To do this, we used the relationship types between use cases to calculate the size of the functional processes (use case) and the size of the entire system

3.1 Planning: Case study

In order to define the goal of our empirical study, we used the Goal/Question/Metric (GQM) template [15], which is described as follows:

To analyze the Acceptance Model proposed ***for the purpose of*** evaluating RmFFP ***with respect to*** their acceptance in the practice ***from the viewpoint*** of the researcher ***in the context of*** software engineering professors using a measurement procedure for requirements specifications.

From this goal, the following research questions were addressed by this study:

RQ1: is perceived usefulness of the RmFFP measurement procedure really influenced by certain intrinsic factors?

RQ2: is perceived usefulness of the RmFFP measurement procedure really influenced by certain extrinsic factors?

RQ3: is perceived ease of use the RmFFP measurement procedure really influenced by certain intrinsic factors?

RQ4: is the intention to use really a result of the perceptions experienced by the subjects using the RmFFP measurement procedure?

Selection of subjects. The subjects were 20 professors from various Peruvian universities. They were enrolled in the United Nations summer school on "Advanced Techniques in Software Development", February - March 2007. The careful selection of participants was based on academic qualifications, teaching or industrial experience, technical background, and specific interest in software engineering. The empirical study was organized as a part of the "Measurement and Software Quality" course given during the summer school.

Variables and Hypotheses. Using the framework proposed by Juristo and Moreno [16], we identified three types of variables:

- **Response variables:** variables that correspond to the outcomes of the empirical study. For this study, we considered certain factors and constructs of the Model as response variables: Perceived Ease of Use (PEOU), Perceived Usefulness (PU), Intention to Use (IU), Job Relevance (JR), Quality of Results (QR), Tangibility of Results (TR), and Minimum Actions (MA). We omitted the extrinsic factor: "experience" and the external factors, which will be considered in further studies. As these outcomes should be measurable, we used a 5-point Likert scale format.
- **Factors:** variables that affect the response variable. In our study, this variable corresponds to the Models-based Measurement Procedures, and as single treatment: the RmFFP procedure [5]
- **Parameters:** variables that we do not want to influence the experimental results: level of practitioner's experience using a measurement procedure; complexity of conceptual models to be measured.

The following hypotheses regarding the research questions were considered:
H1: Perceived Usefulness is determined by the quality of results of the RmFFP measurement procedure.
H2: Perceived Usefulness is determined by the tangibility of results of the RmFFP measurement procedure.
H3: Perceived Usefulness is determined by job relevance using the RmFFP measurement procedure for the software practitioner.
H4: Perceived ease of use is determined by the minimum number of actions required using the RmFFP measurement procedure.
H5: Intention to use is determined by usefulness perceived.
H6: Intention to use is determined by perceived ease of use.

3.2 The Collection Data Method

First, we gave an introduction on how to apply the RmFFP measurement procedure by means of illustrative examples. Finally, we verified the knowledge learned by the participants by working through an assigned application. The time used for the training session was 4 hours distributed over two days. Then, each subject used the RmFFP measurement guide to measure a requirements specification of a Car Rental application with thirty-five use cases. The time allowed for this task was unlimited.

Finally, each subject was asked to complete a specially-designed survey to evaluate RmFFP acceptance. The time allowed for this task was also unlimited.

Instrumentation. A survey instrument[1] was designed to measure the response variables, with twenty closed questions. These questions consisted of 6 items used to measure PEOU; 2 items to measure PU; 3 items to measure IU; 4 items to measure JR; 2 items to measure QR; 1 item to measure TR; and 2 items to measure MA. Table 1 presents the four items used for the job relevance factor.

Table 1. Items formulated for measuring the job relevance factor

Construct	Description	Items
Job relevance	It is possible for a measurement procedure not to be perceived as useful even though the procedure provides accurate results, possibly because the use of the measurement procedure is not relevant for the job type of the software practitioner concerned.	1. Using the measurement procedure, the performance of my job will improve. 2. The use of the measurement procedure is relevant for my job. 3. Using the measurement procedure could increase the effectiveness of the development of my tasks. 4. I would use a measurement procedure, if I had to manage a software project

Responses to the instrument were based on a 5-point Likert scale ranging from (1), strongly disagree, to (5), strongly agree. The order of the items was randomized and some questions negated to avoid monotonous responses.

We also used a set of training materials, such as: a set of instructional slides on RmFFP procedure; an example of the application of RmFFP, and a measurement guide.

3.3 Data Analysis and Interpretation

As we can see in Figure 1, the intention to use a measurement procedure is influenced by perceptions of usefulness and ease of use; which can be influenced by certain type of factors. We identified several relationships, which were defined above in the six hypotheses (H1-H6). In this section, we analyze them by applying the regression analysis technique.

H1: Quality of results → Perceived usefulness. The regression equation resulting from the analysis is: *PU = 2.376 + 0.477*QR*.

The regression had a high significance level ($p < 0.01$), which means that H1 was confirmed. The determination coefficient ($R^2 = 0.316$) showed that 31.6% of the total variation in perceived usefulness can be explained by variation in quality of results.

H2: Tangibility of results → Perceived usefulness. The regression equation resulting from the analysis is: *PU = 3.208 + 0.236*TR*.

[1] http://www.dsic.upv.es/~nelly/survey2.pdf

The regression had a null significance level ($p > 0.1$), which means that H2 was not confirmed.

H3: Job Relevance → Perceived usefulness. The regression equation resulting from the analysis is: *PU = 2.86 + 0.348*JR*.

The regression had a medium significance level ($p < 0.05$), which means that H1 was confirmed. The determination coefficient ($R^2 = 0.186$) showed that 18.6% of the total variation in perceived usefulness can be explained by variation in job relevance.

H4: Minimum actions → Perceived ease of use. The regression equation resulting from the analysis is: *PEOU = 2.733 + 0.314*MA*.

This regression had a null significance level ($p > 0.1$), which means that H4 was not confirmed.

H5: Perceived usefulness → Intention to use. The regression equation resulting from the analysis is: *ITU = 1.628 + 0.577* PU*.

The regression had a medium significance level ($p < 0.05$), which means that H5 was confirmed. The determination coefficient ($R^2 = 0.166$) showed that 16.6% of the total variation in intention to use can be explained by variation in perceived usefulness.

H6: Perceived ease of use → Intention to use. The regression equation resulting from the analysis is: *ITU = 2.881 + 0.298* PEOU*.

The regression had a null significance level ($p > 0.1$), which means that H6 was not confirmed.

Table 2 below summarizes the regression analysis results in terms of the predictive power (R^2) and significance level of the model (p), and the confirmation of the casual relationships.

Table 2. Regression analysis results

Causal hypotheses	Predictive power	Significance. level*	Confirmed?
H1: QR → PU	31.6%	High	Yes
H2: TR → PU	--	Null	No
H3: JR→ PU	18.6%	Medium	Yes
H4: MA→ PEOU	--	Null	No
H5: PU → IU	16.6%	Medium	Yes
H6:PEOU→IU	--	Null	No

Note that three hypotheses out of six were confirmed using a regression analysis (H1, H3, and H5). This means, that the perceived usefulness is determined by the quality of results, and by the job relevance using RmFFP for the software practitioner. In addition, the intention to use RmFFP is determined by the perceived usefulness.

3.4 Validity evaluation

It is important to ensure that the obtained results are valid, we present the more important threats related to our empirical study in Table 3.

* Null: $\alpha > 0.1$, Low : $\alpha < 0.1$, Medium: $\alpha < 0.05$, High: $\alpha < 0.01$, Very high: $\alpha < 0.001$

Table 3. Type of threats to the validity of the results obtained in our empirical study

Type of threats	Description
Conclusion validity	• *Random heterogeneity of subjects:* All the subjects selected for the empirical study had approximately the same level of background. We are aware that this homogeneity reduces the external validity of our empirical study. • *Reliability of measures:* We are aware that the measures based on perceptions are less reliable than objective measures, since it does not involve human judgment. However, to diminish this threat, we carried out a reliability analysis on the survey used, which is explained below.
Construct validity	• *Inadequate pre-operational explanation of constructs:* To ascertain whether the constructs are sufficiently defined, and, hence the experiment is sufficiently clear, we conducted a reliability analysis on the survey, calculating reliability using the Chronbach alpha technique. The generic value obtained was 0.85 indicating that the items included in the survey are reliable. However, a design adjustment on the questions corresponding to the constructors PU, MA and QR would be required for further empirical studies, since their corresponding Cronbach alpha values were lower than 0.7 ([19]).
Internal validity	• *Instrumentation:* This is the effect caused by the artefacts used in the study execution. The requirements specification of the Car Rental System was reviewed; and the measurement guide was verified in advance with a small group of people in order to improve its understandability.
External validity	• *Interaction of selection and treatment:* This is the effect of not having a representative population in the experiment with which to generalize. In our case, we are aware that more studies with a larger number of subjects would be appropriate to reconfirm the initial results obtained.

4. Conclusions and further work

This paper provides a brief introduction to a theoretical model to evaluate the acceptance of measurement procedures from an individual perspective. The model includes three types of factors that influence perceptions of usefulness and ease of use (intrinsic, extrinsic and external factors). An empirical study has been carried out to verify causal relationships that include the intrinsic and extrinsic factors. The analysis shows that perceived usefulness is influenced by the job relevance of the people that use a measurement procedure. However, with respect to intrinsic factors, only the quality of results could affect the perception of usefulness. Perceived ease of use cannot be determined by the minimum actions factor. Furthermore, the results show that the intention to use a measurement procedure can be influenced more strongly by perceived usefulness than by perceived ease of use.

We plan to make further adjustments to the questions on the survey to improve the reliability of certain constructs, such as PU, MA, and QR. In addition, we are aware that further experimentation with industry practitioners will be appropriate in order to reconfirm these initial results. Finally, as further empirical studies, we also intend to consider the influence of software practitioners' experience on the acceptance of model-based measurement procedures.

Referencias

[1] Kasunic M., State of Software Measurement Practice Survey, Carnegie Mellon, Software Engineering Institute, 2006, www.sei.cmu.edu/sema/presentations/stateof-survey.pdf

[2] Abrahão S., Gomez J., Insfran E. Mendes E., A Model-Driven Measurement Procedure for Sizing Web Applications, Conference on Model-Driven Engineering Languages and Systems (MODELS 2007), Nashville, TN, USA, September 30-Octuber 5, 2007, LNCS Springer, 2007.

[3] Abrahao S., Poels G., Pastor O. A Functional Size Measurement Method for Object-Oriented Conceptual Schemas: Design and Evaluation Issues. Software & System Modelling, 5(1): 48-71, Springer Verlag, 2005.

[4] Azzouz S., Abran A., "A Proposed Measurement Role in the Rational Unified Process and its Implementation with ISO 19761: COSMIC-FFP" in Software Measurement European Forum, Rome, Italy, 2004.

[5] Condori-Fernández N., Abrahão S., and Pastor O., On the Estimation of Software Functional Size from Requirements Specifications, Journal of Computer Science and Technology (JCST), Springer, 22(3): 358-370, 2007.

[6] Marín B., Pastor O., Giachetti G.: Automating the Measurement of Functional Size of Conceptual Models in an MDA Environment, 9th International Conference Product-Focused Software Process Improvement, Italy, June 2008, pp. 215-229.

[7] R.B. Cooper and R.W Zmud, "Information Technology Implementation Research: A Technological Diffusion Approach", Management Science, 36(2):123-139, 1990

[8] W. G. Chismar, S. Wiley-Patton, Does the Extended Technology Acceptance Model Apply to Physicians?, 36th Annual Hawaii International Conference on System Sciences, IEEE Computer Society, Big Island, USA, January 2003, pp. 160-167.

[9] Chau P.Y. K., An empirical investigation on factors affecting the acceptance of CASE by systems developers, Journal on Information and Management, Elsevier, 30(6): 269-280, 1996.

[10] Umarji M.and Emurian H., Acceptance Issues in Metrics Program Implementation, Proceedings of the 11th IEEE International Software Metrics Symposium METRICS 05, IEEE Computer Society, 2005, Washington,USA, pp. 10-29.
[11] Gopal A., Krishnan M.S., Mukhopadhyay T., Impact of Institutional Forces on Software Metrics Programs, IEEE Trans. on Software Eng, 31(8):679-695, August 2005.
[12] Gopal A., Krishnan M.S., Mukhopadhyay T., and Goldenson, Measurement Programs in Software Development: Determinants of Success, IEEE Transaction on Software Eng., 28(9):863-875, 2002.
[13] Condori-Fernández N., Pastor O., Towards a Theoretical Model for Evaluating the Acceptance of Model-Driven Measurement Procedures, Proceedings of the 20th International Conference on Software Engineering & Knowledge Engineering, SEKE 2008, San Francisco, USA, July 1-3, 2008, pp. 22-25 .
[14] Davis F. D., "Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology", MIS Quarterly, vol. 3, no. 3, 1989.
[15] Basili V. R. and Rombach H. D., The TAME Project: Towards Improvement-Oriented Software Environments, IEEE Transactions on Software Engineering, 14(6):758-773, 1988.
[16] N. Juristo, Moreno A, Basics of Software Engineering Experimentation, Kluwer Academic Publishers, Boston, 2001.
[17] Condori-Fernández N., Pastor O., An Empirical Study on the Likelihood of Adoption in Practice of a Size Measurement Procedure for Requirements Specification, Sixth International Conference on Quality Software (QSIC 2006), October 2006, Beijing, China, pp. 133-140.
[18] Pastor, O., Molina, J. Model-Driven Arquitecture in Practice. Valencia, Springer Berlin Heidelberg, New York, 2007.
[19] Garson D., Scales and standard measures from statnotes, North Carolina State University, Copyright 1998, last updated March 2008. http://www2.chass.ncsu.edu/garson/pa765/standard.htm.
[20] ISO, ISO/IEC 19761 Software Engineering-COSMIC-FFP-A Functional Size Measurement Method, International Organization for Standardization_ISO, Geneva, 2003.
[21] Moody D. L., The method evaluation model: a theoretical model for validating information systems design methods, 11th European Conference on Information Systems, ECIS 2003, Naples, Italy 16-21 June 2003.