Fraunhofer Institut
Experimentelles
Software Engineering

GQM-Handbook and Overview of GQM-plans

A comparison of three GQM-methods and an Overview of GQM-plans

Authors:
Alessandra Anacletto
Teade Punter
Christiane Gresse von Wangenheim

IESE-Report No. 008.03/E
Version 1.0
April 2003

A publication by Fraunhofer IESE

Fraunhofer IESE is an institute of the Fraunhofer Gesellschaft.
The institute transfers innovative software development techniques, methods and tools into industrial practice, assists companies in building software competencies customized to their needs, and helps them to establish a competetive market position.

Fraunhofer IESE is directed by
Prof. Dr. Dieter Rombach
Sauerwiesen 6
D-67661 Kaiserslautern

Abstract

The *Goal/Question/Metric* Paradigm (GQM) is a goal-oriented approach for the measurement of products and processes in Software Engineering, which has been applied successfully in several companies. The foundation of the GQM Paradigm has been established by Basili and Weiss in 1984 and since then evolved significantly.

This document intends on the one hand to provide support for the development of GQM-based measurement programs by unexperienced people as well as a reference guide for experienced people. The document consists basically of two parts: one describing the GQM process for planning and executing a measurement program and the second part is an overview on various measurement programs. The GQM process presented in this document is the result of the unification of three GQM processes as described in Solingen&Berghout 1999, Gresse von Wangenheim 2002 and Epg 2002. In addition, further lessons learned from GQM measurement programs established by the IESE (Germany) and UFSC (Brazil) are included especially considering guidelines for the application of measurement in small software companies. However, this document is not supposed to be a complete and detailed description of the GQM process. Its objective is rather to provide an overview by integrating various experiences.

This document is the result of the scientific mission of Alessandra Anacleto during July 1st 2002 and September 30th 2002 at the IESE in the context of the international cooperation project "Tailoring Software Measurement for Small and Medium Enterprises" between the Fraunhofer IESE (Germany) and the Federal University of Santa Catarina (Brazil) [GS00]. The goal of the cooperation project is to facilitate the establishment of software measurement in small software development organizations in order to enable efficient quantitative project management and quality and productivity improvement as a basis for the success and competitiveness of the company.

Table of Contents

1 Introduction

Software measurement has been recognized as a key infrastructure to improve the quality and productivity of software [BCR94b]. In order to understand the current situation and build organisation-specific models, we need to measure quantitative and qualitative data on software quality and productivity and potential influence factors. Only based on data collected in the specific environment, quality models can be build, which reflect the present situation in the specific organisation as a basis for effective project management and quality improvement.

In order to make precise statements about quality of software products and processes, this quality has to be described in a quantitative way. However, this is quite difficult for many quality aspects of software products and processes, e.g., such as userfriendliness or complexity. The problem is the identification of relevant characteristics of products or processes, which describe a desired quality aspect in a particular context and its interpretation in the specific context. In general, there exists dozens of measures that can be applied to a wide variety of project, process and product attributes, as, for example, effort, personnel characteristics, change and defect data, process definition, or logical product characteristics [Bas92a, FP97]. And, measurement can be done in many different ways, for example, size of a software module can be measured by counting lines of code, computing function points or counting the number of classes. Thus, the problem is to decide what should be measured in a specific context, who and when the data should be collected and how the measures should be used for management decisions and improvement opportunities. These issues need to be explicitly defined in a Measurement plan.

Thus, in order to operationalise the analysis and interpretation of goals to be achieved through measurement, a measurement approach is required. A measurement approach is a method for deciding what to measure how and how to interpret the collected data. Basically, there exist two types of measurement approaches [PR94]: top-down or bottom-up. Bottom-up approaches start with measurable observations and build up to management objectives and goals, whereas, top-down approaches help to derive useful measures from goals and to interpret the collected data in the context of the goals of interest [BCR94b]. Top-down approaches have shown to support the adequacy, consistency and completeness of the Measurement plan and to help to manage the complexity of measurement programs [Rom91]. Different top-down measurement approaches have been developed [RJ94], such as, the *Quality Function Development Approach* (QFD) [KA83], the *Software Quality Metrics Approach* (SQM) [Mur80] and the *Goal/Question/Metric Approach*

(GQM) [BCR94b, BR88, BW84]. Those top-down approaches focus on the relation of the measures to the measurement goals, yet differ significantly in terms of the scope of supported measurement goals, guidance for the identification of the relevant metrics and their potential uses. In that sense, the GQM approach has been shown to be the most flexible and widely applicable of the three approaches [Rom91].

This document intends on the one hand to provide support for the development of GQM-based measurement programs by unexperienced people as well as a reference guide for experienced people. The document consists basically of two parts: one describing the GQM process for planning and executing a measurement program and the second part is an overview on various measurement programs applied by IESE and Centro GeNESS. The GQM process presented in this document is the result of the unification of three GQM processes as described in Solingen&Berghout 1999, Gresse von Wangenheim 2002 and Epg 2002. In addition, further lessons learned from GQM measurement programs established by the IESE (Germany) and UFSC (Brazil) are included especially considering guidelines for the application of measurement in small software companies.

First the GQM approach is briefly presented in terms of its history, its objectives and references. A comparison between three processes, in which this handbook is based, is presented. The defined process is presented according to the following structure:

- For each phase, its objectives are presented and its relationship with respect to the different processes is presented in a table.
- For each activity in the context of its specific phase, its objectives are presented and its relationship with respect to the different processes is presented in a table. The tasks that should be realized for each activity are presented and some guidelines about how these tasks should be conduct, or some tips about the specific task are given. Finally, the expected output artifacts are presented for each activity.

This document is not supposed to be a complete and detailed description of the GQM process. Its objective is rather to provide an overview by integrating various experiences. For further details on specific activities we refer to the respective GQM process description from where the activity has been derived.

2 The GQM Approach

The *Goal/Question/Metric* Paradigm (GQM) [BCR94b, BW84] has been proposed as a goal-oriented approach for the measurement of products and processes in Software Engineering. It has been applied with success in several companies, such as NASA Goddard Space Flight Center (USA) [Bas92], Robert Bosch GmbH (Germany) [Cemp96], Allianz Lebensversicherungs-AG (Germany) [GRR94], Digital SPA (Italy) [Cemp96], Motorola [Das92] (USA), Schlumberger (Netherlands) [Cemp96], SIA (Italy) [Cug97], Tenovis [AKK00], Bauer&Partner [HKFP02], Dräger [BHP+97], Tokheim [Sol97] and small start-up software companies at the incubator Centro GeNESS (Brazil) [AG02].

The GQM Paradigm is based upon the assumption that for an organization to measure in a purposeful way it must identify and precisely specify the organizational and project-specific goals. Those goals have to be traced to the data that are intended to define the goals operationally and provide a framework for analyzing and interpreting the data with respect to the stated goals. Therefore, the analysis task is explicitly specified by a measurement goal. The goals are defined in an operational way by refining them in a top-down fashion into a set of measures via a set of questions and models, which formalize the informally stated questions and guide the definition of relevant measures. This refinement is precisely documented in a GQM plan, providing a rationale for the selection of the underlying metrics. The collected data is analyzed and interpreted in a bottom-up fashion in the context of the GQM goal and questions considering the limitations and assumptions underlying each metric (see Figure 1).

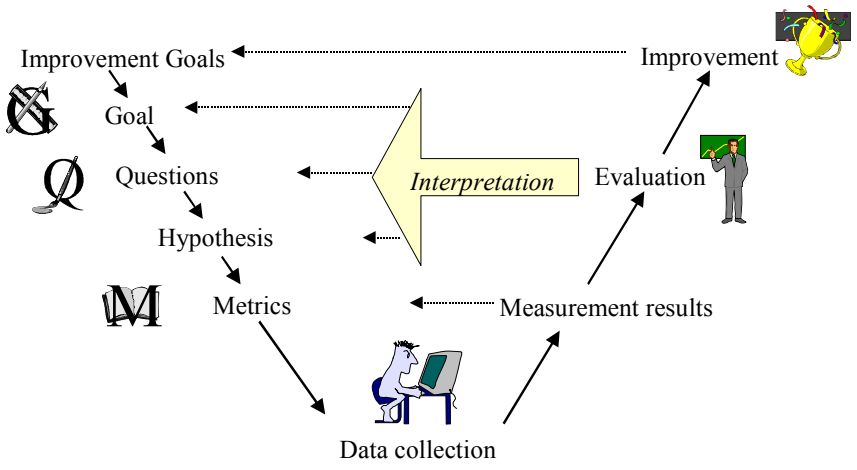

Figure 1 Overview of GQM-based Improvement

The definition and interpretation of the measurement data has to be tailored to the needs and characteristics of the specific context in order to be valid. Therefore, the stakeholders in the software project, e.g., the manager and developers, have to be actively involved in order to capture their implicit understanding and know-how.

Following these principles, the approach has several advantages: it supports the identification of useful and relevant metrics as well as it guides the analysis and interpretation of the collected data. It allows an evaluation of the validity of the obtained models and conclusions. In addition, GQM helps to avoid resistance against measurement through the involvement of the persons related to the object of study throughout the measurement program´s planning and execution [Rom91].

3 A Comparison of GQM Methods

In this chapter the history of the GQM process is briefly presented according to its evolution from [BW84] when the concept of goal oriented measurement was formulated to recent enhancements. The three methods selected to be the basis for this handbook are shortly presented describing their similarities and differences.

3.1 GQM Method – Its history from Basili & Weiss (1984) to Gresse (2002)

Basili and Weiss in their document "A Methodology for Collecting Valid Ssoftware Engineering Data" established the foundations of goal oriented measurement in 1984 [BW84]. Since that, the GQM approach was defined, which is a goal driven software measurement approach, and has evolved a lot until now. The first evolution was the "Experience Factory" [BR88], in which the concept of keeping experiences to permit reusing them later was introduced. These two documents were the basis for the GQM process.

In 1992, Basili for the first time defined the GQM process [Bas92a]. In the same year, Rombach defined the structure for GQM plans [Rom92] and in the next year a formal description of GQM plans was established by Differding [Dif93]. Based on the GQM process defined by Basili and on the experience factory, in 1995, the process was refined during the CEMP project by Gresse, Hoisl and Wuest [GHW95]. In the next year, SEI defined a handbook, based on Basili's GQM process where the goal-driven software measurement is presented focused on goal definition [PGF96].

Recently, two new evolution of the process were conduct mainly based on the CEMP experiences: in 1999 van Solingen and Berghout refined the GQM process based on their experiences in the CEMP project and by applying GQM at Schlumberger [SB99]. In 2001, Hamann, Kempkens and Zettel modelled a GQM process especifically for web-application projects in the context of the project IPSE PAMS [HKZ01]. Recent enhancement of the GQM process are mostly directed into formalizing the development of the GQM plan based on explicit quality models, as e.g., by Gresse [Gre02a] or others [CD99,MBB+98,OJ97,GM97]. Other enhancement focus principally on the reuse of experiences, such as the systematic reuse of measurement products described by Gresse [Gre02b] or the maintenance of mesurement plans described by Differding [Dif01]. A timeline graph for the basic evolutions is shown in Figure 2.

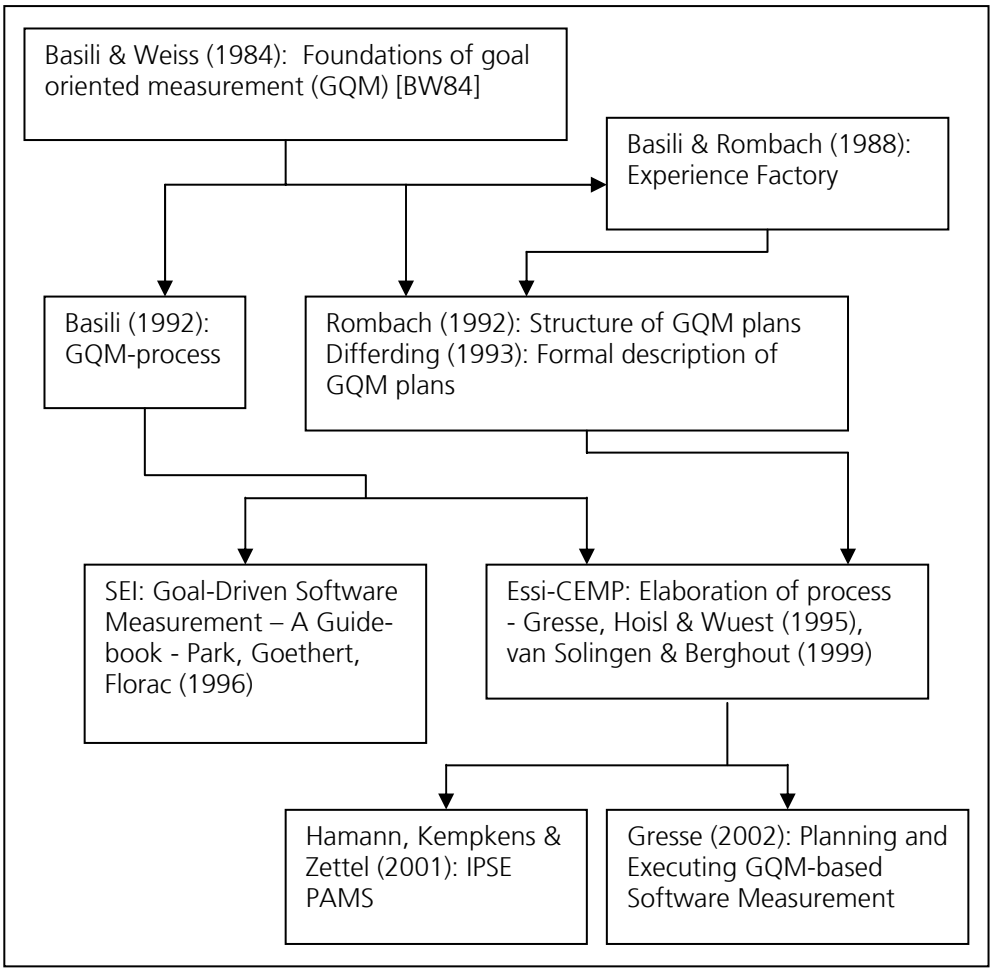

Figure 2 GQM timeline

For this handbook, the three more recent processes were considered: van Solingen & Berghout [SB99], Hamann, Kempkens & Zettel [HKZ01] and Gresse [Gre02a]. All these processes have the same basis: the CEMP project enabling a unification of the processes and, in addition, reflect our experiences in applying the GQM approach in practice.

3.2 Similarities and Differences in the Selected Processes

Table 1 summarises the comparison between the three processes phases. Note that each column is referred to each process and in each line the similar phases for each specific process are cited. Some phases are repeated in two lines, which means that this phase is related to more than one phase of other processes.

Table 1 GQM Phases – a comparison of three methods

	Gre02a	HKZ01	SB99
GQM1	- Prestudy	- Planning	- Planning
GQM2	- Identification of GQM goal(s) - Development of GQM plan(s)	- Planning	- Definition
	- Development of Data collection plan	- Instrumentation - Data collection and storage	- Definition - Data collection
GQM3	- Data collection, validation and storage	- Data collection and storage	- Data collection
GQM4	- Data analysis and interpretation	- Data analysis and report generation	- Interpretation
GQM5	- Post-mortem analysis and capturing of experiences	- Data analysis and report generation	

During the first phase, in GQM1, activities related to initiation of the measurement program are kept. The organisation and its software projects are characterised, their goals are identified and people are trained and motivated. There are no important differences between processes in this phase.

In the second phase, GQM2, the measurement goals are defined and formalised into quality and variation factors resulting in the GQM plan. The GQM goals are specialised into questions and these questions into relevant metrics. Finally, the Data collection plan is defined containing the data collection procedures and instruments (which are defined/constructed during this phase, too). The GQM plan and the Data collection plan form the Measurement plan, which is the plan to be followed during the execution of the measurement program. Some differences can be noted between these three processes in this phase, the most important ones are presented here:

- In GQM2 [Gre02a] is suggested the development of models, which are supposed to help in the metrics definition and to guide the analysis later.
- In GQM2 [HKZ01] and [SB99] is suggested the definition of indicators, which are supposed to help in the metrics definition and they are a basis for the analysis activity to be kept later.

In the third phase, the execution of the measurement program starts based on the Measurement plan defined in the first phase. During this phase, in GQM3, data is collected according to the procedures defined in the Data collection plan. Collected data is validated and stored in accordance to the Data collection plan. The most important difference between processes is that in GQM3

[HKZ01] and [SB99] is suggested the preparation of a measurement database, where the collected data can be stored.

Collected data is then analysed during the fourth phase, GQM4, supported by the indicators defined in the GQM plan. The interpretation activity is realised in a feedback session. There are no great differences between processes in this phase.

Finally, the last phase, GQM5, is related to capture experiences and package results. This phase is suggested by GQM5 [Gre02a] and [HKZ01].

The union of this three process results in a process that is presented in a flow diagram (Figure 3). It does not mean that this union composes a better process, the objective is rather to group the information from these processes in one place. According to this diagram, the introduction to the measurement program is the first phase to be executed, when characterisation activities are executed and all involved people are trained and motivated. The next phase is the definition phase. Here, measurement goals are defined and refined into metrics and data collection procedures and instruments are defined. The following phase is data collection, when the measurement program is effectively executed according to the Measurement plan defined in the previous phases. Collected data is then analysed and interpreted during feedback sessions. After this phase it should be defined if the measurement program is terminated or if it should proceed. If the program is terminated, then the packaging phase is executed, otherwise, it is necessary to define, if the measurement program has to be updated. This is, generally, be decided during the feedback session, when the measurement program should also be reviewed. If no change is required, then data collection can proceed unchanged, otherwise the Measurement plan should be reviewed and according be updated.

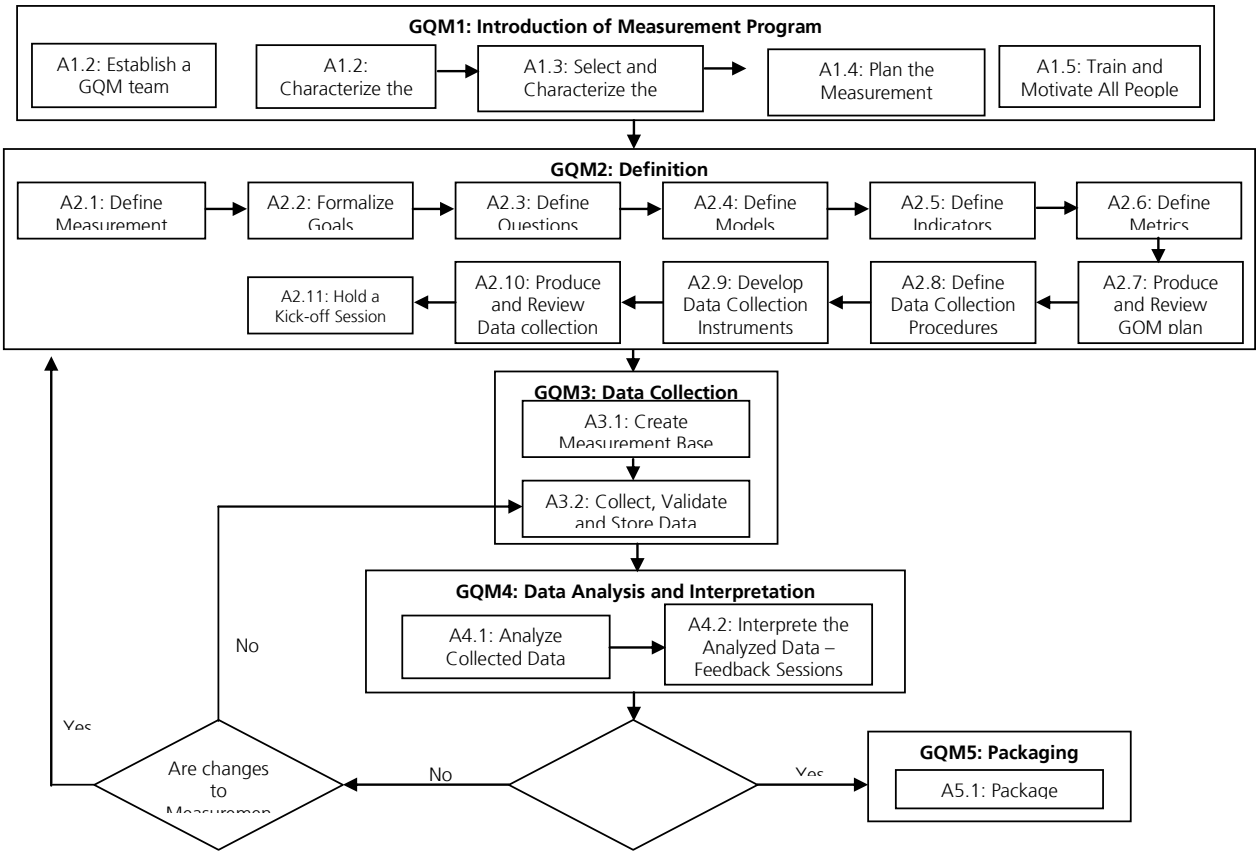

Figure 3 Process Flow Diagram

4 Expected Results of a GQM Based Measurement Program

In this section, we present the principal measurement products to be produced during a measurement program.

The measurement program results are stored in a document named Measurement plan. This Measurement plan presents the organization improvement goals and issues related to the measurement program execution. It is composed of two most important parts: the GQM plan and the Data collection plan. The GQM plan includes the organization and project caracterization, goal definition and its derivation into metrics. The Data collection plan describes the data collection procedures and instruments needed. Figure 4 presents the Measurement plan.

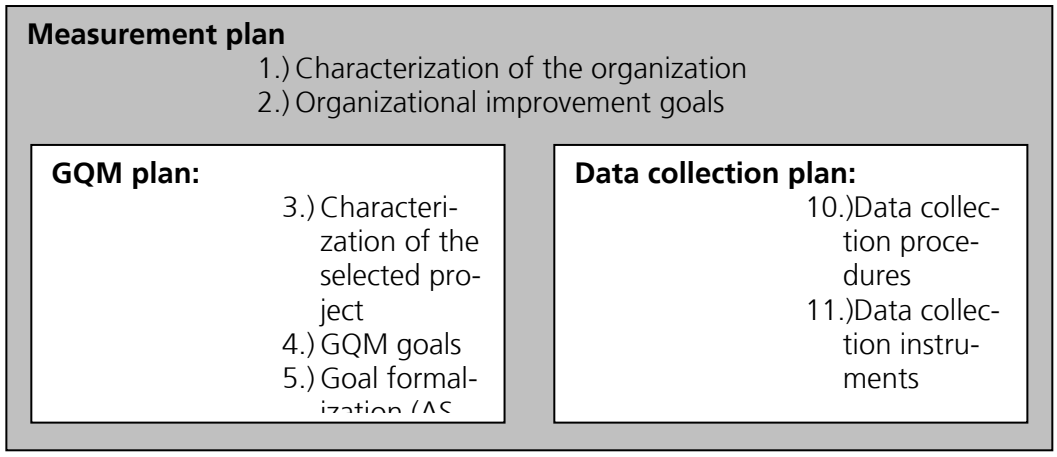

Figure 4 Measurement plan

The Measurement plan can include other necessary sections, for example, a section on the analysis results after feedback sessions can be kept.

The next sections describe the activities of this handbook according to the following structure:

GQMx: Identifies the phases that were defined unifying the processes in 0, "x" is a number, which identifies the specific phase (GQM1 – GQM5).

Ax.y: "A" means "Activity", in this way the main activities of each phase (x) are presented. "y" is a number, which identifies the specific Activity (A1.1 – A5.1) of phase "x".

Tx.y.z: "T" means "Task", in this way the tasks that should be executed during each activity (x.y) are presented. "z" is a number, which identifies the specific Task (T1.1.1 – 5.1.1).

Gx.y.z.k: "G" means "Guideline", in this way the guidelines to help better executing each task (x.y.z) are presented. "k" is a number, which identifies the specific Guideline (G1.1.1.1 – G5.1.1.5).

Figure 5 presents an example of this structure.

GQM1(*Phase name*): Introduction of Measurement Program
A1.1(*Activity name*): Characterize the organization
T1.1.1(*Task description*): Characterize the organization in which the measurement program takes place (e.g., wrt. application domain and business sector) [Gre02a].
G1.1.1.1(*Guideline description*): The organizational and project characterization can be supported by questionnaires. Note that

Figure 5 Structure example

In general, all tasks should be executed in order to establish a successful GQM-based measurement program. The guidelines provide some tips that can facilitate the conduction of the tasks. Special guidelines focusing on the establishment of measurement programs in small companies are provided, indicated by "For small companies" in the beginning of sentence.

4.1 Introduction of Measurement Program

This first phase prepares the organization for starting software measurement. During this phase a GQM team is established, the organization and its projects are characterized. Normally, one software project is selected to introduce the measurement and all people involved in the measurement program are trained and motivated.

Some activities of this phase are only necessary when measurement is being introduced in the organization for the first time. This is especially related to characterization activities, e.g., "Characterize the Organization", which do not need to be repeated every time a measurement program is established in a company. However, the characterization of the project where the measurement program is applied has to be done for each project.

The measurement program itself is considered a project. So, during this phase the measurement project is planned (schedule, budget, milestones, etc.) and evaluated by high-level management to be accepted.

For small companies: All activities from this phase and the two first activities from the second phase (goal definition and formalization) can be realized in one meeting with all stakeholders in order to reduce the respective effort [AG02].

Related Phases:

Gre02a	HKZ01	SB99
- Prestudy	- Planning	- Planning

4.2 Establish a GQM Team [SB99]

The objective of this activity is to define which persons are responsible for the measurement activities. This team can be composed by, for example, by people from a Quality Assurance department, Software Engineering Process Group, or SPI department. The team should be independent from project teams and have no personal interest in the measurement results. The purpose of this team is to guarantee the continuity of a measurement program and to support measurement activities.

For small companies: it is not necessary to establish a GQM team. One person in the project/company can be responsible part-time for measurement activities.

The GQM team has detailed knowledge of the GQM process and is responsible for the planning and performance of the measurement program. Most of the effort required for measurement is spent by the GQM team. In comparison, the project team is responsible for the software development in the actual project and only participates in some steps of the GQM process [GRR96, GHW95].

Related activities:

Gre02a	HKZ01	SB99
		- Establish GQM team

Output Artifacts:

GQM team established/responsible people defined

4.3 Characterize the Organization

The objective of this activity is to characterize the organization, for example, regarding its business sector, the current organization chart and potential customers. The specific improvement goals for the organization are also stated.

This activity does not need to be executed every time a measurement program is started. Once characterized and having well-defined improvements goals, it is only necessary to review this characterization and the improvement goals in order to be sure that the measurement program is following the present organization's viewpoint.

Related activities:

Gre02a	HKZ01	SB99
- Prestudy	- Planning	- Select improvement areas

Tasks to be Executed:

- Characterize the organization in which the measurement program takes place (e.g., wrt. application domain and business sector) [Gre02a].
- The organizational and project characterization can be supported by questionnaires. Note that these questionnaires should be adapted to the specific organization. Examples of questionnaires are given in Appendix I.
- State the existing problems (e.g., high defect rate) [Gre02a].
- Identify organizational improvement goals (e.g., to reduce the number of the defects in the software system).

Output Artifacts:

- Organization Characterization
- Organization improvement goal(s)

4.4 Select and Characterize the Project

The objective of this activity is to characterize the project in which the measurement program is going to be executed. This may include information as, for example, about the size of the project and staff. The specific project goals are stated, too.

This activity should be always executed in order to have a well-defined context in which measurement is applied.

Related Activities:

Gre02a	HKZ01	SB99
- Prestudy	- Planning	- Select application project and establish project team - Create project plan

T1.1.1: Characterize potential pilot projects (e.g., wrt. duration, application domain and critical quality factors) [Gre02a].

G1.1.1.1: The organizational and project characterization can be supported by questionnaires. Note that these questionnaires should be adapted to the specific organization. Examples of questionnaires are given in Appendix I.

G1.1.1.2: For small companies: If there is only one project (or only one typical project in the organization), the tasks from this activity should be grouped, because it is not necessary to select a project.

T1.1.2: Select a pilot project.

G1.1.2.1: When introducing GQM based measurement, the pilot project should be selected according to some criterions:

- The project should be a typical project in the organization.
- It should be possible to motivate the project team for the application of a new technology.
- The project should need improvement.
- Consider the resources dedicated to measurement. The project duration and the team size should be reasonably small.
- The project should not have many risks.

T1.1.3: Align measurement objectives and improvement ideas of project team [SB99].

T1.1.4: Review or produce software process model [SB99].

G1.1.4.1: The existence of a software process model is essential to plan and execute a measurement program [BDT96]. If there is no model that reflects the software process for the actual project, a descriptive process model should be developed during this activity.

G1.1.4.2: The process is necessary to e.g., by defining categories (e.g., model on effort distribution per phase – the phases should be known) or later on defining the data collection procedures, the when and by whom is defined through this process model.

Output Artifacts:

- Characterization of the selected project in which measurement is applied
- Software development process model

A1.2: Plan the Measurement Program

The measurement program itself is considered a project. During this activity, this project is planned regarding to schedule definition, resource allocation and preconditions to execute, among other things. Finally, it is advised to follow this plan during the execution of the project.

4.4.1.1.1 Related Activities:

Gre02a	HKZ01	SB99
		- Create project plan

Tasks to be Executed:

T1.2.1: Define the schedule to be followed during the execution of the measurement program. This includes the definition of deliverables, resource allocation, effort (from GQM and project team) and budget to be spent, and all management decisions to hold the project.

T1.2.2: Define preconditions (e.g., software development process has to be modeled) and constraints (e.g., measurement effort < 1 man-year) to the measurement program [Gre02a].

T1.2.3: Gather all information in a Measurement management plan. This document should involve [SB99]:

- Management abstract, which presents the measurement program in approximately 20 lines.
- Introduction, which presents the scope of the measurement program, and the relation of the improvement objectives to the software development project goals.
- Characterization, which describes the outcomes of the characterizations that were held within the program on the organizational, project, project team and GQM team levels.
- Schedule, which presents the managerial issues for the measurement program, such as timeline, list of deliverables, resource allocation and cost-benefit analysis.
- Organization, which describes the relevant organizational, project team and GQM team structures of the measurement program.
- Management process, which presents priorities, management reporting procedures and risk control activities.
- Training and promotion, which presents the activities planned for training the project team and for promotion and dissemination of the results over the organization.

G1.2.3.1: This Measurement management plan will primarily serve as a proposal to management in order to receive management approval of and commitment to the measurement program. Once the proposal is accepted, the GQM team maintains the Measurement management plan. During the execution, the measurement management plan serves as a basis for monitoring and controlling the measurement program. If necessary, the measurement management plan has to be updated appropriately during the planning and execution of the measurement program.

Output Artifacts:

- Measurement management plan

A1.3: Train and Motivate All Involved People

The objective of this activity is to prepare all involved people for the planning and execution of the measurement program. During this activity all people are motivated and trained according to their specific roles in the program.

4.4.1.1.2 Related Activities:

Gre02a	HKZ01	SB99
- Prestudy	- Planning	- Training and promotion

Tasks to be Executed:

T1.3.1: Motivate all people involved in the project, including high-level management, project management, developers and all interested people.

G1.3.1.1: It is very important that people participating in the measurement program are motivated and also agree with the introduction of a GQM-based measurement program. Therefore, these people have to be attracted and convinced of the GQM approach. This can be done (e.g. in the context of a seminar on measurement) by presenting the benefits of a measurement program, how their work is supported, which problems could be solved and - if available - telling success stories from other projects [GRR96, GHW95].

G1.3.1.2: Motivating is constantly necessary. It is important that involved people know the importance of applying the measurement program, so that they will be more compliant with measurement, providing good data and giving their feedback during interpretation sessions.

G1.3.1.3: For small companies: the activities 10 and 11 can be grouped. In small companies, it is not interesting to have many meetings in order to reduce the costs with measurement.

T1.3.2: Train all participants of the measurement program according to their specific roles.

G1.3.2.1: The training should cover the GQM paradigm (for a basic understanding) and an overview on the GQM method, focusing the process steps in which they will actively participate [GRR96, GHW95].

Output Artifacts:

- Motivated and trained people

GQM 2: Definition

This phase builds a base for the effective execution of the measurement program. During this phase, relevant metrics are derived and procedures (methods) and instruments (tools) for their collection are defined.

Related Phases:

Gre02a	HKZ01	SB99
- Identification of GQM goal(s) - Development of GQM plan(s)	- Planning	- Definition
- Development of Data collection plan	- Instrumentation - Data collection and storage	- Definition - Data collection

A2.1: Define Measurement Goals

The objective of this activity is to define the goal(s) to be achieved with the measurement program considering the knowledge acquired wrt. organization and project characterization and objectives. Some preconditions to start with this activity are:

- The organization should be characterized and its goals defined.
- The project in which the measurement program will be applied should be selected, characterized and its goals defined.
- All people should be trained according to their roles in the measurement program.

If one of these points is not well defined than the introduction activity should be executed first, even if partially.

For small companies: the next two activities – define measurement goals and goal formalization – can be realized in the same meeting.

4.4.1.1.3 Related Activities:

Gre02a	HKZ01	SB99
- Identification of GQM goal(s)	- Define measurement goals	- Define measurement goals

Tasks to be Executed:

T2.1.1: Define measurement goals. A template, which points out the relevant dimensions of a GQM goal, is structured as follows [BCR94]:

Dimension	Definition	List of Examples	Example
Object	What will be analyzed?	processes, products, resources	Software development process
Purpose	Why will the object be analyzed?	characterization, evaluation, monitoring, prediction, control, improvement [BDR96]	Characterization
Quality Focus	What property of the object will be analyzed?	cost, correctness, defect removal, changes, reliability, userfriendliness, maintainability	Reliability
Viewpoint	Who will use the data collected?	user, senior manager, project manager, developer, system tester, quality assurance manager	Software developer
Context	In which environment does the analysis take place?	Organization, project, problem, processes, etc.	Company IntelliCar/ ABS

G2.1.1.1: When defining goal dimensions the aspects listed below should be followed:

- *Object*: it should focus on products or processes that need to be better understood. The process model can help to identify and define relevant objects.

- *Purpose:* it should be adapted to the comprehension level of problems in the company (e.g.: before starting to modify it is necessary to characterize the current situation). The objective should be adapted to the process stability and performance controlling, too.

When introducing GQM the purpose of a GQM program has to be selected with extreme care [GRR96]. Possible purposes are:

- Understanding, characterisation: focusing on a basic understanding and the setting of baselines.

- Plan, control, monitor: when baselines have been set, measurement can support the planning of a software project, control of the ongoing project and monitoring current trends.

- Improve, evaluate: when baselines have been set, improvements can be focused and modifications due to improvement actions can be assessed.

- *Quality focus:* it should be consistent with the organizational improvement goal priority and focus on more urgent weaknesses related to software process.

- *Viewpoint:* If GQM is being used for the first time it is interesting to define GQM goals in three different viewpoints [GRR96, GHW95]:

- Technical people (e.g.: developers)

- Project management

- High-level management

In this way all concerning people are involved and get the measurement program benefits. In the other side, the measurement program will be increasingly accepted and everyone will cooperate with it.

G2.1.1.2: Potential measurement goals have to be identified with great care by all stakeholders of the measurement program. This is important, as measurement will only be accepted, if the interests of all people involved (e.g., management representatives, developer) are reflected and considered by the measurement program [Epg02].

G2.1.1.3: In order to define the measurement goals formally, try to identify potential goals first informally in a brainstorming session. Use the presented template in the next step to formalize those informally formulated goals [GRR96, GHW95].

G2.1.1.4: A mechanism to support goal definition and selection in a meeting, is by asking the "seven questions" stated below [SB99]:

- What are the strategic goals of your organization?

- What forces have an impact on your strategic goals?

- How can you improve your performance?

- What are your major concerns (problems)?

- What are your improvement goals?

- How can you reach your improvement goals?

- What are possible measurement goals, and what are their priorities?

T2.1.2: Prioritize the GQM goals according to organization and pilot project needs [GHW95].

T2.1.3: Select the more important goals.

G2.1.3.1: When starting with measurement, it is important to define only few goals (usually it is advised only four goals) in order to keep the costs low and to evaluate the benefits of the measurement before amplifying the program throughout the organization [GHW95].

Output Artifacts:

- GQM goal(s)

A2.2: Formalize Goals

Once the measurement goals are defined, a GQM plan is developed for each GQM goal. Therefore, each GQM goal is formalized. The objective of this activity is to formalize the quality focus into quality factors and variation factors, and the object of study into its specific process/product model.

4.4.1.1.4 Related Activities:

Gre02a	HKZ01	SB99
- Goal formalization	- Identify QF + VF - Fill abstraction sheets	- Conduct GQM interview - Review or produce software process models

Tasks to be Executed:

T2.2.1: Perform interviews (or workshops) with several representatives of the viewpoint to refine the quality focus defined in the goal into quality factors and variation factors. Use abstraction sheets (0) to summarize the main issues and goal dependence. This information is presented in five sections on a page [GRR94]:

- Quality Factors: What does the quality focus specified in the selected GQM goal mean to the interviewee? What are the measured properties of the object, according to the project members?
- Baseline Hypothesis: What is the project member's current knowledge, with respect to these quality aspects (measured aspects)?
- Variation Factors: Which (environmental) factors do the project members expect to have an effect on the measured properties?
- Impact on Baseline Hypothesis: How do these variation factors influence the measured properties?
- Feedbak: indicates improvement suggestions in the relation to the GQM goal.

Object	Purpose	Quality Focus	Viewpoint	Context
Sw development process	Characterization	Reliability	Sw developer	Company x

Quality factors	Variation factors
QF1: total number of defects QF2: distribution of defects per criticality QF3: % of defects detected in inspections QF4: total rework effort (in hours)…	Process conformance: VF1: type of coding inspection Domain understanding: VF2: experience of development team …
Baseline hypotheses	**Impact on baseline hypotheses**
QF1: total # of defects: 100 QF2: 75% uncritical, 25% critical QF3: 20% QF4: total rework effort: 1000 h …	VF1: ad-hoc code inspection are less effective than other types VF2: more experienced development team introduces less defects…
Feedback All developers should be trained wrt. the software development process.	

Figure 6 An Example of Abstraction Sheet [Gre02a]

G2.2.1.1: Interviews with various people representing the viewpoint stated in the GQM goal should be done in order to capture all relevant information.

G2.2.1.2: During these interviews it is possible to discover that the GQM goals selected are not adequate to the objects/purposes to be measured. In this case, the GQM goals should first be redefined [GRR96].

G2.2.1.3: The interviews should be performed in a manner, that all different conceptions of processes performed or terms definition from people (addressed by the viewpoint clause) can be captured, e.g., by single interviews. If this is impossible, interviews can take place in small groups depending on the dynamics of the relevant group of people, wrt the possibility of open and controversial discussions [GRR96, GHW95].

G2.2.1.4: The interviewer should regard to some issues that can manage the interview to be well succeeded:

- The interviewees should be motivated, trained and informed about the measurement program and which information is expected from them during the interviews [GRR96, GHW95].

- In the beginning confidence should be assured and how the information will be used should be explained.

- The interviewer need to be well prepared about the specific goal, the software process and the terminology used in the pilot project. Experiences with similar measurement programs can help in the preparation of the interview.

- The interviewer task is to acquire knowledge, s/he should not influence the interviewee in any direction [GRR96].

G2.2.1.5: Organizational models, e.g. the software process model can be used like a reference in the interview, if this is considered useful by the interviewee [GRR96].

T2.2.2: According to the defined object of study in the GQM goal, define or review its respective process/product model.

G2.2.2.1: If there is no process or product model with repsect to the defined object, a model should be defined.

G2.2.2.2: There should be an agreement on the models by all people involved in the measurement program. Make sure the process models describe in which way the work is really done and not the ideal way in which it should be done.

Output Artifacts:

- Abstraction sheet
- Process/product model refining the object of study

A2.3: Define Questions

After defined and formalized the goal, questions are derived in order to help to define which information is required in order to achieve the goal. These questions are used in order to derive relevant metrics.

4.4.1.1.5 Related Activities:

Gre02a	HKZ01	SB99
- Development of GQM questions	- Define questions	- Define questions and hypothesis - Review questions and hypothesis

Tasks to be Executed:

T2.3.1: For each quality factor/variation factor defined in the abstraction sheet at least one question will be derived/specified.

G2.3.1.1: Since GQM plans usually consist of large numbers of questions, questions categories [BDR96, BR88] have been defined in order to guide their derivation and to structure GQM plans.

G2.3.1.2: Questions can be further refined, structured in a question hierarchy [GRR96].

G2.3.1.3: Questions regarding variation factors should explicitly express the hypothesized relationship between the variation factor and the quality factor(s) as motivation for the question.

G2.3.1.4: To support an optimal interpretation from collected data to answering questions to achieve the goal, the questions should be defined at an intermediate level of abstraction between metrics and goals.

G2.3.1.5: The description of the organizational environment (including, e.g., software process model) can provide information required to formulate questions and metrics (terminology, definition of models, etc.) [GRR96].

G2.3.1.6: If necessary, comments should be given for each question regarding to specific information about terminology, in which circumstances that question is of interest, etc.

T2.3.2: For each question formulate potential answers as hypotheses.

Output Artifacts:

- GQM questions

A2.4: Define Models

The objective of this activity is to define models that will operationalise the GQM questions, quantify the various attributes of the object being studied and define precisely how comparisons, evaluations, and predictions are to be performed [BDR96].

If the GQM team (or people responsible for the measurement program planning) are experienced in measurement, or if the measurement program is very simple, this activity does not need to be executed.

4.4.1.1.6 Related Activities:

Gre02a	HKZ01	SB99
- Development of GQM models		

Tasks to be Executed:

T2.4.1: Define the model detailing each question.

G2.4.1.1: A GQM model can include the following dimensions [Gre02a]:

- Model description: describe the objective and usage of the model.
- Assumptions: specify assumptions on which the model is based.
- Terminology: defines terminology used in the formalization of the model.
- Context-dependent properties: describing the intuitive understanding of the ordering relation with respect to the quality factors in the empirical world.
- Analysis: describes how the data collected wrt. the attributes to be measured will be analyzed in order to feed the model.
- Entities/attributes: lists the entities/attributes to be measured and defines the properties of the attributes based on our intuitive understanding.

G2.4.1.2: An example of a descriptive model is [Gre02a]:

Descriptive Model: *Effectiveness of inspections*
The model describes the effectiveness of inspections at company X, project ABS.
Terminology: Defect is a synonym for the term fault, which is an incorrect step, process or data definition in a computer program. Document is term subsuming any type of document produced during the software development process.
*Assumption(s):*The actual defect density is comparable across software modules.
Context-dependent properties: if during an inspection i1 there are at least as many defects detected per size of software module as in another inspection i2, then inspectionEffectiveness (i1) >= inspectionEffectiveness (i2).
Analysis: distribution of effectiveness across inspections performed
Effectiveness = (number of defects detected during inspections)/size of software modules

Entity/Attribute(s):
- Inspection: number of defects detected
property1. The number of defects detected is non-negative
property2. The number of defects detected is null if none detected
- Software module: size (-> descriptive model on software module size)

G2.4.1.3: Attributes are formalised via a set of generic properties that characterise their metrics (e.g., mathematical properties such as additivity) [BEM95].

G2.4.1.4: These models can be of different types (e.g., descriptive or predictive) depending on the purpose of the GQM goal or the objective of the respective GQM question [BDR96].

G2.4.1.5: Models have to be defined precisely and unambiguously. That means, that terms used in the model definition have to be explicitly specified in consistency with respect to the organizational terminology.

Output Artifacts:
- GQM models

A2.5: Define Indicators

Define visualization of metrics (indicators) in order to support answering questions, this means that example diagrams are drawn. This is a starting point for the development of feedback material and can be used to motivate the project team with respect to commitment to the measurement program.

4.4.1.1.7 Related Activities:

Gre02a	HKZ01	SB99
	- Define indicators	- Produce analysis plan

Tasks to be Executed:

T2.5.1: Identify the specific measurement data items that will be displayed [Epg02].

G2.5.1.1: The baseline hypotheses of the project team members are an important start point, which can be compared with the actual data.

T2.5.2: Define the format for presentation of data items [SB99, Epg02].

G2.5.2.1: An indicator can be designed in many ways, e.g.: tables or graphics.

G2.5.2.2: Graphical representation is usually the most useful format for issue-driven metrics, since the most useful project information is generally obtained by analyzing relationships between metrics.

G2.5.2.3: Indicators illustrate to the project team which charts they should expect during data interpretation.

Output Artifacts:

- Indicators

A2.6: Define Metrics

Based on the questions, models and indicators, metrics are defined. The objective of this activity is to determine relevant metrics to be collected and to define their main properties.

In some documents [Gre02a, Dif01, PGF96] the term "measure" is used rather than "metric", as metric connotes a generic distance measure in the mathematical and physical sciences. Here, we use the term "metric" also due to the name of the approach: "Goal/Question/**Metric**".

4.4.1.1.8 Related Activities:

Gre02a	HKZ01	SB99
- Development of GQM measures	- Define measures	- Define measures

Tasks to be Executed:

T2.6.1: Refine the questions/models into metrics. A metric can be defined in terms of [Gre02a]:

- Unit: specifies how the attribute is measured.
- Scale: helps to select adequate methods for the data analysis and indicates permissible transformations.
- Range: indicates a set of permissible values and may help to identify abnormal values.
- Protocol: describes the way in which the metric is to be collected.

G2.6.1.1: The metric properties (scale, unit…) should be explicitly defined in order to support the collection of correct data during the execution of the measurement program.

G2.6.1.2: When defining the metric properties (scale, unit…), the future data collectors should be involved. They are the experts and know which are the correct ranges/units/… to be used [FKLR00].

G2.6.1.3: If the company has not done any measurement activities before, it is important to start with a small number of metrics rather than trying to measure everything [HDK00]. Then try to measure them first, analyze the measurement and then define other metrics that have to be specified according to the measurement goal.

G2.6.1.4: Metrics should be defined to provide all the quantitative and qualitative information for answering the questions in a satisfactory manner

G2.6.1.5: If models are defined (according to [Gre02a]), for each entity/attribute defined in a GQM model, a metric is defined, which maps an empirical attribute to the formal model.

G2.6.1.6: Once all these measurements are collected with respect to the defined metrics, sufficient information should be available to answer the questions completely [SB99].

G2.6.1.7: Metrics can be classified in [FP97], for example:

- Objective vs. subjective measurement:

- Objective measurements: They are environment independent and permit that different people reproduce the same measure. For example, the effort spent in an activity. These measurements are more precisely than subjective ones.

- Subjective measurements: These measurements are environment dependent. The measures can vary depending on the data collector's judgement. For example, what is considered good for one can be considered bad for another (e.g., when measuring user satisfaction).

- External vs. internal measurement:

  - External measurements: These measurements are related to the environment of the object, considering the its behavior more important then the entity itself. For example, reliability, usability or cost-effectiveness.

  - Internal measurements: These measurements are those that can be measured exclusively in terms of the object itself. It can be measured by examining the entity separate from its behavior. For example, size or complexity.

- Direct vs. indirect measurement:

  - Direct measurements: Measurement of an attribute of an entity that involves no other attribute or entity. For example, code length (LOC) or effort spent (person-month).

  - Indirect measurements: Measurement of an attribute of an entity that depends on other attribute or entity to be measured. For example, to determine the programmers productivity it is necessary to know: LOC produced / effort spent (person-month).

Output Artifacts:

- GQM metrics

A2.7: Produce and Review GQM Plan

The objective of this activity is to gather all information defined until now for each goal into its specific GQM plan. This plan is the basis for the execution of the measurement program and guides the analysis and interpretation sessions when links between goal/questions/models/indicators/metrics/... are discussed.

4.4.1.1.9 Related Activities:

Gre02a	HKZ01	SB99
- Review of GQM plan	- GQM plan integrate	- Produce GQM plan

Tasks to be Executed:

T2.7.1: For each defined goal, group all the information defined until now (quality factors and variation factors, questions, hypothesis, models, indicators, metrics…) into the specific GQM plan.

G2.7.1.1: The GQM plan has to be refined iteratively. During the first iteration step the GQM plan usually cannot be completely developed up to the level of metrics and the definition of the indicators. The process (and eventually the preceding interview process) will be iterated as long as the GQM plan does not pass the subsequent review [GRR96].

G2.7.1.2: For each measurement goal defined a specific GQM plan should be produced.

G2.7.1.3: It is interesting to define a glossary of used terms. All terms used should be well understood and agreed by all people involved in the measurement program.

G2.7.1.4: In order to control the main measurement activities, it is interesting to explicitly report information about updates in the Measurement plan (GQM plan and Data collection plan), e.g., in form of a table, including date, which updates were made and by whom.

T2.7.2: Review the plan to ensure that defined goals, questions, models, indicators and metrics are consistent and complete in relation to the process and product models of the respective project.

G2.7.2.1: Consistency and completeness checks have to be performed throughout the entire definition phase. According to the size of this plan more reviews must be kept in order to provide a more consistent plan.

G2.7.2.2: The GQM plan is approved, if no contradictions are detected. Any problem detected has to be resolved. This may require a (partial) repetition of the previous steps until the GQM plan is approved [GHW95].

G2.7.2.3: For small companies: If the defined metrics are simple, the review of the GQM plan by the project team can be done later together with the review of the Data collection plan.

Output Artifacts:

- GQM plan

A2.8: Define Data Collection Procedures

The objective of this activity is to determine how the defined metrics in the GQM plan should be collected, by whom and when.

4.4.1.1.10 Related Activities:

Gre02a	HKZ01	SB99
- Definition of data collection procedures	- Create Measurement plan	- Produce Measurement plan

Tasks to be Executed:

T2.8.1: Group all derived metrics into a list.

G2.8.1.1: Repeated occurrences of metrics wrt. different models are mapped into one unique data collection procedure [GHW95].

T2.8.2: For each metric, define the procedure to be followed during its collection. This includes:

- Define who is the responsible for data collection.

G2.8.2.1: If data can be collected automatically through a tool, it is not necessary to define a role/person who will provide this data.

G2.8.2.2: The selection of data collectors should be done according to the following criteria [BDR96]:

- Expertise: Who has the technical/managerial expertise to provide the data accurately?
- Bias: Is there any reason for the data provider to show any bias in the information s/he provides?
- Access: Who has access to the object being measured?
- Cost: Can the time spent on measurement have costly effects on the project?
- Availability: Is the person available to spend time on data collection?
- Motivation: How committed is the person to the measurement program?

- Define who is responsible for the quality assurance of the collected data and for the handling and storage of the data (someone from the GQM team or a person responsible for measurement).
- Define when data should be collected. For example, how often developers are asked to fill in a survey-questionnaire (daily, monthly).
- Define how data should be collected. This includes the definition of instruments (tools, forms, interviews, etc) necessary for data collection.

G2.8.2.3: Existing forms, such as problem reports or tools available, should be considered by integrating the data collection procedures, if possible [Gre02a].

G2.8.2.4: In order to keep the data collection overhead minimal, measures that can be collected at the same time by the same role and the same instrument should be grouped together to be collected by the same instrument [GHW95].

Output Artifacts:

- Data collection procedures

A2.9: Develop Data Collection Instruments

During this activity the instruments defined in the data collection procedures are developed. Depending on the type of data collection instrument, either a tool has to be developed, forms have to be designed or structured interviews have to be planned. The data collection instruments have to be tailored to the development environment, which requires detailed knowledge on the organisational structure, workflow, information flow, standards and terminology.

Each task defined to this activity is related to an instrument. According to the instruments defined in the data collection procedures specific tasks are realised.

4.4.1.1.11 Related Activities:

Gre02a	HKZ01	SB99
- Definition of data collection instruments	- Prepare data collection - Tailor data collection tool	- Produce Measurement plan

Tasks to be Executed:

T2.9.1: If manual forms are needed, they have to be carefully designed. Manual form consists of a paper where the information is manually filled in and afterwards electronically stored.

G2.9.1.1: A form, in general, contains some identification information. For all measures to be collected on the form, a question is formulated. If applicable, the expected answers should be explained by stating unit and range.

G2.9.1.2: In order to enable the collection of reliable data, the fields have to be carefully formulated, all unknown terms explained, and only short answers should be requested and checklists should be used, when possible [GHW95].

G2.9.1.3: Data collection forms should be designed so that the data can be easily provided [GRR96, GHW95]:

- The data to be provided on collection forms should be concise and clearly defined.

- Organisational standards and terms should be used on the collection forms to prevent misunderstandings.

- Used classification categories on the collection forms have to be explained.

G2.9.1.4: If data is collected by GQM team, a much simpler version of a form can be designed without detailed explanations, as the fields and answers categories should be well known by the GQM team. An example of such form is the Data Extraction Sheet ("*Extrato de Dados*") as defined in [AG01].

T2.9.2: If electronic forms are needed, they need to be developed [SB99]. Electronic forms are an improvement from manual forms. The advantage is that the collected data does not need to be electronically stored after being collected as it has to be done with data collected on manual forms. But, the data collection effort continues the same for project team, because they have to fill it in these forms.

G2.9.2.1: Some possibilities for electronic data collection forms are e-mail, web pages and database applications.

G2.9.2.2: A form, in general, contains some identification information. For all metrics to be collected on the form, a question is formulated. If applicable, the expected answers should be explained by stating unit and range. This is especially interesting for data collected via electronic forms, because it may allow partial automatization of data validation, depending on the tool used to report the data [GRR96, GHW95].

T2.9.3: If interviews are needed, they need to be planned.

G2.9.3.1: An interview plan describes the main steps of the interview and lists the data to be acquired during the interview [Gre02a].

T2.9.4: If automated data collection tools are needed, select and prepare the existing ones, or if necessary, develop them [SB99]. Automated collection tools permit data collection without having to fill in forms. They can reduce the data collection time, but do not limit data collection to automated tools. The most valuable information generally comes from people and not from tools.

Output Artifacts:

- Data collection instruments

A2.10: Produce and Review Data Collection Plan

The objective of this activity is to gather information about data collection procedures and instruments in a document named Data collection plan. This plan will document the process used for data collection.

4.4.1.1.12 Related Activities:

Gre02a	HKZ01	SB99
- Review of Data collection plan	- Train collection procedures	- Trial period

Tasks to be Executed:

T2.10.1: Group the data collection procedures and instruments in a document called Data collection plan.

G2.10.1.1: This term "Data collection plan" was suggested by [Gre02a], substituting the term "Measurement plan" in order to differentiate it from the complete plan of the measurement program including the GQM plan and the Data collection plan.

T2.10.2: Review the Data collection plan regarding its completeness and consistency with the GQM plan(s) and the Measurement plan as a whole.

G2.10.2.1: During the review the following aspects should be considered [GHW95]:

- Is the Data collection plan consistent with the GQM plan(s)?
- Are all metrics from the GQM plan(s) considered in the Data collection plan?
- Are all data collection procedures and responsibilities for data collection, validation and storage specified?
- Are the data collection points of time part of the project plan?
- Do the data collection instruments defined exist in the project plan or have been developed?
- Are the assumptions that are documented in GQM plan(s) consistent with the project plan?

The Data collection plan is approved, if no contradictions are detected. Any incompleteness or inconsistency detected has to be resolved.

T2.10.3: Train the data providers with respect to the data collection procedures, in order to yield valid, complete, and consistent data.

G2.10.3.1: Try to have a trial measurement period during which the defined data collection procedures, tools and forms can be tested [SB99].

Output Artifacts:

- Data collection plan

A2.11: Hold a kick-off session

The objective of this activity is to get the agreement of all people involved in the measurement program to start the execution of the program.

For small companies: in general, few people are involved in the measurement program in this context. It is easier to get the agreement of people by talking to them informally and often a formal meeting is unnecessary.

4.4.1.1.13 Related Activities:

Gre02a	HKZ01	SB99
		- Hold a kick-off session

Tasks to be Executed:

T2.11.1: Hold a kick-off session to get an agreement with the project team that data collection will indeed start [SB99].

G2.11.1.1: During this session all people involved in measurement program should be present.

G2.11.1.2: During this session, once again the goals of the measurement program are explained, and the project team members are instructed with respect to data collection procedures, tools and forms. At the end of the session, everybody should agree that data collection can actually start.

Output Artifacts:

- Agreement for data collection by project personnel
- Data collectors trained for data collection

GQM 3: Data Collection

In this phase, the execution of the planned measurement program starts. During this phase, a base to store collected data is developed. The data is collected, validated and stored.

Gre02a	HKZ01	SB99
- Data collection, validation and storage	- Data collection and storage	- Data collection

A3.1: Create Measurement Base

4.4.1.1.14

The objective of this activity is to develop a measurement base where the collected data will be stored and processed.

4.4.1.1.15 Related Activities:

Gre02a	HKZ01	SB99
	- Prepare database	- Create metrics base

Tasks to be Executed:

T3.1.1: Define a measurement base: a database where the collected data will be stored and processed [SB99].

G3.1.1.1: Tools to support this base are:

- Spreadsheet tools, where data can be stored, maintained and processed. The use of spreadsheets is interesting because of their flexibility to support goals adjusting, adding or removing.

- Database management systems (DBMS), where data can be organized, stored and retrieved. The advantages of using a DBMS are the availability of languages to define the data structure, insert, modify and delete data, and extract refined data. Constraints such as checks on cross-references among records can be defined to ensure consistency of data. Formats, ranges, valid values, and more can be checked automatically as they are input [FP97].

G3.1.1.2: This base can be created earlier in the process, during the planning phase, although it depends on the GQM plan.

Output Artifacts:

- Measurement base

A3.2: Collect, Validate and Store Data

4.4.1.1.16

The objective of this activity is to collect data using the defined instruments (forms, tools, etc.) in accordance to the data collection procedures. This data is delivered to the GQM team (or the responsible for measurement activities) and then validated. Finally, this data is stored in the measurement base.

4.4.1.1.17 Related Activities:

Gre02a	HKZ01	SB99
- Data collection, validation and storage	- Provide data - Validate data - Store paper questionnaires	- Collect and check data collection form - Store measurement data in metrics base

Tasks to be Executed:

T3.2.1: Provide data according to the defined measurement procedures.

G3.2.1.1: Providing data can mean to fill out a questionnaire, enter data into a collection tool, run an automated tool or to perform interviews.

G3.2.1.2: This activity is realized by the responsible for data collection as defined in the data collection procedures.

G3.2.1.3: It is interesting to continuously support and motivate the project team to collect data, especially in the initial phase of the measurement program, by stating the benefits of measurement to them [FKLR00].

T3.2.2: Validate the collected measurement data.

G3.2.2.1: At least the following points should be checked [GRR96]:

- *Legibility*: Are the answers legible? This question is only relevant for collection sheets filled out by hand.

- *Completeness*: Have all required collection forms submitted? Are there any questions omitted that have to be answered?

- *Correctness*: E.g. several answers checked in a multiple choice question, although only one answer is allowed.

- *Plausibility*: Are the values of correct type, e.g. integer, decimal, text? Do the values lie within a valid range?

G3.2.2.2: The validation should be done as soon as possible after the data collection in order to enable the correction of invalid data [Gre02a].

G3.2.2.3: The validation may be at least partially automated for data that is collected on-line [GRR96].

T3.2.3: Store data that was captured on paper questionnaires electronically.

G3.2.3.1: Data that was not collected electronically (electronic forms or automated tools) must be electronically stored in the measures basis.

G3.2.3.2: Storing the collected and validated data from the collection sheets into the measurement database should be done with great care, since this process is a source for introducing errors. This process may be superfluous or (at least partially) automated for data that is collected on-line [GRR96].

Output Artifacts:

- Validated data stored

GQM 4: Data Analysis and Interpretation

During this phase the collected (and stored) data is processed, analyzed, organized in presentation slides and presented for interpretation during feedback sessions.

Gre02a	HKZ01	SB99
- Data analysis and interpretation	- Data analysis and report generation	- Interpretation

A4.1: Analyze Collected Data

4.4.1.1.18

The objective of this activity is to effectively process the stored data, organize them in graphs/tables and prepare the presentation slides based on the GQM plan.

4.4.1.1.19 Related Activities:

Gre02a	HKZ01	SB99
- Data analysis	- Retrieve data - Analyze data	- Define analysis sheets and presentation slides

Tasks to be Executed:

T4.1.1: Define the analysis sheet [SB99]. The analysis sheet is composed by three layers:

- *Raw data*: contains raw data copied from the measurement base.
- *Processed data*: contains processed data, it means that relevant data must have been selected, calculated and sorted (e.g., some measures are the average of other measures). This layer is a basis for creating graphs and tables.
- *Graphs and tables*: the result from processed data is transformed into graphs and tables for presentation later on.

G4.1.1.1: These graphs and tables must be designed according to the defined indicators.

T4.1.2: Define analysis slides. This is a presentation tool where graphs and tables resulting from the analysis sheet will be prepared for presentation during the data interpretation activity.

G4.1.2.1: The data presentations should indicate the GQM question they intend to answer, the applied analysis method and the number of underlying data points. In addition, in order to stimulate discussions during the interpretation, the collected data can be compared with the hypotheses stated in the GQM plan.

Output Artifacts:

- Analyzed data
- Presentation slides

A4.2: Interpret the Analyzed Data – Feedback Sessions

4.4.1.1.20

The objective of this activity is to interpret the collected data, to verify the hypotheses stated in the GQM plan, to identify improvement possibilities and to evaluate the measurement program.

4.4.1.1.21 Related Activities:

Gre02a	HKZ01	SB99
- Interpretation	- Prepare feedback session - Hold feedback session	- Prepare feedback session - Organize and hold feedback session - Report measures resulting

Tasks to be Executed:

T4.2.1: Prepare a feedback session in order to interpret analysis results.

G4.2.1.1: Typical actions are: reserve room, invite people, select results to be discussed, review presentation slides and choose presentation methods.

G4.2.1.2: The presentation material should be handed out to the participants at least one half to one day before the feedback session takes place. If the presentation material does not get ready in time, it is recommended to postpone the feedback session instead of skipping the self study and performing the feedback session with unprepared people [GRR96, GHW95].

G4.2.1.3: Based on requests resulting from the self-study of the presentation material in advance or already performed feedback sessions new demands wrt. the presented material can arise by the participants. The presentation material should be modified accordingly [GRR96].

G4.2.1.4: It is advised to present 15 to 20 slides at most in a single feedback session [SB99].

G4.2.1.5: Feedback sessions should be held periodically (e.g., every 4-8 weeks) depending on the availability of new data and the respective GQM goal [GHW95].

T4.2.2: Interpret analyzed data during the feedback session.

G4.2.2.1: Interpretation includes validation with respect to the environment, identification of problems, root cause analysis and suggestions for improvements.

G4.2.2.2: It can be useful to have on-line access to the measurement database during the feedback session. Thus, some requests for additional information may be satisfied immediately by on-line queries to the database [GRR96].

G4.2.2.3: Analysis of data could also include supplementary validation of the data. For instance, a data provider can remark that some data points are only estimates or some data points are missing due to illness, vacation etc. Such information has to be taken into account when interpreting the measurement data [GRR96].

G4.2.2.4: Three types of roles can be identified in a feedback session [GRR96, GHW95]:

- The *moderator* presents the slides and leads the discussions.

- The *people representing the viewpoint of the GQM goal* analyze the slides and draw conclusions.

- The *members of the GQM team* estimulate the discussions, address interesting observations and record notes during the feedback session.

G4.2.2.5: During the feedback session all representatives of the viewpoint in the GQM goal and the data collectors should be present.

G4.2.2.6: Mutual trust among all participants is an essential element of a feedback session [SB99].

G4.2.2.7: Make sure that measurement analysis does not become focused on individual performances [SB99].

G4.2.2.8: It is advised for a feedback session not to take longer than 3 hours, because it requires a high level of concentration by the attendees [GRR96, GHW95].

G4.2.2.9: The GQM team should not interpret data, this task must be done only by the project team (or people of the viewpoint in the GQM goal), because they are the experts.

T4.2.3: Evaluate the measurement program itself during the feedback session, too [GRR96]. The evaluation includes the analysis of, e.g.,:

- If any measurement data turns out to be of little or no use, it can be decided not longer to collect this data.
- If the collection procedures are too intrusive, too time-consuming, the collection procedures should be simplified.
- Different forms of data presentations may be proposed in order to facilitate the interpretation.
- A shift of focus of interest may take place during the measurement program. If more detailed measurement data is needed, the feedback sessions are the right place to discuss the expansion of the measurement program.

In each of these cases the Measurement Management plan, GQM plan and Data collection plan need to be updated respectively.

T4.2.4: Report interpretations of measurement results [SB99]. This report should contain all relevant observations, interpretations, conclusions and action points that were formulated during the feedback session.

G4.2.4.1: This report is then distributed to the project team.

G4.2.4.2: Privacy aspects wrt. the data have to be considered carefully. It is important that data privacy issues are discussed openly and any concerns of all participants are taken seriously. In order to ensure privacy only accumulated data should be presented [GRR96].

G4.2.4.3: The implementation of the modification is essential, otherwise the measurement program only implies additional effort or overhead without any benefits [Gre02a].

T4.2.5: According to the results of the feedback session, define if the measurement program should continue. Some points to be considered are:

- If people (high-level management, project management, developers…) involved in the program are not satisfied with the measurement program.

- If the established goal was attained.

If the measurement program is terminated and the next activity (Package Results) is executed. Otherwise, if changes to the Measurement plan were requested, it must be updated, so that data collection can proceed. If no changes were requested, the measurement program continues with A3.2: data collection.

Output Artifacts:

- Feedback session report

GQM 5: Packaging

This is the last phase of the measurement program once the execution of the measurement program is finished. During this phase, the measurement plan, all data collected and experiences gathered during the measurement program are analysed and stored so that they become reusable in future software projects and measurement programs [BCR94, GB97, Gre02a]. Packaging experiences includes the development and enhancement of standards, training, and development policies.

Gre02a	HKZ01	SB99
- Post-mortem analysis and capturing of experiences	- Data analysis and report generation	

A5.1: Package Results

4.4.1.1.22

Results and experiences gained by the measurement program wrt. the object of study as well as to the measurement program itself are analyzed and packaged [GRR96].

4.4.1.1.23 Related Activities:

Gre02a	HKZ01	SB99
- Post-mortem analysis and capturing of experiences	- Package results	

Tasks to be Executed:

T5.1.1: Package results of the measurement program. This may include the development or update of [SEL94]:

- *Software management policies and guidelines*: Much of the information that has been collected and synthesized by the analysis can be fed back into the organization in the form of models, planning aids, and guidelines. When packaged into well-designed policies and guidebooks, this information can improve a manager's ability to plan a software project, monitor its progress, and ensure the quality of its products.
- *Software development and maintenance standards*: Based on the analysis results, standards for software development and maintenance defining the products, methods, tools, data collection procedures, and certification criteria that have been identified as beneficial to the organization may be defined.
- *Software training*: The measurement experience can drive the planning and execution of the training curriculum of the organization.

- *Tools and automated aids:* Tools and automated aids which have been developed with the objective to facilitate software management, development, maintenance, or data collection processes.
- *Reports of process studies*: describing the goal, the methods employed, the results measured, the conclusions drawn and lessons learned wrt. project management, software process, etc. They are vital sources of information when the time comes to integrate study recommendations with other standard practices before packaging them as policies, guidebooks, courses, or tools.
- *Measurement plan*: all products developed during the measurement program, such as GQM goals, metrics etc. are potentially reusable in a similar measurement program [Gre02a].

G5.1.1.1: Each experience package must be associated with the context in which the models/lessons learned are valid and/or can be applied. They have to be tailorable to the needs, extendible, understandable and accessible for future software projects [BR88,GRR96].

G5.1.1.2: The experience package can represent various kind of experiences, such as [BR88]:

- Product packages which have as their central element a life-cycle product, associated with the information needed to reuse it and the lessons learned in reusing it (e.g., programs, architectures, designs).
- Process packages which have as their central element a life-cycle process, associated with the information needed to execute it and the lessons learned in executing it (e.g., process models, methods).
- Relationship packages which have as their central element a relationship or a system of relationships among observable characteristics of a software project. These packages are used for analysis and/or forecast of relevant phenomena (e.g., cost and defect models, resource models).
- Tool packages which have as their central element a specific tool, either constructive (e.g., code generator, configuration management tool) or analytic (e.g., static analyzer, regression tester)
- Management packages which have as their central element any container of reference information for project management (e.g., management Handbooks, Decision Support Models.
- Data packages which have as their central element a collection of defined and validated data relevant for a software project or for activities within it (e.g., project databases, quality records)

G5.1.1.3: Experience pacakages can be represented in a variety of forms [BR88]:

- Equations defining the relationship between variables, (e.g., Effort = a * $Size^b$)
- Histograms or pie charts of raw or analyzed data, (e.g., % of each class of fault)
- Graphs defining ranges of «normal» (e.g., graphs of size growth over time with confidence levels)
- Lessons learned associated with project types, phases, activities (e.g., reading by stepwise abstraction is most effective for finding interface faults), or in the form of risks or recommendations, (e.g., definition of a unit for unit test in Ada needs to be carefully defined)
- Models or algorithms specifying the processes, methods or techniques (e.g., an SADT diagram defining design inspections with the reading technique as variable dependent upon the focus and reader perspective)

- Examples of resource models and baselines including cost models, resource allocation models for staffing, schedule, and the relationship between resources and various factors that affect resources, e.g. specific methods, customer complexity, the application, the environment, and defect classes.

The purpose of the experience package is to provide useful and easy-to-use experience to appropriate software projects on demand.

G5.1.1.4: The systematic learning and packaging of the experiences has to be separated from the software development [BCR94a].

Output Artifacts:

- Experience package(s)

Document Information

Title:	GQM-Handbook and Overview of GQM-plans A comparison of three GQM-methods and an Overview of GQM-plans
Date:	April 3003
Report:	IESE-008.03/E
Status:	Final
Distribution:	Internal