Quality Models in Practice: A Preliminary Analysis*

Stefan Wagner, Klaus Lochmann, Sebastian Winter
Technische Universität München
Garching, Germany
{wagnerst,lochmann,winterse}@in.tum.de

Andreas Goeb
SAP Research
Darmstadt, Germany
andreas.goeb@sap.com

Michael Klaes
Fraunhofer IESE
Kaiserslautern, Germany
michael.klaes@iese.fhg.de

Abstract

This paper presents the findings of a survey on quality models in practice conducted among four software companies in Germany. In the first phase of the study, 25 quality managers and users of software quality models were interviewed regarding the use of quality models, quality assurance techniques, and problems arising from the current situation in their companies. We present qualitative and quantitative findings as well as our plans for the second study phase including an international online questionnaire.

1. Introduction

The quality of software systems is a concept with many aspects and is therefore hard to define precisely. Quality models (QMs) are a common way to handle quality in a structured manner. Following the definition of Deissenboeck et al. [3] a QM is *a model with the objective to describe, assess and/or predict quality*. This means that a QM can be provided in different forms, e. g. international standards, checklists, or implicitly by certain tools. QMs have been subject to research for many years resulting in a manifold and heterogeneous landscape of QMs [6]. The spectrum covers, for example, general product quality standards, domain-specific models and test models.

1.1. Research problem

The fact that there are so many different models describing software quality as a whole or certain aspects of it leads to the question which of these models are actually used in practice and for which purposes. Their usage seems to differ widely in practice, while detailed information about the actual usage and resulting problems is missing. Furthermore, it is not clear how such QMs contribute to the quality assurance processes they are used in.

1.2. Contribution

In this paper, we present an analysis of the use of QMs in four German companies. It does not focus on special aspects of software product quality but covers several domains of software development (standard software, custom development, and embedded systems) and different company sizes. We identified the main classes of currently used QMs, the quality assurance techniques applied with these QMs, as well as problems related to the models.

1.3. Related work

In 1993 Davis et al. [2] published the results of a postal survey on the practice in software quality but most of the questions were related to methodical topics rather than the use of models for product quality. However, their results show that 41 % of the respondents claimed not to have any formal QA method at all, and from the remaining respondents the majority (55 %) used their own method for QA rather than an external one.

More recently, Jung et al. [5] published a survey of ISO/IEC 9126. They investigated whether the categorization of quality attributes is correct and reliable. This resulted in an improved categorization to provide guidance for revising the standard.

A survey on software testing practices in Australia was conducted in 2004 by Ng et al. [7] among 65 different organizations. Their results show which testing methodologies and techniques as well as tools are used within these organizations. The survey also includes information about metrics, standards, and trainings related to software testing. The survey identifies current practices and therefore aims in the same direction as our survey. However, the scope is smaller as we consider QMs in general.

*This work has partially been supported by the German Federal Ministry of Education and Research (BMBF) in the project QuaMoCo (01 IS 08023B/D).

2. Methodology

This section provides an overview of the study design, research questions, and the study execution.

2.1. Study definition

We concretize the study ideas discussed in the introduction using the Goal-Question-Metric goal template as proposed, e.g., by [8]:

Analyze	*product quality models*
for the purpose of	*characterization*
with respect to their	*usage (where and for what)* and *possible improvement potentials*
from the viewpoint of	*quality managers* and *quality model users*
in the context of	(1) *specific industry partners* (2) *software developing industry*

We refine the study goal by defining the following three research questions:

RQ1. What QMs (including standards and laws) are used in which contexts?

RQ2. How is quality assurance influenced by QMs?

RQ3. What are current problems with QMs and which potentials for improvement exist?

2.2. Study design

The study is planned as an explorative survey in which practitioners are asked for currently applied approaches, best practices, and perceived improvement potentials. In order to maximally benefit from the study, we decided to split the study into two phases. Phase 1 involves only quality managers and quality model users from four specific companies, phase 2 extends the scope by quality mangers and quality model users from software development organizations world wide.

Phase 1 consists of personal interviews in which open as well as closed questions are asked. In phase 2 an online questionnaire is used focusing on a selection of closed questions. This paper focuses on the finalized phase 1 of the study, which serves not only as a prestudy for phase 2 by providing input for the selection of the most interesting questions and relevant answer options, but also allows a more detailed elicitation of qualitative information as a result of the chosen research method.

2.3. Study population

In phase 1, we do not assume to obtain statistical relevant results with high external validity due to the limited number of interviews and narrow scope of the study. Therefore, we do not invest much effort in arranging a probability sampling over the population of interest; instead, we use a *convenience sampling* approach for selecting the interviewees from the considered population. In order to obtain a maximum level of coverage in our population with a limited number of interviews, the interviewees are selected within *clusters* based on the different organizations (that develop software in different domains) and two different roles: general manager (GM), and quality manager (QM).

Table 1: Distribution of interviews

Organization	**Number of interviews**		
	GM	QM	$\sum$
Capgemini sd&m AG	1	2	3
itestra GmbH	2	0	2
SAP AG	11	4	15
Siemens AG	3	2	5

2.4. Study implementation

In the implementation of phase 1, we prepared the material required for conducting the study and piloted it. Since we planned to conduct the interviews in a face-to-face mode and collected qualitative as well as quantitative information, we had developed a questionnaire with semi-structured (i.e., questions known in advance) and structured parts (i.e., question and response categories known in advance). The questions in the questionnaire were identified and refined in a series of workshops. The final version of the questionnaire contained 15 open questions and 12 closed questions. The questions were ordered along the research questions (RQ1–3) and asked primarily for *first-hand experience* of the interviewee as suggested by Fowler [4].

Since different persons conducted the interviews, an interview guide had been prepared in order to reduce the impact of particular interviewer's characteristics on the interview result. The guide tells the interviewer how to conduct the interview in general (two interviewers, one who primarily asks questions and one who takes notes that should be afterwards reviewed by the interviewee). In addition, guidelines were provided on how to introduce the interview session. The questionnaire was initially tested internally in order to detect its potential weaknesses and estimate the time required for the following interviews.

2.5. Study execution

In total, 25 interviews were conducted. The first five interviews were used for pilot testing with real interviewees and led to a second version of the questionnaire with some minor revisions. Because only minor revisions were made,

all 25 interviews were used for the further analysis. Tab. 1 shows the distribution of the interviews over the different organizations. The predominant roles were general managers, project managers and quality managers. Most of the interviews required between 40 and 60 minutes. All interviews were documented using interview reports. In order to analyze the qualitative information collected, a coding schema [1] was developed and evaluated with respect to inter-coder reliability. Three of the interviews were randomly selected and re-coded by three different researchers from three different organizations. The Fleiss' kappa value for the agreement of the coders is .54. This result is satisfactory considering that we had four different coders in total that coded independently. The evaluated coding schema was then used to code the answers for all open questions in all interview reports.

3. Results

The interviews gave a broad spectrum of insights. However, we focus on the most important ones with respect to our research questions.

3.1. Qualitative results

For the question on the usage of standards, laws, and QMs (RQ1), we found that a wide variety is in use. In 25 interviews, 31 different QMs were mentioned. The companies apply general standards such as the "relevant ISO standards" but also QMs proposed by academia. In some domains and depending on the customer demands, domain-specific standards are in use: "depends on the project, e.g., may be required by customer in contract as part of requirements". Internally also checklists are used to specify quality: "Checklists are built on the basis of ISO 9126 and adjusted to a particular project." Finally, company-specific standards and defect classifications are in use, especially in the larger companies. They are among others used in customer communication: "Matrix for customer approval".

The role of this variety of QMs in QA is analyzed for RQ2. Again, the usage of QMs differs widely in different organizations and even in different projects. However, the main uses of QMs are as a requirements source and in quality assessment. For example, in some organizations the quality requirements are defined in company-specific standards: "The requirements mainly come from the company-wide standards"[1]. Other companies use QMs to further refine and specify quality requirements. General as well as company-specific standards are in use for quality assessments of software products: "Quality models play a central role in quality analysis projects. They are used to be able to make a quality statement."[1] Other uses of QMs include aggregation of quality evaluations and measurements, communication with customers, benchmarking software and tools, and decision making support.

According to the individual usage of QMs, several problems are identified in practice (RQ3). Four interviewees reported problems with the operationalization of current QMs: "Operationalization (break down to technical attributes) is difficult" and "The quality model is not operationalized enough. There is a gap between framework and execution." Problems are also described with classifying defects: "A company wide defect classification was attempted but does not seem to be possible and the issues are too diverse." Common quality attributes ("-ilities") have associated problems as well. It was stated that it is hard to cover all attributes and that the measurability of many attributes is limited. This led one interviewee to the conclusion: "The -ilities are good for management talk only."

3.2. Quantitative results

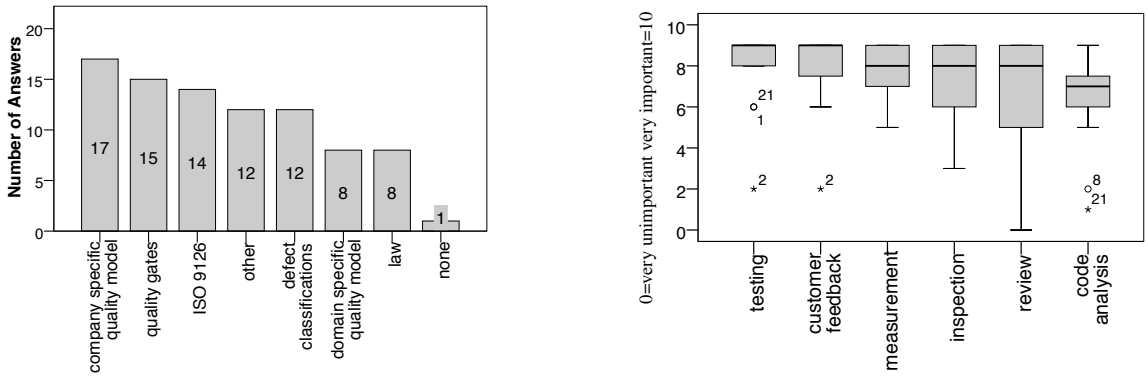
In order to address RQ1 it was asked which QMs are used in the interviewee's company. A set of QMs was provided, from which the interviewee could select in a multiple choice manner, and it was also allowed to name additional QMs. The results for the predefined set of qualtiy models is given in Fig. 1(a). There it can be seen, that company-specific models are the most used QMs. Since the qualitative interviews revealed that most company-specific models are created by customizing standards and laws, a cross-tab analysis was done. This revealed that of the 17 interviewees using company-specific QMs, 16 used them in conjunction with (on average more than two) other models. The possibility to name additional QMs was used frequently. The QMs that were named by more than one interviewee are listed (in brackets the counts are given): US-GAAP[2] (4), SOX[2] (3), MISRA (3), ISO 9001-2000 (2), and IEC 61508 (2).

A question on the importance of quality analysis techniques was asked to shed light on RQ2. The options were (1) reviews, (2) inspections, understood as reviews with a formally defined process, (3) code analysis, understood as tool-based static analysis of code, (4) testing, understood as dynamic software test, (5) measurement, understood as collection and analysis of various data, and (6) customer feedback. Fig. 1(b) shows the results as a bar-plot.

Regarding improvement potentials (RQ3), "defining practically usable quality models" was rated as the field of improvement with the greatest potential, "financial analyses of quality improvements" ranked second. On rank three and four "defining evaluation criteria" and "quantifying quality" follow.

[1] Translated from German.

[2] Laws regarding accounting rules used for financial statements.



(a) Quality models

(b) Assurance techniques

Figure 1: Quantitative results for the usage of quality models and quality assurance techniques

4. Threats to validity & discussion

4.1. Construct validity

We have not carried out observations in the analyzed companies but relied on interviews of people working there. Because of that, our analysis might differ from a direct observation.

4.2. External validity

In this study only four companies were included. Therefore, the results can only be generalized to some extent. However, the companies are from very different domains, from standardized software product development, through custom software development and consulting, to integrated hardware/software development. However, it must be noted that the interviews were not equally distributed over the participating companies as can be seen in Tab. 1.

4.3. Internal validity

The interviews were executed by interviewers that are active in the field of software quality and are involved in our research project on QMs. Hence, it may have occurred that the asked questions tended to direct the answers in a certain direction. Since most interviewees were project managers and persons responsible for quality management, the answers may be biased. To mitigate this threat, the interviews were mainly carried out by interviewers from the same companies. A second threat comes from the selection of the interviewees, that was not done randomly. Since the goal was to interview people that have to do with quality assurance, the results may be biased.

4.4. Discussion

Regarding RQ1 the study revealed that a large number of various QMs is in use. However, company-specific models seem to be popular and often complemented with standards and laws. The use of QMs for quality assurance (RQ2) differs widely in different companies and projects. They are used as a source for requirements and as a basis for quality analysis, e.g. checklists for quality gates. The mostly used quality assurance techniques are dynamic software tests and direct customer feedback. Employed to a lesser degree are measurements, reviews, and inspections. Static code analysis was only considered least important for quality assurance.

Regarding improvement potentials (RQ3), the results of the qualitative analysis are mainly confirmed by the quantitative one. “Defining practically usable quality models” was rated as the field of improvement with the greatest potential. Also on rank three and four are potentials that have to do with quantification and the definition of measurable evaluation criteria follow. In addition, on rank two financial analyses of quality improvements is seen to comprise large improvement potentials.

5. Conclusions

We conducted interviews in companies of various business domains to study the current use of QMs in practice. We were able to identify which QMs are actually used in practice. The purpose and concrete application of the QMs in quality assurance was analyzed. Besides the widespread use of QMs in practice we could still identify serious challenges that need to be addressed by future research. The operationalization of QMs, i.e. making them work in a realistic environment and producing quantified results, is a main problem.

We are currently working on a broader online survey that builds on the experiences and results of this preliminary study. We aim at a world-wide distribution of the questionnaire to have a broader foundation of the results.

References

[1] M. J. Corbin and L. A. Strauss. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publ., Los Angeles, 3. ed. edition, 2008.

[2] C. J. Davis, A. C. Gillies, P. Smith, and J. B. Thompson. Current practice in software quality and the impact of certification schemes. *Software Quality Journal*, 2(3):145–161, 1993.

[3] F. Deissenboeck, E. Juergens, K. Lochmann, and S. Wagner. Software quality models: Purposes, usage scenarios and requirements. In *WoSQ'09*. IEEE CS Press, 2009.

[4] J. F. Fowler. *Improving survey questions: Design and evaluation*. Sage Publ., Thousand Oaks, 1995.

[5] H.-W. Jung, S.-G. Kim, and C.-S. Chung. Measuring software product quality: A survey of ISO/IEC 9126. *IEEE Software*, 21(5):88–92, 2004.

[6] M. Klaes, J. Heidrich, J. Muench, and A. Trendowicz. CQML Scheme: A classification scheme for comprehensive quality model landscapes. In *Euromicro SEAA '09*, 2009. To appear.

[7] S. Ng, T. Murnane, K. Reed, D. Grant, and T. Chen. A preliminary survey on software testing practices in Australia. In *ASWEC'04*. IEEE CS Press, 2004.

[8] C. Wohlin. *Experimentation in software engineering: An introduction*. Kluwer Academic, Boston, Mass., 2000.