



## D5.3 Evaluation of pilot cases (V1)

V0.1

<b>Programme</b>	H2020		
<b>Funding scheme</b>	RIA - Research and Innovation action		
<b>Topics</b>	ICT-10-2016 - Software Technologies		
<b>Project number</b>	732253		
<b>Project name</b>	Quality-aware rapid software development		
<b>Project duration</b>	1 <sup>st</sup> November 2016 – 31 <sup>st</sup> October 2019		
<b>Project website</b>	<a href="http://www.q-rapids.eu">www.q-rapids.eu</a>		
<b>Project WP</b>	WP5 – Evaluation		
<b>Project Task</b>	<i>Task 5.3 – Execution of the BITTIUM use case</i> <i>Task 5.4 – Execution of the SOFTEAM use case</i> <i>Task 5.5 – Execution of the ITTI use case</i> <i>Task 5.6 – Execution of the NOKIA use case</i>		
<b>Deliverable type</b>	X	R	Document, report
		DEM	Demonstrator, pilot, prototype
		DEC	Websites, patent fillings, videos, etc.
		OTHER	Material that does not belong to any specified category
		ETHICS	Ethics requirement
<b>Contractual delivery</b>	19/04/2018		
<b>Delivered</b>	19/04/2018		
<b>Responsible beneficiary</b>	Organisation		
<b>Dissemination level</b>	X	PU	Public
		CO	Confidential, only for members of the consortium (including the Commission Services)
		EU-RES	Classified Information: RESTREINT UE (Commission Decision 2005/444/EC)
		EU-CON	Classified Information: CONFIDENTIEL UE (Commission Decision 2005/444/EC)
		EU-SEC	Classified Information: SECRET UE (Commission Decision 2005/444/EC)



Authors	
Organisation	Fraunhofer IESE
Short name	Liliana Guzmán, Anna Maria Vollmer, and Silverio Martínez
Organisation	UoO
Short name	Pertti Karhapaa and Woubshet Behutiye
Organisation	UPC
Short name	Lidia López
Reviewers	
Organisation	Softeam
Short name	Antonin Abherve
Organisation	UoO
Short name	Woubshet Behutiye

Version history			
Version no.	Date	Description	Author
V0.1	31/01/2018	Outline	Liliana Guzmán (Fraunhofer IESE)
V0.2	28/02/2018	Object of study, evaluation goals, and evaluation methodology	Liliana Guzmán (Fraunhofer IESE), Pertti Karhapaa (UoO), and Lidia López (UPC)
V0.3	05/03/2018	Within-case analysis for each company	Liliana Guzmán and Anna Maria Vollmer (Fraunhofer IESE)
V0.4	16/03/2018	Review and update of within-case analysis for each company	Liliana Guzmán, Anna Maria Vollmer, and Silverio Martínez (Fraunhofer IESE), Pertti Karhapaa and Woubshet Behutiye (UoO), Lidia López (UPC)  Jari Partanen (Bittium), Michal Choras (iTTi), Sanja Aaramaa (Nokia), and Antonin Abherve (Softeam)
V1.0	16/03/2018	Executive summary, introduction, threats to validity, and conclusions	Liliana Guzmán (Fraunhofer IESE)
V1.0	16/04/2018	Move within-case analysis from main document to an private annex to keep confidentiality of company-specific results	Liliana Guzmán (Fraunhofer IESE)
V1.0	17/04/2018	Add cross-case analysis for the three companies which consent the report of the evaluation results	Liliana Guzmán and Anna Maria Vollmer (Fraunhofer IESE)
V1.0	17/04/2018	Deliverable review.	Woubshet Behutiye (UoO)
V1.0	19/04/2018	Deliverable review/Update deliverable	Antonin Abherve/Liliana Guzmán



Definition of the key terms and abbreviations	
PoC	Proof-of-Concept



## Contents

The list of tables .....	5
Executive summary .....	6
1. Introduction.....	7
1.1 Motivation .....	7
1.2 Intended audience.....	7
1.3 Scope .....	7
1.4 Relation to other deliverables .....	7
1.5 Structure of the deliverable .....	7
2. Object of study: Proof-of-Concept .....	7
3. Evaluation Methodology .....	8
3.1 Overview.....	8
3.1.1 Evaluation Goals and Questions .....	8
3.1.2 Evaluation Strategy.....	9
3.2 Evaluation of the Q-Rapids tool.....	9
3.3 Evaluation of the Q-Rapids process model.....	11
3.4 Data Preparation and Analysis Methods.....	12
4. Results: Cross Case Analysis .....	12
4.1 Appropriateness of the Q-Rapids Tool .....	12
4.1.1 Sample .....	12
4.1.2 Appropriateness of the Q-Rapids Quality Model .....	12
4.1.3 Appropriateness of the Q-Rapids tool's functionalities and features .....	13
4.1.4 Needs for Improvements.....	14
4.1.5 Threats to Validity .....	14
4.1 Appropriateness of the Q-Rapids Process Model.....	15
4.1.1 Sample .....	15
4.1.2 Appropriateness of the Process Model for "Setting-Up Q-Rapids" .....	15
4.1.3 Needs for Improvements.....	15
4.1.4 Threats to Validity .....	16
5. Conclusions.....	16
References.....	17



---

## The list of tables

Table 1: Instruments and Materials .....	10
Table 2: Instruments and Materials .....	11
Table 3: Appropriateness of the Q-Rapids quality model .....	13
Table 4: Appropriateness of the Q-Rapids functionalities and features .....	13
Table 5: Suggestions of improvements for the Q-Rapids tool .....	14
Table 6: Appropriateness of the Q-Rapids process model regarding the step “setting-up Q-Rapids” .....	15
Table 7: Suggestions of improvements for the Q-Rapids process model .....	16



## Executive summary

The Q-Rapids framework aims to provide the software industry with methods and tools for handling quality in the context of rapid software development (RSD). Q-Rapids relies on a generic data-driven, quality-aware, rapid development process characterised by integrated management of quality requirements and functional requirements. Q-Rapids aims to be generic and suitable for managing the variety of quality requirements in different application domains and project settings.

To assess the impact of Q-Rapids framework, we designed and performed a formative evaluation of the proof-of-concept of the Q-Rapids framework (hereafter referred as to Q-Rapids PoC). The Q-Rapids PoC provided product owners with tool support for analysing strengths and drawbacks of software products as well as for analysis of potential risks regarding the continuous software development (i.e., blocking situations). Moreover, the Q-Rapids PoC provided stakeholders with a process model for setting-up Q-Rapids, namely for defining the Q-Rapids quality model, determining data sources and providing access to data, and installing the Q-Rapids tool.

We aimed at evaluating the appropriateness of the Q-Rapids tool and of the Q-Rapids process model. We performed semi-structured interviews in which product owners used the Q-Rapids tool to analyse the quality of a software product and blocking situations based on real project data and give their feedback on the strengths and drawbacks of the Q-Rapids tool. We also performed semi-structured interviews with the stakeholders responsible for setting-up Q-Rapids and asked them to provide feedback on the Q-rapids process model.

Regarding the Q-Rapids tool, the evaluation result shows it encompasses a useful, understandable, and relevant quality model. Moreover, product owners perceive the Q-Rapids tool as reliable and efficient and the related dashboard as useful, easy to use, and easy to navigate. Regarding the Q-Rapids process model, stakeholders perceive its notation as understandable and the process model for the step “setting-up Q-Rapids” as complete and useful. Stakeholders also observed that the Q-Rapids process model reflects the specified activities, namely defining the Q-Rapids quality model, determining data sources and providing access to data, and installing the Q-Rapids tool. Participants also commented on suggestions of potential improvement of the Q-Rapids tool and Q-Rapids process model.



## 1. Introduction

### 1.1 Motivation

The acceptance of the Q-Rapids framework by end-users depends on, amongst others, understanding its impacts and associated risks in different application domains and project settings. Thus, we plan formative and summative evaluations to support the incremental development of the Q-Rapids framework and providing empirical evidence on its impacts. To achieve this, we defined the expected quality and impact of the Q-Rapids framework in deliverable *D5.1 Industry requirements and validation plan*. Moreover, we designed general empirical studies for evaluating the proof-of-concept, consolidated version, and final version of the Q-Rapids framework in deliverable *D5.2 Validation design*. In this document, we report the evaluation on the proof-of-concept of the Q-Rapids framework (hereafter referred as to Q-Rapids PoC).

### 1.2 Intended audience

This deliverable provides the researchers of WP1, WP2, WP3, and WP4 with the evaluation methodology used to get early feedback about the Q-Rapids PoC in different application domains and settings. Moreover, it provides them with suggestions for improving the Q-Rapids PoC in subsequent project phases.

### 1.3 Scope

The scope of this document includes the methodology used during *Task 5.3 -5.6 Execution of the use cases* to evaluate the Q-Rapids PoC and the corresponding results.

### 1.4 Relation to other deliverables

In this document, we further specify and enhance the empirical design reported in deliverable *D5.2 Validation plan* in order to evaluate the Q-Rapids PoC.

The Q-Rapids PoC consists of the versions of the Q-Rapids tool (including the Q-Rapids dashboard, quality model, and data collection and analysis engine) and the Q-Rapids process model at project month 15. Further description of the Q-Rapids dashboard can be found in the deliverable *D3.2 Dashboard proof-of-concept*. Further descriptions of the Q-Rapids quality model as well of the Q-Rapids data collection and analysis engine can be found in the deliverable *D1.2 data analysis and gathering proof-of-concept*. The Q-Rapids process model is available in the project repository.

### 1.5 Structure of the deliverable

This deliverable is organised into the following sections: in Section 2, we will provide an overview of the Q-Rapids PoC. In Section 3, we will briefly describe the evaluation design used to evaluate the Q-Rapids PoC. In Section 4, we will summarize the results of the evaluation of the Q-Rapids PoC and discuss the threats to validity of the performed evaluation. Finally, we will provide an overview of the next steps regarding the further development and evaluation of the Q-Rapids framework.

Given confidentiality concerns, we report the results of evaluating the Q-Rapids PoC in each company in a confidential annex, namely *D5.3 Evaluation of pilot cases (V1) – Annex (Confidential)*.

## 2. Object of study: Proof-of-Concept

The Q-Rapids PoC includes the first version of the Q-Rapids tool as well as the Q-Rapids process model.

The **Q-Rapids tool** is a data-driven, quality-aware tool for rapid software development in which quality and functional requirements are managed together. It aims at increasing the productivity of the development team and shortening time to market of software products and services whilst ensuring appropriate levels of quality. The Q-rapids tool comprises the Q-Rapids dashboard, the Q-Rapids quality model, and the Q-Rapids



data collection and analysis engine. At the time of the evaluation, the Q-Rapids tool provided product owners with support for analysing the quality of a software product and blocking situations.

The **Q-Rapids process model** describes the Q-Rapids process from the following viewpoints: (1) functional (what is done?), (2) organizational (who is doing it and in which unit it is done?), (3) behavioral (when and how it is done?), and (4) informational (what data elements participate and what relationship they have?). The evaluation of the Q-Rapids PoC focused on the “setting-up Q-Rapids” step. This process step consist of the activities of “Defining the quality model”, “Determining data sources and providing access to data”, and “Installing Q-Rapids tool” together with all relevant sub-activities and elements of those activities.

A complete description of the Q-Rapids dashboard functionality is available as User’s Guide that can be downloaded from the Q-Rapids project website (downloads section), jointly to a video tutorial<sup>1</sup> of the dashboard. The technical description of the Q-Rapids dashboard can be found in the deliverable *D3.2 Dashboard proof-of-concept*. The description of the Q-Rapids quality model as well of the Q-Rapids data collection and analysis engine can be found in the deliverable *D1.2 data analysis and gathering proof-of-concept*. The Q-Rapids process model is available in the project repository.

## 3. Evaluation Methodology

### 3.1 Overview

#### 3.1.1 Evaluation Goals and Questions

Regarding the **Q-Rapids tool**, we aimed at understanding its appropriateness for analysing the strengths and drawbacks of a software product as well as for analysing potential risks regarding the continuous software development (i.e., blocking situations) from the perspective of product owners. Moreover, we aimed at identifying needs for improvements regarding the Q-Rapids tools.

Appropriateness of the Q-Rapids tool means the degree to which product owners consider the underlying quality model as understandable, complete, useful, and relevant. Moreover, it means the degree to which product owners perceive the functionalities and features of the Q-Rapids tool as easy to use, efficient, reliable, useful, relevant, enjoyable, and easy to navigate. We also considered visualisations as part of the Q-Rapids tool appropriateness. Thus, we defined the following evaluation questions:

- Q1. **Appropriateness of the Q-Rapids quality model** – To what extent product owners perceive the Q-Rapids quality model as understandable, complete, useful, and relevant?
- Q2. **Appropriateness of the Q-Rapids tool’s functionalities and features** – To what extent product owners perceive the functionalities and features of the Q-Rapids tool as easy to use, efficient, reliable, useful, relevant, enjoyable, and easy to navigate?
- Q3. **Needs for improvements** – What needs to be improved to increase the appropriateness of the Q-Rapids quality tool?

Regarding the **Q-Rapids process model**, our evaluation goal was to understand the degree to which practitioners consider the process model for “setting-up Q-Rapids” as understandable, complete, useful and compatible with existing organisational processes. Practitioners encompass here the responsible persons for setting-up and installing the Q-Rapid tool in a given company. We also aimed at identifying needs for improvements regarding the Q-Rapids process model. Thus, we defined the following evaluation questions:

---

<sup>1</sup> <https://youtu.be/2m-dmJZiYBA>



**Q1. Appropriateness of the Q-Rapids process model** – To what extent practitioners perceive the Q-Rapids process model for “setting-up Q-Rapids” as understandable, complete, useful, and compatible?

**Q2. Needs for improvements** – What needs to be improved to increase the appropriateness of the Q-Rapids process model regarding the step “setting-up Q-Rapids”?

### 3.1.2 Evaluation Strategy

We designed two empirical studies for evaluating the Q-Rapids PoC, namely:

- **Evaluation of the Q-Rapids tool:** Semi-structured interviews in which product owners independently tried the Q-Rapids tool and commented on its appropriateness for analysing the strengths and drawbacks of a software product as well as for analysing potential blocking situations during rapid software development.
- **Evaluation of the Q-Rapids process model:** Semi-structured interviews in which practitioners independently explored the process model for “setting-up Q-Rapids”, and commented on its understandability, completeness, usefulness, and compatibility with existing organisational processes.

### 3.2 Evaluation of the Q-Rapids tool

**Procedure:** To answer the evaluation questions regarding the Q-Rapids tool, we performed semi-structured interviews with product owners. Each interview was conducted as follows:

First, we explained the evaluation goals and procedure to the participant and asked him/her to sign the informed consent.

Second, we trained the participant in the Q-Rapids tool. Using a standardised training video, we introduced the quality model as well as the functionalities and features of the Q-Rapids tool. The standardised video served to ensure equal treatments among different interviews and reduce experimenter bias. At the end of the training, we asked the participant how well prepared he or she felt to use the Q-Rapids tool and answered all questions the participants asked regarding the Q-Rapids tool.

Third, the participant used the Q-Rapids tool to explore the Q-Rapids tool, namely to analyse strengths and drawbacks in the software product as well as potential blocking situations during the software development. We encouraged the participant to think aloud and mention positive aspects and negative aspects of the Q-Rapids tool. This served to get a better understanding of the participant's insights on the Q-Rapids tool. Fourth, we collected further feedback regarding the appropriateness of the Q-Rapids tool by using a structured questionnaire. Finally, we asked the participant for open feedback on the Q-Rapids tool and the evaluation.

An observer documented the progress of each interview using a predefined observation protocol. The observer kept records of the participants' comments and questions on the Q-Rapids tool and of deviation from the evaluation procedures. This was intended to understand the quantitative evaluation of the Q-Rapids tool and to facilitate later analysis of possible threats to validity such as experimenter bias. The interviews were conducted by different experimenters and observers.

**Instruments and Materials:** Table 1 provides an overview of the data collection instruments and materials used during the evaluation. The corresponding instruments and materials are stored in the project repository.

We operationalized the appropriateness of the Q-Rapids tool based on the Likert-scales introduced in [1], [2], [3], [4], [5], and [6]. Each Likert-scale includes up to four statements to be rated using a response scale



from 1: strongly disagree to 5: strongly agree and an additional “I don’t know” option. We instantiated the selected questions according to the purpose and content of the Q-Rapids tool.

Table 1: Instruments and Materials

Instrument/Material	Purpose	Overview of content
Evaluation guideline	To provide experimenters and observers with a detailed description of the evaluation procedures and guidance to manage potential confounding factors.  To ensure an equal treatment among different interviews independent of the experimenter	Document explaining how to conduct the evaluation and how to manage potential confounding factors.
Informed consent	To enable participants to make an informed decision about accepting/declining their participation in the evaluation	Research purpose, study plan, data collection procedures, storage and use, access to raw data and results, planned publications, and procedures for data anonymity and confidentiality
Video Tutorial regarding the Q-Rapids tool	To train the participant in the Q-Rapids tool	Video explaining how to use the Q-Rapids tool and how to interpret the visualized data
Task description	To provide the participant with a description of the assigned task	Description of the tasks to be performed by the participant
Q-Rapids tool quick guide	To provide the participant with a quick overview of the Q-Rapids tool functionality	Document summarizing the options to access to the features and describing the charts used in the data views
Q-Rapids quality model	To provide the participant with a definition of the elements of the quality model	Document comprising the specification of the quality factors and metrics included in the Q-Rapids quality model
Q-Rapids tool evaluation questionnaire	To measure the appropriateness of the Q-Rapids tool	Questionnaire on the perceived understandability, completeness, usability, and relevance of the Q-Rapids quality model.  Questionnaire on the perceived ease of use, efficiency, reliability, usefulness, relevance, enjoyment and easy of navigation of the functionalities and features of the Q-Rapids tool
Q-Rapids tool observational protocol	To enable analysis of strengths and drawbacks regarding the Q-Rapids tool  To facilitate the analysis of threats to validity	Start time, attendees, end time, activities/events/deviations/disruptions, timestamp, memos

**Population and Sampling:** The target users of the Q-Rapids tool are mainly product owners. Taking into account the project plan and constraints, we drew a convenient sampling including product owners of the companies involved in the Q-Rapids project [8]. At the time of the evaluation, the participants were team members of the project selected in each company for evaluating Q-Rapids (See *D5.1 Industry requirements and validation plan*).

**Execution:** In December 2017, we installed the Q-Rapids tool in each company. In parallel, we trained the experimenters and observers responsible for performing the evaluation at each company. After collecting project data for at least two weeks in each company, we evaluated the Q-Rapids tools following the procedures described above in January 2018. We scheduled each evaluation session up to 120 minutes taking into consideration the availability of eventual participants.



### 3.3 Evaluation of the Q-Rapids process model

**Procedure:** To answer the evaluation questions regarding the Q-Rapids process model, we performed semi-structured interviews with several stakeholders. Each interview was conducted as follows: First, we explained the evaluation goals and procedure to the participants and asked them to sign the informed consent. Second, we presented a standardised video explaining how to read the process model. We answered any questions the participants had about the process model notation and specification. The standardised video was prepared to ensure equal trainings among different interviews and reduce experimenter bias. Third, we asked the participants to inspect the process model and evaluate the “setting up Q-Rapids” steps with the help of the provided questionnaire. Fourth, we asked each participant to write three most critical suggestions of improvements. Finally, we summarized the interview and quickly described the next steps.

**Instruments and Materials:** Table 2 provides an overview of the data collection instruments and materials used during the evaluation. The corresponding instruments and materials are stored in the project repository.

We operationalized the appropriateness of the Q-Rapids process model based on the Likert-scales introduced in [3] and [4]. Each Likert-scale includes up to four statements to be rated using a response scale from 1: strongly disagree to 5: strongly agree and an additional “I don’t know” option. We instantiated the selected questions according to the purpose and scope of the Q-Rapids process model.

Table 2: Instruments and Materials

Instrument/Material	Purpose	Overview of content
Evaluation guideline	To provide experimenters and observers with a detailed description of the evaluation procedures and guidance to manage potential confounding factors.  To ensure an equal treatment among different interviews independent of the experimenter	Document explaining how to conduct the evaluation and how to manage potential confounding factors.
Informed consent	To enable participants to make an informed decision about accepting/declining their participation in the evaluation	Research purpose, study plan, data collection procedures, storage and use, access to raw data and results, planned publications, and procedures for data anonymity and confidentiality
Video Tutorial regarding the Q-Rapids process model	To introduce the Q-Rapids process model to the participant and explain him/her how to read it	Video explaining the process model for “setting-up Q-Rapids”, how to inspect the model
Q-Rapids process model	To represent / illustrate the Q-Rapids process for the participant	Representation of the Q-Rapids process (activities, input/outputs, steps, prerequisites, roles) relevant in Q-Rapids)
Feedback questionnaire	To measure the appropriateness of the Q-Rapids process model	Questionnaire on the perceived understandability, completeness, usability, and compatibility of the process model for “setting-up Q-Rapids”.
Q-Rapids tool observational protocol	To enable analysis of strengths and drawbacks regarding the Q-Rapids process model  To facilitate the analysis of threats to validity	Start time, attendees, end time, activities/events/deviations/disruptions, timestamp, memos

**Population and Sampling:** To evaluate the process model for “setting-up Q-Rapids”, we drew a convenient sample including stakeholders involved or affecting decisions in the sub-activities of the setting up of Q-



Rapids process [8]. That is, roles involved in defining the Q-Rapids quality model, determining data sources, providing data access, or installing the Q-Rapids tool. Thus, we selected project managers, technical leads, quality managers, software architects and testers.

**Execution:** In December 2017, we trained the experimenters and observers responsible for performing the evaluation at each company. Then, we evaluated the Q-Rapids process model following the procedures described above in January 2018.

### 3.4 Data Preparation and Analysis Methods

After each interview, the experimenter and observer transcribed the participants' answers and the observation protocol into an interview protocol consisting of two standardised Excel templates. This served to keep the data analysis consistent among companies. Then, we carried out the quantitative and qualitative analysis.

We first carried out within-case analyses of the quantitative and qualitative data collected for each company. Then, we compared, related, and integrated the results among the companies (cross-case analyses) [9].

In this deliverable, we report descriptive statistics including the sample size (N), median (Mdn), minimum (Min), maximum (Max), and modal value (Mode) for the quantitative analyses. Regarding the qualitative analysis, we used data-driven thematic analysis [10] to analyse participants' feedback on the Q-Rapids quality model. At least two researchers derived themes – i.e., explicitly mentioned suggestions of improvement – inductively by coding and interpreting all observation protocols independently of each other. Then, they compared their results and solved any deviations. Moreover, we performed several peer-review meetings including all experimenters, observers, and analysts to review the interpretations of the elicited qualitative data. This served to keep the qualitative analyses grounded on the collected evidence and rigorous.

## 4. Results: Cross Case Analysis

At the time of submitting this deliverable, we had the consent of three out of four companies to report the corresponding evaluation results in an aggregated form. Thus, we include in this section only the results regarding three out of four use cases in which we carried out the evaluation of the Q-Rapids PoC.

### 4.1 Appropriateness of the Q-Rapids Tool

#### 4.1.1 Sample

In total, one product owner, four project managers, and one developer participated in the evaluation. They had at least three years of working experience in their companies (Mdn = 10, Min = 3, Max = 30) and at least half a year of working experience in their current role (Mdn = 10, Min = 0.5, Max = 30).

#### 4.1.2 Appropriateness of the Q-Rapids Quality Model

Table 3 presents the perception of the participants on the Q-Rapids appropriateness. The majority of the participants claimed the assessed metrics included in the implemented Q-Rapids quality model are understandable (N = 5, Mdn = 4, Min = 2, Max = 5, Mode = 4). One participant pointed out: *"I need more clarification and details"* to understand the assessed metrics. This participant and another recommended including the actual values of the metrics – i.e., raw values before the normalization through utility functions. In general, most of the participants had difficulties in understanding the normalized values as stated by one participant: *"I don't understand [...] what it means. I know 1 is good and 0 is bad, but what about when it is 0.91?"* Moreover, they perceived the quality factors (N = 5, Mdn = 3.5, Min = 2, Max = 5, Mode = N.A.) and the strategic indicators (N = 5, Mdn = 3.75, Min = 3, Max = 5, Mode = N.A.) as moderately understandable. In general, the participants proposed to avoid negated formulations of the factors or metrics to increase their understandability. For instance, to state complex files instead of non-complex files.



Table 3: Appropriateness of the Q-Rapids quality model

Participants perception on ...*	N	Mdn	Min	Max	Mode
Relevance	5	4	3	4.5	4
Right level of Detail	5	3	3.5	5	3
Understandability: Metrics	5	4	2	5	4
Understandability: Quality Factors	5	3.5	2	5	N.A.
Understandability: Strategic Indicators	5	3.75	3	5	N.A.

\*: Each quality aspect was measured using a valid and reliable Likert-scale. Each item was rated using a five-point response scale from 1: strongly disagree to 5: strongly agree and incl. the option "I don't know".

All participants agreed the current Q-Rapids quality model is relevant for their work (N = 5, Mdn = 4, Min = 3, Max = 4.5, Mode = 4), but they perceived the Q-rapids quality model fairly provide information at the right level of detail (N = 5, Mdn = 3, Min = 3.5, Max = 5, Mode = 3). They recommended to link the provided information of the quality factors as well as the assessed metrics with further information sources (e.g., issue reports) in order to better support the decision making process. The participants agreed that integrating several heterogeneous data sources is an added value for supporting the decision making processes in their companies.

#### 4.1.3 Appropriateness of the Q-Rapids tool's functionalities and features

Table 4 provides an overview of the perception of the participants on the Q-rapids tool's functionalities and features. The participants agree the current version of the Q-Rapids tool is usable (N = 5, Mdn = 4, Min = 2.5, Max = 5, Mode = 4) and relevant for their work (N = 5, Mdn = 4, Min = 3, Max = 5, Mode = 5). One participant argue the Q-Rapids tool provide them with information for improving discussions and with support for a closer work between managers and developers. This participant pointed out: *"This tool make closer the manager to the development team, you can rationalize the subject"* (i.e., the product and project). The same participant also observe the Q-Rapids tool help to identify project issues. Another participant perceive the integration of data from several data sources as an important added value from the Q-Rapids tool.

Table 4: Appropriateness of the Q-Rapids functionalities and features

Participants perception on ...*	N	Mdn	Min	Max	Mode
Usability	5	4	2.5	5	4
Relevance	5	4	3	5	5
Efficiency	5	4	3.5	5	4
Reliability	5	4	3	5	5
Results traceability	5	4	4.5	4	4
Ease of Use	5	4	3	5	5
Navigation	5	4	3	5	4
Visualization	5	3	2.5	5	3
Enjoyment (**)	5	2	0	3	N.A.

\*: Each quality aspect was measured using a valid and reliable Likert-scale. Each item was rated using a five-point response scale from 1: strongly disagree to 5: strongly agree and incl. the option "I don't know".

\*\* : Seven-point response scale from -3 very negative to 3 very positive.

Furthermore, the participants are convinced they can work efficiently using the Q-Rapids tool when analysing the strengths and drawbacks of a software product or when analysing potential risks regarding the continuous software development (N = 5, Mdn = 4, Min = 3.5, Max = 5, Mode = 4).

All participants consider the information provided by the Q-Rapids as traceable (N = 5, Mdn = 4, Min = 4.5, Max = 4, Mode = 4) and four out of five as reliable (N = 5, Mdn = 4, Min = 3, Max = 5, Mode = 5). One participant perceives the metrics as reliable, but not the strategic indicators or the quality factors. Regarding the historical view, another participant emphasized, *"... the trend is the most important thing! We are constantly changing things so we need to see that."*



The participants observe that the Q-Rapids tool is easy to use (N = 5, Mdn = 4, Min = 3, Max = 5, Mode = 5) and navigate (N = 5, Mdn = 4, Min = 3, Max = 5, Mode = 4) and it includes reasonable visualizations (N = 5, Mdn = 3, Min = 2.5, Max = 5, Mode = 3). In particular, the participants consider the digging navigation and the historical views as very positive features of the Q-Rapids tool. One participant commented: *"I appreciate the digging feature to inspect the rationale behind the elements."* Another participant stated: *"It is a very good idea adding the historical views ... the user can analyse trends ... The historical views helps you to identify "bad criteria", when you see that there is no evolution in one factor, you can talk to the pertinent person."* Another participant recommended amongst others to improve the used visualization and to provide more information directly in the tool by using e.g., tooltips or legends explaining all elements inside a graph. The participant stated that support for interpreting values (e.g., the coloured areas of the gauge charts) or guidance of how to draw decisions should be improved. For instance, she or he asked: *"what does it mean if a gauge chart indicates 30% of blocking, is this good or bad? And how can someone make decisions based on that?"*.

In general, the participants assessed the current version of the Q-Rapids tool as enjoyable.

#### 4.1.4 Needs for Improvements

Table 5 provide an overview of the suggestions of improvements mentioned by the participants during the evaluation of the Q-Rapids tool.

Table 5: Suggestions of improvements for the Q-Rapids tool

Suggestions of improvements for the Q-Rapids tool	N
Add definitions of metrics, quality factors and strategic indicators (not self-explanatory).	5
Include support for determining thresholds values: thresholds can vary during the project	4
Enhance the visualisation of quality factors - i.e., radar diagram	3
Keep consistent visualizations (e.g. colours) by the graphical/textual view when navigating through the quality model. Include graphical views for metrics	3
Make clear the added value regarding existing tools such as SonarQube	2
Add link to raw data for supporting the decision making process	2
Adding information for properly interpreting visual clues, e.g., meaning of "zero-value", exact values (i.e., values before normalization), thresholds, measurement units, and chart legends	2
Add scrolling, sorting and filter functionalities (e.g. selecting and ordering metrics).	1
Consider different points in time for data analysis, e.g., sprints, milestones, releases (instead of daily).	1
Add functionality for allowing the user to add/ configure the quality model (strategic indicators, metrics)	1

#### 4.1.5 Threats to Validity

We evaluated the Q-Rapids tool by drawing a convenient sample of product owners (Selection Bias) working in several application domains and software development environments. Thus, our results are tied to the context of the companies involved in the Q-Rapids project. Furthermore, the evaluation included only one treatment – using the Q-Rapids tool – (Mono-Operation Bias) and the evaluation result shows low statistical power (Conclusion validity). The tasks performed by the participants had a limited scope and focused on analysing strengths and drawbacks of software products as well as blocking situations, but not on making decisions. Therefore, the results can only be interpreted as an indication of the appropriateness of the Q-Rapids tool. The Q-Rapids quality model might serve as a basis for supporting the analysis of strengths and drawbacks of software products and blocking situations in companies applying rapid software development in a similar setting. Further evaluations in different company settings including a larger sample of decision makers and alternative treatments are required for generalizing results to other organisations applying rapid software development.

Moreover, we observed two important confounding factors during the execution of the interviews. First, one participant has difficulties in understanding the questionnaires related to the evaluation of the Q-Rapids



quality model. To avoid internal and conclusion validity issues, we omitted his answers during the data analysis. Second, during one interview non-announced project team members (outside the target population) attended. The corresponding experimenter and observer reported they are convinced these attendees did not influence the participant perception.

## 4.1 Appropriateness of the Q-Rapids Process Model

### 4.1.1 Sample

In total, one consultant, one technical leader, two developers, and two managers participated in the evaluation. They had at least two years and a half of working experience in their companies (Mdn = 9, Min = 2.5, Max = 19) and at least five years of working experience in their current role (Mdn = 10, Min = 5, Max = 10).

### 4.1.2 Appropriateness of the Process Model for “Setting-Up Q-Rapids”

Table 6 presents the perception of the participants on the Q-Rapids process model. The participants consider the notation used for specifying the Q-Rapids process model as understandable (N = 6, Mdn = 4, Min = 3.5, Max = 5, Mode = 4). All participants perceive the specification of the “setting up Q-Rapids” step as useful (N = 6, Mdn = 4, Min = 2, Max = 5, Mode = 4), understandable (N = 6, Mdn = 4, Min = 3.5, Max = 5, Mode = 4), and complete (N = 6, Mdn = 4, Min = 3, Max = 4, Mode = 4). However, they claim the current model is inconsistent (N = 6, Mdn = 2.5, Min = 2, Max = 4, Mode = 4). One participant stated that the Q-Rapids process model might be not so easy to read for people that do not know the used notation. Furthermore, the participant explains: *“It” - i.e., the process model specification - “is easy to comprehend ... specially after having seen the introductory video.”* Moreover, this participant points out: *“After reading the process model specification people will be not ready to install Q-Rapids”*. The participant suggests to link to further information or documents, e.g., an installation checklist or a detailed “how to” guidance for eliciting the quality model, getting access to data sources, or installing connectors, etc.. This could be especially helpful for new users.

Table 6: Appropriateness of the Q-Rapids process model regarding the step “setting-up Q-Rapids”

Participants perception on ...*	N	Mdn	Min	Max	Mode
Notation Understandability	6	4	3.5	5	4
Model Usefulness	6	4	2	5	4
Model Understandability	6	4	3.5	5	4
Model Completeness	6	4	3	4	4
Model Consistency	6	2.5	2	4	4

\*: Each quality aspect was measured using a valid and reliable Likert-scale. Each item was rated using a five-point response scale from 1: strongly disagree to 5: strongly agree and incl. the option “I don’t know.”

In general, all participants agree on the need of having “how-to” guidelines for several activities such as synchronizing Q-Rapids tool with connectors, customizing connectors, linking data sources with the quality model, defining the quality model, and solving data integration issues. Two participants also emphasize the need for including examples in order to increase the understandability of the Q-Rapids process model. One participant explained *“Everywhere we should have examples, e.g. defining quality model, we need examples of each step, it is much more easier to read and think about what it means for me, e.g. this quality model thinking it’s really hard when thinking what this means, it is much better to have examples.”*

### 4.1.3 Needs for Improvements

Table 7 provides an overview of the suggestions of improvements mentioned by the participants during the evaluation of the Q-Rapids process model for the step “setting-up Q-rapids”.



Table 7: Suggestions of improvements for the Q-Rapids process model

Suggestions of improvements for the Q-Rapids tool	N
Include guideline for setting up the quality model (procedures, methods, roles, responsibilities, etc.)	5
Include examples, glossary, overview of roles (e.g., hierarchy)	2
Include guideline for customizing connectors	2
Include guideline for linking data sources with the quality model	2
Include guideline for identifying required data sources	2
Include links to guidelines not included in the process specification	2
Include guideline for synchronizing Q-Rapids tool with connectors	1
Include pre-requisites for starting using Q-rapids	1
Include role pre-requisites regarding knowledge, skills, etc.	1
Revise activity sequence (parallelize e.g. setting up connectors and installing dashboard)	1
Include "estimated time/effort" per process activity	1

#### 4.1.4 Threats to Validity

We evaluated the Q-Rapids tool by drawing a convenient sample of stakeholders responsible for setting-up Q-Rapids in a given company (Selection Bias). Thus, our results are tied to the context of the companies involved in the Q-Rapids project. Furthermore, the evaluation included only a walkthrough through the Q-Rapids process model for one specific step, namely “setting-up Q-Rapids” (Mono-Operation Bias) and the evaluation result shows low statistical power (Conclusion validity). Therefore, the results can only be interpreted as an indication of the appropriateness of the process model for the step “setting-up Q-Rapids”.

## 5. Conclusions

In this deliverable, we summarise:

- The evaluation design for the Q-Rapids PoC, namely the Q-Rapids tool and the Q-Rapids process model for the step “setting-up Q-Rapids”
- The results regarding the appropriateness of the Q-Rapids tool and the Q-rapids process model
- The suggestions of improvements provided by the evaluation participants regarding Q-Rapids tool and the Q-rapids process model.

These results enable the further development of the Q-Rapids framework. They will be discussed and prioritized by the project partners taking into account the project goals and constraints.

Here it is important to remark that during *Task 5.3 - 5.6 Execution of the use cases* we also designed and conducted several evaluations to support the development of the Q-Rapids PoC between the project months 9 and 15. The scope of these evaluations includes following object of study and evaluations:

- Q-Rapids dashboard: Four online workshop to get early feedback on the mock-ups of the Q-Rapids dashboard (June-August 2017), two onsite feedback workshops to get early feedback on the strategic indicators included in the Q-Rapids dashboard (September 2017), and one feedback workshop to get early feedback on an executable mock-up of the Q-Rapids dashboard (October 2017).
- Q-Rapids quality model: Two onsite feedback workshop (September 2017) to get early feedback on the quality factors and metrics included in the Q-Rapids quality model.
- Q-Rapids process model: Two onsite feedback workshop (September 2017) to get early feedback on the Q-Rapids process notation and model.



## References

- [1] V. McKinney, K. Yoon, and F. M. Zahedi, "The measurement of web customer satisfaction: An expectation and disconfirmation approach," *Information Systems Research*, vol. 13, no. 3, pp. 296–315, 2002.
- [2] R. R. Nelson, P. A. Todd, and B. H. Wixom, "Antecedents of information and system quality: An empirical examination within the context of data warehousing," *Journal of Management Information Systems*, vol. 21, no. 4, pp. 199–235, 2005.
- [3] D. L. Goodhue and R. L. Thompson, "Task technology fit and individual performance," *MIS Quarterly*, vol. 19, no. 2, pp. 213–236, 1995.
- [4] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decision Sciences*, vol. 39, no. 2, pp. 273–315, 2008.
- [5] P. Xu and B. Ramesh, "Impact of knowledge support on the performance of software process tailoring," *Journal of Management Information Systems*, vol. 25, no. 3, pp. 277–314, 2008.
- [6] L. Goel, N. A. Johnson, I. Junglas, and B. Ives, "From space to place: Predicting users' intentions to return to virtual worlds," *MIS quarterly*, vol. 35, no. 3, pp. 749–772, 2011.
- [7] J. Cresswell and V. L. P. Clark, *Designing and conducting mixed methods research*, 2nd ed. London: SAGE Publications, 2011.
- [8] J. Daniel, *Sampling essential. Practical guidelines for making sampling choices*. SAGE Publications, 2012.
- [9] M. Miles and M. Huberman, *Qualitative data analysis*, 2nd ed. London: Sage Publications, 1994.
- [10] V. Braun and V. Clark, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, pp. 77–101, 2016.