Rejoinder to the opposition reports on "Topic Title X"

Author X

1. Introduction

This rejoinder summarizes the corrections which were applied on the thesis report, based the comments given by the two opponents. Both opponent reports were thoroughly analyzed and if clarifications were needed, the opponent was contacted by email to get further explanations on his comment. In the following, the reports are identified as OR1 (by Reviewer A) and OR2 (by Reviewer B).

The rejoinder is presented in a tabular form whereby the columns are:

- ID: An identifier for the opponent comment, TA for Reviewer A and respectively BE for Reviewer B, which is used in some cases in the "Description" column to refer to other comments.
- Comment: The *literal* transcription of the comment given by the opponent.
- Action: The action taken by the authors of the thesis report:
 - o Fix the comment identified an issue and is corrected in the final thesis report
 - No fix the comment identified an issue and is *not* corrected in the final thesis report.
 The motivation is given in the column "Description".
 - Not valid the comment is proven to be wrong.
- Description: Elaborates the action taken by the thesis authors and gives motivations for not correcting an issue identified by the opponent.

2. Rejoinder for OR1 (by Reviewer A)

Table 1: Content issues

ID	Comment	Action	Description
TA1	Authors have very extensively discussed	Fix	A reading guideline is added at the beginning of
	the background of SPI. However, the		the thesis report (Section 1.4) instead of
	content in this chapter seems too		removing the suggested material. The guideline
	elaborative. On one hand it is good to		defines the optional, background material which
	begin with the basics as person new to		may be useful for readers not familiar with the
	the field can also read and understand		topic.
	the concepts presented in the study. On		
	the other hand, the huge detail can make		
	it difficult for the reader to stay focused		
	and eventually lose interest. It is		
	suggested to shorten content a bit in		
	chapter 2 for example measurement		
	frameworks/ evaluation methods.		

^{*}Note that references in the "Description" column refer to the final thesis report while the references prefixed with "R" are given in Section 5 of this document.

TA2	This chapter highlights the related work done in the past. However, this section seems rather short than the background section. It is suggested to enhance the motivation by putting some more concrete arguments e.g. relating missing things in previous work with your aims and objectives.	Fix	The OMSD model [82] is discussed to expand the scope of the related work (Chapter 3). However, the authors think that concrete arguments for conducting this research are already given. The following paragraph is added to make this explicit: "This thesis is aimed to build a framework for evaluating the outcome of SPI initiatives based on the model proposed in [28] and validate it. This implies also that it is aimed to address the issues regarding SPI evaluation (see Section 9.1, Table 47). The following paragraph shows how the above discussed research relates to the concepts proposed to solve these issues and motivates the implementation of SPI-MEF." The last paragraph points out aspects, which are not considered by the discussed papers, and which are aimed to be addressed by this thesis
TA3	By doing this the repetition of sentences i.e. "the study lacked appropriate time, methods of measurement, factors that threat to validity" can be avoided. These arguments can also be split and presented individually with reference to the specific previous studies, which did not address these aspects.	Fix	work. The redundant statements in Chapter 3 are removed.
TA4	In chapter 4 the adopted research methodology for this study is described in detail. Authors have succeeded in explaining fully the methods of conceptual analysis and expert opinion, the appropriate ways and steps to conduct them. However, the motivation for using "Conceptual analysis" regarding SPI-MEF is missing.	Fix	A motivation for using conceptual analysis is added in Section 4.1.1 and in the first paragraph of Section 4.2.
TA5	It is suggested to motivate the use of this method "Conceptual Analysis" and reduce general details of its steps as the relevant texts where these steps have been performed are not found.	Not valid	According to the research methodology workflow (Figure 9), conceptual analysis is applied two times. The description how the analysis steps were performed can be found on Page 27 (first and last paragraph).
TA6	The details describing what is "face-to- face interviews and telephonic	No fix	There were no details given except the <i>definition</i> of personal and telephone interview (two

	interviews" can also be skipped. It is sufficient to know the reason why it was required to perform interviews in two different ways.		sentences on Page 24).
TA7	It is suggested to shorten content evaluation result template, the dashboard. Every color with its three phases does not require explicit description, the pictorial representation seems sufficient to discuss the concept.	No fix	The authors deem the picture alone as <i>not</i> sufficient as it may be difficult to interpret correctly. Therefore, an explicit guideline is given how to interpret and read the dashboards.
TA8	Further, it is not understandable if the "side-effects" are referring to the "confounding variable" in section 5.2.2.1 page 38.	Not valid	"Side-effects" are not referring to "confounding variables". It should be clear from the context in Section 5.2.2.1 that side-effects occur after the treatment and are dependent variables, while from the definition of confounding factors (Section 5.2.3.3) it is clear that they are independent variables: "Confounding factors hide, distort or amplify the measured effects in the target entity and the improvement cannot be accurately assessed."
TA9	A few things remained unclear in this chapter e.g. what was this "investment unit" (IU)? How are authors able to assign some random numeric value to this IU?	Fix	A definition of "investment unit" was indeed missing and is added now in Section 5.3.3.2.
TA10	Moreover, how authors are able to use arithmetic operations on ordinal scale Impact Rating (IR)? The later question was also raised in later chapters where static validation is discussed, however no effective solution to this issue is found in the report. The concept of holistic view seems very promising however it is suggested to replace it by some other scale/means. The other way can be if author motivates more strongly the rationale of seeing it as an indication of improvement rather a metric.	Fix	The controversy if an ordinal scale can be treated as an interval scale boils down to the argument of being correct in a strict theoretical sense or being useful in practice. As Knapp [R2] points out, there are two schools of thought: the "conservative" group argues that inferential statistics, like mean and standard deviation, are not applicable on an ordinal scale [R2]. The "liberal" group, on the other hand, argues that, although no <i>true</i> interval scale is given, differences between categories can be regarded as equal [R2]. Furthermore, several studies (references given in [R2]) have empirically shown that it matters little if an ordinal scale is treated as an interval scale [R2]. Knapp [R2] cites S.S. Stevens, who defined the different measurement scales and whose work builds the foundation of the "conservative"

			camps' arguments. Stevens [R3] writes in the definition of the ordinal scale: "In the strictest propriety the ordinary statistics involving means and standard deviations ought not to be used with these scales, for these statistics imply a knowledge of something more than the relative rank-order of data. On the other hand, for this 'illegal' statisticizing there can be invoked a kind of pragmatic sanction: In numerous instances it leads to fruitful results. While the outlawing of this procedure would probably serve no good purpose, it is proper to point out that means and standard deviations computed on an ordinal scale are in error to the extent that the successive intervals on the scale are unequal in size. When only the rank-order of data is known, we should proceed cautiously with our statistics, and especially with the conclusions we draw from them." Based on this argument, the authors of the thesis report are confident that the proposed method to create the holistic view is at least practicable and useful. The method description given in Section 5.3.3 suggests that more than bare rank-order of data is given in the proposed impact rating. The following statement is added in Section 5.3.3.1: "On the other hand, it is pointed out by Stevens [107] that it can be practicable to treat an ordinal scale as interval scale. Therefore, the authors deem the application of statistical methods on the SVI, as the calculation of the mean, as permissible."
TA11	A long list of confounding factors is listed but how to control them; other than putting "to-be compared" entities in one group; is not suggested. If the underlying purpose was just to make readers aware of what factors can be confounding then it is suggested to present and emphasize the later aim more precisely.	Not valid	Emphasis is already given in Section 5.2.2.3 just before the list of confounding factors: "SPI-MEF aims to create awareness of the potential threats caused by confounding factors and, since they are dependent on the context in which the improvement initiative is carried out, exemplify those which are most commonly encountered."
TA12	However, some improvement can be made in chapter 6 e.g. repetition in table 16 can be avoided by referring to appropriate section. In table 16, row	Fix	The descriptions of improvement practices in Table 16 are removed and a reference to Section 6.1 is added.

	"Description", the "improvement practices" are repeated from section 6.1 as it is occupying a lot of space.		
TA13	One typical style has been observed while referring any reference at almost all places except chapter 3, e.g. "are presented in [28]" or "given in [xyz]". This makes hard for the reader to go back and forth in references and read what is meant at a certain point. Instead of this it is suggested to rephrase in a more readable way e.g. "are presented in [28]" can be described as "are described by Cheng et al [28]". In chapter 3, the referencing seemed fine but then it gave a clear hint of a different writing style.	No fix	The authors cannot identify a different writing style in Chapter 3. It is not true that references are always given as " presented/given in [xyz]". The proposed style change would lead basically to a complete rewrite of major parts of the report. Since the opponent did not motivate the suggestion, except by personal preference, it is not applied.

Table 2: General comments

ID	Comment	Action	Description
TA14	Despite its effectiveness and promising usefulness, few aspects remained unclear i.e. to what extent the framework would be tailor-able to the organizations who adopt it? It means which artifacts, levels, steps of SPI-MEF can be customized for an organization with different parameters like size, maturity, resources etc. Moreover, how much cost/ effort are required to effectively introduce the framework in any organization. Does it require specialized staff or group to learn how to use and train others involved in SPI? It was observed that implementation detail of the developed framework is out of scope of this study. However, the cost to learn the concepts of SPI-MEF in order to adopt it could have been discussed.	Fix	See comment BE2.
TA15	It was clear to the reader that heuristic about the lag and degradation factor was eliminated due to obvious, well stated reasons in the report. However, it is not	No fix	Specifically the concepts of Lag and Degradation Factor are already mentioned in future work (Section 9.2) for further investigation since the static validation did not reveal a conclusive result

disc	cussed very clearly how one can	on their	applicability	in	practice	(see	Section
iden	ntify the lag/degradation factor and	7.5.5).					
eval	luate a particular SPI initiative in						
case	e of parallel ongoing SPI initiatives						
(whi	ich may also be affecting each						
othe	er's performance).						
	-						

Table 3: Language structure and style

ID	Comment	Action	Description
TA16	Use of past tense and active voice is generally recommended in thesis report writing. Whereas the tense mostly found is in passive voice e.g., "in this thesis xyz is done" It shows author is intending to use present/past perfect tense, which however is not explicitly used. It is understandable to use present tense in describing fundamental concepts like describing difference between semi-structured/ structured interviews. But it is suggested to use active voice and past/present perfect tense consistently for the rest of the report.	No fix	The opponent did not motivate his claim that the active voice in generally recommended. On the other hand, the authors recognize that the usage of the active voice has its advantages (the text is more personal, vivid and possibly easier to read), and they discussed which form to use before they commenced the write-up of the report. They decided to use the passive form since it was anticipated that a major part of the report would comprise procedure descriptions (Chapter 5 and 6) and to present data (most of Chapter 7). For this purpose, the passive voice is favorable over the active voice [R1].
TA17	Use of long sentences has also been noticed. It might be inevitable at some places for the author due to contextual constraints. However, some sentences can be split to make them more readable e.g. "Using a single metric to measure a certain attribute can affect the validity of the forthcoming evaluation and to countermeasure this issue, the measurement concept incorporates cross-examination which can be used to increase data validity in two ways". Also, "In the next phase, the evaluation team is advised to act more actively by collecting data in the target entity, paying attention not to interrupt the ongoing development process while doing it."	Fix	The authors identified several sentences which needed improvement according to the opponents comment. The list of changes is documented in Appendix A.
TA18	Some attention is required to structure the sentences e.g. "For a more in-depth	Fix	The authors assume that the opponent refers here to a wrong ordering of adverb and verb. The

	discussion on baseline establishment, refer to Section 8.4.2.1 in [28] which provides also several references to publications which address baselines in general and present also advanced baseline construction techniques". It is a long sentence and for the bold phrases, it should be "also provides" and "also presents". Also, "The gap shows then which refinements in the measurement program need to be implemented". It should be "then shows".		authors agree that the ordering in the cited sentences is wrong and corrected the passages in Section 5.2.3.1 and 5.1.2. However, the authors regard this issue as minor and applied no further corrections in the thesis report.
TA19	Also, "Typical confounding factors that shall be taken into consideration for evaluation planning or during the evaluation", it should be "should".	Fix	"Shall" is replaced where appropriate by "should" on pages 44, 47, 48, 50, 53, 64, 79, 80, 81, 92 and 93.
TA20	Some spelling mistakes are also observed e.g. "well-know", "planed" on page 6, 'fulfil' on page 17.	Fix	The corrections proposed by the opponent are applied on pages 4, 6 and 57.

Table 4: Miscellaneous

ID	Comment	Action	Description
TA21	Chapter 5, page 37, evaluation scoping example running scenario, the 'Organization level' is discussed but missing in the example.	Fix	It was not intended to show the "Organizational level" in the example and it was listed erroneously in the running scenario on Page 37.
TA22	In figure 17, the 'Evaluation' text is not in the green box. It is not understandable if it is left without box on purpose.	Fix	"Evaluation" is not within a box since it is a constant as opposed to the other items in the diagram whose actual instances are variable. This is now clarified in Section 5.2.2.2.
TA23	On page 17, "which in his formal definition has to fulfil the representation condition of measurement [48]". It is not clear who is intended.	Fix	Sentence (it's actually on Page 57) is changed in "[] which in its formal definition has to fulfill the representation condition of measurement [48]."

3. Rejoinder for OR2 (by Reviewer B)

Table 5: Content issues

ID	Comment	Action	Description
BE1	The authors explained that they have been working closely with the authors of the SPI evaluation model referred in their study. Although the authors have considered the primary studies in the base study, their report did not describe the selection process they followed in the selection of the base study. Including their selection activities above mentioned activities they have conducted would be more transparent if the initial selection process and motivation behind their choice is described in the beginning of their report.	No fix	As already explained during the presentation defense, no selection procedure was taken at the commencement of the thesis project since the work is regarded as a continuation of the research performed by Cheng & Permadi [28]. This is stated clearly in the introduction (Chapter 1): "This thesis presents the continuation work of "Towards an Evaluation Framework for Software Process Improvement" [28] that proposed a model for the evaluation of SPI outcome. Motivated by the findings in [28], a measurement and evaluation framework called Software Process Improvement Measurement and Evaluation Framework (SPI-MEF) is constructed based on the proposed model."
BE2	In the proposed SPI-MEF framework, the last phase is the evolution stage which is characterized by the analysis of feedback to improve the evaluation process. The authors explained that the evaluation process would be able to reuse some of the artifacts but they were not able to state how much of it can be reused. The possibility to reuse the artifacts in subsequent evaluations is a useful practice especially considering the heavy nature of the whole process. However, the study does not highlight or gives an estimation of how much of the effort produces reusable artifacts. Estimation of how much of the effort is reusable will hint the implementer to primarily give focus on the reusable artifacts before moving into other artifacts.	Fix	As already explained during the presentation defense, SPI-MEF does currently not provide any support in estimating the effort which is needed to implement it. Since the framework was only statically validated, it is very difficult to come up with cost or effort estimations. Some interviewees argued that applying the framework could adhere with considerable additional effort (Section 7.4.4) which was discussed in Section 7.5.4. However, no concrete effort estimations are given and as a correction, this topic is added as potential future work (Section 9.2) to improve SPI-MEF.
BE3	The authors explained that they are not able to determine scalability since the framework has not been implemented on real projects. Given the fact that SPI initiatives are heavy processes, the introduction of this evaluation process,	Not valid	Nowhere in the report is written that the authors could not determine the scalability of the framework. The authors don't assume that "SPI initiatives are heavy processes" and the opponent does not motivate his claim that this is a fact. Even more, the report clearly states what the authors

	which is also a heavy process, makes the whole process very heavy. Most SPIs are primarily suited for large or medium sized organizations. Some small software firms regarded it as bureaucratic process since it takes a lot of their limited resources. In order for software firms to benefit from this process, estimation of how the framework scales would better serve them in allocating the necessary resource in the initial phases of their evaluation process.		regard as SPI initiatives, namely frameworks, practices and tools. Claiming that all of those involve heavy processes is an oversimplification. The authors agree that scalability is an important aspect but it depends on the concrete SPI initiative which is undertaken and needs to be evaluated. Since it was not the aim of the thesis work to validate the framework against a specific SPI initiative, the authors could not assess its scalability and giving a statement about scalability in general without having actual empirical data was deemed by the authors as inappropriate.
BE4	The authors explained that the use of the framework would require some level of training. The problem seen in using most SPI initiatives is that the processes require specialized personnel. The authors have not indicated what type of skill the process would require but having such information would be essential for proper planning of the evaluation process.	No fix	The opponent does not motivate his claim that SPI initiatives, in general, require specialized personnel. Obviously, this depends on the specific initiative that is undertaken. As SPI-MEF is not targeted to a specific initiative, it is not the aim of the framework to define exactly which skills are needed. On the other hand, certain tasks are generic and the framework suggests the involvement of certain roles in these processes (e.g. gap analysis in Section 5.1.2).

Table 6: Structure

ID	Comment	Action	Description
BE5	The language flow of the report is clear, although there are some repetitions. For instance, the text in Figure 2 was already stated in the body of text in previous section. There were also some over use of quotes in explaining the three levels above Figure 2.	Fix	The text in Section 2.3.1, where the conceptual, operational and quantitative levels of GQM are explained, is rephrased.
BE6	The references were consistently formatted and followed the correct format except for the web links. The web link references need to have a timestamp of last access.	Fix	The timestamps of last access are added for references [88], [89] and [113].

4. Conclusion

The authors of the thesis report would like to thank both the opponents for their time and their valuable input. Out of the 29 identified comments, 18 were addressed, 7 were not fixed, and 4 were deemed as invalid. The corrections were implemented in the final thesis report.

5. References

- [R1] K. Riley, "Passive Voice and Rhetorical Role in Scientific Writing," Journal of Technical Writing and Communication, vol. 21, 1991, pp. 239-257.
- [R2] T.R. Knapp, "Treating ordinal scales as interval scales: an attempt to resolve the controversy," Nursing Research, vol. 39, Apr. 1990, pp. 121-123.
- [R3] S.S. Stevens, "On the Theory of Scales of Measurement," Science, vol. 103, 1946, pp. 677-680.

6. Appendix A

TA17: Correction of long sentences

Page	Original	Revised
5	The first wave was the introduction of the waterfall life-cycle model [37] in 1970 that depicts the activities in software development, whereas the second wave was the conception of process improvement models like the Capability Maturity Model (CMM) in the late 1980's that drive the process maturity movement [32].	The first wave was the introduction of the waterfall life-cycle model [37] in 1970 that depicts the activities in software development [32]. The second wave is regarded as the conception of process improvement models like the Capability Maturity Model (CMM) in the late 1980's that drive the process maturity movement [32].
18	Using a single metric to measure a certain attribute can affect the validity of the forthcoming evaluation and to countermeasure this issue, the measurement concept incorporates cross-examination which can be used to increase data validity in two ways.	Using a single metric to measure a certain attribute can affect the validity of the forthcoming evaluation. In order to countermeasure this issue, the measurement concept incorporates cross-examination which can be used to increase data validity in two ways.
57	Since the aim of the holistic view is to provide an overview of the improvement (the individual evaluations are a more appropriate data source for decision making), SVI has to be seen as an index or score, giving an indication of the improvement, rather than a metric, which in his formal definition has to fulfil the representation condition of measurement [48].	Since the aim of the holistic view is to provide an overview of the improvement (the individual evaluations are a more appropriate data source for decision making), SVI has to be seen as an index or score. It gives an indication of improvement and is not a metric, which in its formal definition has to fulfill the representation condition of measurement [48].

72	The reason for having a wider coverage is that the effects of the improvement have propagated to the Product and Organizational level and the evaluation team assumes at that moment, that the top-level management is willing to provide more resources for the evaluation after seeing the results of the previous two evaluations.	The reason for having a wider coverage is that the effects of the improvement have propagated to the Product and Organizational level. Furthermore, the evaluation team assumes at this moment, that the top-level management is willing to provide more resources for the evaluation after seeing the results of the previous two evaluations.
84	The stakeholders (SPI Coordinator, Development Team and Head of Department) are all pleased with the results and they feel confident that the SPI initiatives will be able to improve their current processes when they are going to be implemented in Phase II to all projects in the Internet Banking System development.	The stakeholders (SPI Coordinator, Development Team and Head of Department) are all pleased with the results. They feel confident that the SPI initiatives will be able to improve their current processes when they are going to be implemented in Phase II to all projects in the Internet Banking System development.
105	There were two restrictions, which had to be taken into consideration: after a first inquiry of the expert candidates, it was clear that the allocated time for a participation was limited (approximately one hour, to show the order of magnitude); second, mostly for the industry experts, a face-to-face meeting was impossible due to their geographical location.	There were mainly two restrictions which had to be taken into consideration. First, after an initial inquiry of the expert candidates, it was clear that the allocated time for a participation was limited (approximately one hour, to show the order of magnitude). Second, mostly for the industry experts, a face-to-face meeting was impossible due to their geographical location.
106	Topic 6 is handled exclusively in the academia interviews since it was regarded as too difficult to judge for the industry interviewees and covers solely theoretical aspects; furthermore, since the topics were ranked beforehand by their priority and Topic 6 was judged as of moderate priority, it was decided to drop it from the industry interview to allow more time for the remaining questions.	Topic 6 is handled exclusively in the academia interviews since it was regarded as too difficult to judge for the industry interviewees and covers solely theoretical aspects. Furthermore, since the topics were ranked beforehand by their priority and Topic 6 was judged as of moderate priority, it was decided to drop it from the industry interview to allow more time for the remaining questions.
117	All interviewees agreed that the approach to evaluate the improvement can be thought as independent from the concrete initiative, that is, it is possible to define an evaluation framework which can be applied in different improvement contexts and SPI-MEF was assessed to be flexible and general enough to fulfill this purpose.	All interviewees agreed that the approach to evaluate the improvement can be thought as independent from the concrete initiative. Hence, it is possible to define an evaluation framework which can be applied in different improvement contexts and SPI-MEF was assessed to be flexible and general enough to fulfill this purpose.
121	One of the classic dilemmas in an organization is to balance the effectiveness of a process and its efficiency; therefore, should there be two initiatives running in parallel which are not in synergy with each other, the improvements will be hard to observe or realize without constant monitoring [INT_M].	One of the classic dilemmas in an organization is to balance the effectiveness of a process and its efficiency. Therefore, should there be two initiatives running in parallel which are not in synergy with each other, the improvements will be hard to observe or realize without constant monitoring [INT_M].

122	In general, it was agreed that the holistic view is beneficial [INT_L, INT_M, INT_N, INT_O and INT_P] as it allows an organization to quickly view the benefits of a given SPI initiative into various measurement levels and hence, enables an organization to relate the improvement to various angles rather than looking into just one good or bad result of a specific measurable entity [INT_M and INT_O].	In general, it was agreed that the holistic view is beneficial [INT_L, INT_M, INT_N, INT_O and INT_P] as it allows an organization to quickly view the benefits of a given SPI initiative into various measurement levels. Hence it enables an organization to relate the improvement to various angles rather than looking into just one good or bad result of a specific measurable entity [INT_M and INT_O].
123	However, the company culture is the key factor that would determine the acceptance of it at the initial stage and then of course top management has to be convinced by the driving force of the change, otherwise an implementation of the framework is derailed before it is even started [INT_M].	However, the company culture is the key factor that would determine the acceptance of it at the initial stage [INT_M]. Then of course top management has to be convinced by the driving force of the change, otherwise an implementation of the framework is derailed before it is even started [INT_M].
125	Therefore, to ease the application of the iron triangle metaphor that was introduced along with the concept of "primary and complementary measures", a pool of metrics for each measurement level grouped according to different success indicators is provided in the final framework so that practitioners can select the needed success indicators and corresponding metrics (see Appendix B).	Therefore, to ease the application of the iron triangle metaphor that was introduced along with the concept of "primary and complementary measures", a pool of metrics is provided in the final framework. Measures are grouped according to measurement level so that practitioners can select the needed success indicators and the corresponding metrics (see Appendix B).
127	Subjective ratings in improvement assessment are used in industry and therefore applicable in the "Holistic View"; the contribution of the framework would therefore be the formalization of that process, and it can be enhanced by describing how the rating can be homogenized between the different stakeholders which are involved in the evaluation.	Subjective ratings in improvement assessment are used in industry and therefore applicable in the "Holistic View". The contribution of the framework would therefore be the formalization of that process, and it can be enhanced by describing how the rating can be homogenized between the different stakeholders which are involved in the evaluation.
128	Therefore, in the final framework, in the "Holistic evaluation" concept, the "Gain/Loss" component was discarded (refer to SPI-MEF summary in Appendix A) and in addition to the combined holistic view of all measurement levels, another holistic view presentation was added for all the viewpoints in each measurement level (see Section 5.3.3.3) so that the evaluation done by each individual viewpoint can also be presented.	Therefore, in the final framework, in the "Holistic evaluation" concept, the "Gain/Loss" component was discarded (refer to SPI-MEF summary in Appendix A). In addition to the combined holistic view of all measurement levels, another holistic view presentation was added for all the viewpoints in each measurement level (see Section 5.3.3.3) so that the evaluation done by each individual viewpoint can also be presented.