UvA-DARE (Digital Academic Repository)

# Scientific grounding of lean six sigma's methodology

de Koning, H.

**Publication date**
2007

**Document Version**
Final published version

Link to publication

**Citation for published version (APA):**
de Koning, H. (2007). *Scientific grounding of lean six sigma's methodology*.

**General rights**
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)*

Download date:24 Aug 2022

# Scientific Grounding of Lean Six Sigma's Methodology

**Henk de Koning**

ISBN 978-90-6464-176-3

Omslagontwerp: Frank ter Weele

# Scientific Grounding of Lean Six Sigma's Methodology

**Academisch Proefschrift**

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel der Universiteit
op vrijdag 19 oktober 2007, te 14:00 uur

door

**Henk de Koning**
geboren te Langbroek

## Samenstelling van de promotiecommissie

| | |
|---|---|
| Promotor: | Prof.dr. R.J.M.M. Does |
| Co-promotor: | Dr. J. de Mast |
| Overige leden: | Prof.dr. C.A.J. Klaassen |
| | Prof.dr. M.R.H. Mandjes |
| | Prof.dr. ir. C.T.B. Ahaus |
| | Prof.dr. S. Bisgaard |
| | Dr. ir. A. van der Wiele |
| | Dr. ir. F.H. van der Meulen |
| | Dr. J. van den Heuvel |

Faculteit der Natuurwetenschappen, Wiskunde en Informatica.

*A scientist says a simple thing in a hard way. An artist says a hard thing in a simple way.*

*Charles Bukowski*

*It is better to be approximately right than exactly wrong.*

*John Tukey*

IBIS UvA BV
INSTITUUT VOOR BEDRIJFS- EN INDUSTRIËLE STATISTIEK

Dit proefschrift is mede mogelijk gemaakt door een financiële bijdrage van het Instituut voor Bedrijfs- en Industriële Statistiek van de Universiteit van Amsterdam (IBIS UvA)

# Preface

The research I did for the past four years resulted in the dissertation in front of you. Equally important, however, is that this research has had a tremendous impact on my professional and personal development. What I have learned in this respect is that is not so much what you do, but more important whom you work with. In this context I would like to thank a few people that all acted – in one sense or the other – as role models for me.

First of all, I would like to thank Jeroen de Mast. It is fair to say this thesis would not have been written at all without his help. Thanks to him for having this very rare combination of creativity and powerful logical thinking. More than once he provided a totally new angle to a problem, the solution of which I considered out of reach. And of course thanks to him for trying patiently to improve my writing skills. If I could only borrow some of his....

Next, I want to thank Ronald Does. He provided a model of confidence and energy and kept a positive outlook, whatever happened. He is able to see opportunities where others see none; maybe therefore he is smiling most of the time.

I owe a great deal to Frank van der Meulen. He helped me out when my morale touched rock-bottom, applied his coaching skills – well-developed, albeit in field hockey, but nevertheless – to guide me through the nasty technical stuff, and, most of all, he became a real friend.

I would like to thank all the other people who co-wrote papers with me: Soren Bisgaard, Thijs Vermaat, Jaap van den Heuvel, John Verver, and Serge Simons. They shared their enthusiasm with me, but remained critical all the same regarding the draft versions of the papers we wrote. My colleagues at IBIS UvA I am indebted for the wonderful atmosphere, which made me feel at home from the very fist moment i walked in. Finally, I would like to thank friends and family. Not only for keeping me motivated, but mostly for the joy they gave in these four years, which rank among the best as far as I am concerned.

Henk de Koning
Amsterdam, August 2007

# Contents

# 1 A research design to study the Six Sigma programme

## 1.1 The Six Sigma programme

The twentieth century saw an incredible development of professional organisations. The impact of technological advances is obvious, but besides these, innovations in management structures and methods have resulted in the highly productive organisations of today. When the race for outperforming competitors on quality and efficiency gained momentum, companies started to copy each other's best practices. Consultants and management gurus quickly jumped in and started giving names to these best practices: total quality management (TQM), just-in-time, business process reengineering, statistical process control, quality circles, lean manufacturing, continuous improvement, et cetera. Out of these methods, principles and approaches time has singled out the ones that really have added value. And while most approaches have been presented as panaceas at one time or another, time has shown that they are in fact complementary.

Six Sigma is not revolutionary. It is built on principles and methods that have proven themselves over the twentieth century. It has incorporated the most effective approaches and integrated them into a full programme. The most recent change has been the incorporation of Lean thinking, which is essentially a collection of best practices focused on making processes more efficient by removing waste. The Six Sigma programme contains several elements (De Mast, Does and De Koning, 2006):

- *A business context*: Six Sigma intends to help companies to survive in a competitive environment by creating cost savings, improving customer satisfaction, and improving organisational competence for innovation and continuous improvement.
- *An organisation structure*: Six Sigma offers a management structure for organizing continuous improvement of routine tasks. Six Sigma prescribes that improvement is done in a project-by-project fashion.
- *A methodology*: Six Sigma offers a method for carrying out improvement projects.

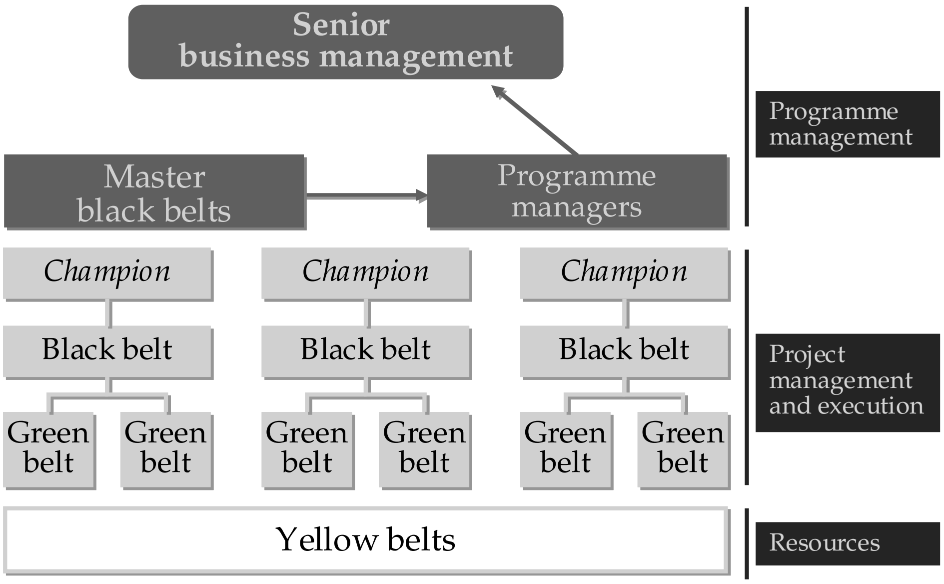

Figure 1.1 *Six Sigma organisation structure.*

These three elements will be elaborated in the next subsections.

### 1.1.1 Business context

The great majority of work in organisations is performed by routine operations (see De Mast, 2007). Manufacturing, backoffice processes, sales, marketing, healthcare are all functions performed (at least partly) in a routine manner. Six Sigma projects deal with the improvement of these routine operations, which we will generally refer to as 'processes', seeking to make them more effective and more efficient.

The direct benefits of Six Sigma projects consist of benefits derived from cost advantages (superior productivity and equipment utilisation, avoidance of capital expenditure, etc.) and customer satisfaction (reduced price sensitivity, growth of revenue or market share). Mid- to long-term benefits of Six Sigma are competence building in manufacturing or service delivery virtuosity and in continual and company-wide improvement and innovativeness.

### 1.1.2 Organisation structure

#### Project management and execution

The organisation structure of Six Sigma is shown in figure 1.1. Six Sigma project leaders are called black belts (BBs) or green belts (GBs). The BBs and GBs are expected to have intimate and detailed understanding of the process and the problem at hand. That implies that most BBs and GBs are recruited from the line organisation, and not from staff departments. To be able to work on structural problem systematically BBs and GBs are (partly) cleared from their regular job. Projects are selected and monitored by so-called Six Sigma champions. The champion is the project owner, in the sense that he is responsible for the process that the project aims to improve. Loosely said, the champion owns the problem, and hires the BB and GBs to solve it. During its execution, a project is reviewed several times by the champion, thus allowing him to adjust the direction that the BB or GB chooses. The project team is complemented by yellow belts (YBs), who bring in relevant knowledge, help collecting data, or perform other tasks.

### Programme management

The Six Sigma initiative is managed as a programme. Programme management consists of one or more master black belts (MBBs), one or more programme managers, and a programme director. The programme managers manage the day-to-day administration of Six Sigma, do the project selection, monitoring and control, and are responsible for the adjustment of the programme's course. The MBBs act as Six Sigma experts, and are the company's resource concerning Six Sigma's method and techniques. Moreover, they deliver the BB and GB training and they support projects. Finally, the programme director is part of the business's senior management and is ultimately accountable for the Six Sigma initiative.

### Training programme and project tracking

The BB training is usually given in four or five modules, with typically a three weeks' period in between successive modules. The BBs execute their first Six Sigma project in these intermittent periods, ensuring that the BBs immediately apply what they have learned to their project. During the training and his first project the BB is supported by the MBB, but he is expected to execute subsequent projects by himself. The GB training is an extract from the BB training, typically six to eight days (in either two or three modules).

The progress of projects is monitored by the champion and – especially during the training – by the MBB. Their main instrument for this purpose are reviews. Projects are typically reviewed four times, once after each of the phases Measure, Analyse, Improve and Control (see the next subsection for an explanation of the Six Sigma phases). During a review it is the champion's responsibility to assure that the project is still on the right track from the perspective of the organisation. The MBB assesses whether the BB or GB follows the Six Sigma method, and correctly applies relevant techniques.

A more detailed, prescriptive account of the Six Sigma project organisation is offered in De Mast, Does and De Koning (2006).

### 1.1.3 Six Sigma's method

Six Sigma seeks to elevate problem solving and quality improvement to a more professional level by training BBs and GBs in an attitude that can be described as scientific. Improvement actions are not based on perception and anecdotal evidence. But neither are they based on the notion of the omniscient specialist who, sitting behind his desk, derives a remedy by making clever deductions from his expert knowledge. Six Sigma's methodology incorporates several principles based on scientific method, such as defining problems in precise, operational terms, and grounding problem diagnosis in data analysis (for an overview of these principles, see De Mast and Bisgaard, 2007). These principles are embedded in the project phase structure that Six Sigma prescribes. Six Sigma projects follow five phases, called Define, Measure, Analyse, Improve, and Control (DMAIC). This step-by-step procedure for improvement projects provides BBs and GBs with a checklist that helps them to ask the right questions and deploy appropriate tools. It also helps structure progress reports and facilitates project tracking. In addition to DMAIC, Six Sigma offers an extensive collection of tools and techniques BBs and GBs use to attain intermediate results during DMAIC deployment.

### 1.1.4 Outline of this chapter

The subject of this thesis are the methodological aspects of the Six Sigma programme. These are taken to include a description of the type of goals that can be pursued with the method, but all other elements implied by the Six Sigma programme – the organisational structure, programme management, training and project tracking – are considered beyond the scope of this thesis. The objective of this thesis is loosely speaking the study of the validity and applicability of the methodological aspects of the Six Sigma programme.

This objective immediately confronts us with the problem that Six Sigma's methodology has many aspects that belong to different disciplines in science, such as statistics, methodology, management science, economics and quality engineering. Many of these aspects can be studied using standard research approaches, but there will be aspects for which we cannot fall back on a standard approach. In these cases we are forced to work-out a research design ourselves. In order to do so, in this chapter a scientifically sound approach for studying the validity and applicability of the Six Sigma method is worked-out. Several research methodologies are considered, whereupon a grounding research approach is developed. A comparison of the results of a literature review and the proposed research plan shows that current literature on the methodological aspects of Six Sigma does not meet scientific standards of precision and consistency. After this, the objective of the thesis is defined in more detail. A discussion about the motivation for this kind of research within the domain of mathematics and the outline of the thesis conclude this chapter. This chapter is for a large part based on De Koning and De Mast (2005).

## 1.2 Research methodology

The Six Sigma programme's method guides project leaders through a quality improvement project. We can therefore characterise the subject of study – Six Sigma's method – as a system of *prescriptions*: guidelines that tell a project leader what to do in order to reach a certain goal. It will be clear that a study of the Six Sigma method cannot be undertaken following the formal type of research that is common in mathematics, where theorems are derived by rules of deduction from a set of axioms.

### 1.2.1 Empirical research

One could consider to study the Six Sigma method following the approach of *empirical research*. In that case prescriptions (or rather, their application and the outcome of their application) are regarded as empirical phenomena. Measuring the success of their application, one could single out the successful elements of the Six Sigma method from the less successful. Although the study of records of past uses is an important element of the approach that we envisage, it is not sufficient. Merely recording which prescriptions correlate with successful applications and which do not, gives no explanation of the way the Six Sigma method works. To gain insight in how successful prescriptions work, one must understand the internal logic of the Six Sigma method.

### 1.2.2 Rational reconstruction

A second approach would be to understand the Six Sigma method as an attempt to reconstruct the unspoken “know-how” that skilled project leaders have collected during years of experience with quality improvement projects in the form of heuristics, best practices, and intuition. A major part of this know-how is “tacit” knowledge. This is knowledge which works in the background of consciousness and directs attention and action, but which is not made explicit or linguistically codified (Polanyi, 1958). The Six Sigma method could be regarded as an attempt to structure and explicate this tacit knowledge in order to facilitate the transfer of this knowledge to less experienced project leaders. Such an attempt is called a rational reconstruction and the related type of research is reconstruction research. Although the method of rational reconstruction is quite commonly used (for example in research in philosophy, law, information science), the literature on rational reconstruction is surprisingly meagre. The following paragraph is partly based on Kamlah (1980) and Davia (1998).

A rational reconstruction presents a given problematic complex (the object of reconstruction) in a similar, but more precise and more consistent formulation (the product of reconstruction) (Poser, 1980). The given problematic complex is typically intuitive, tacit knowledge. The simplest form of rational reconstruction is explication: the formulation of exact definitions for loosely defined concepts. Linguistic research is often reconstruction research (where one attempts to make explicit the grammatical rules that native speakers of a language know intuitively), as well as research in law (trying to reconstruct intuitive notions of right and wrong) and aspects of mathematics (e.g., the axiomatic set-up of probability as an attempt to formalise intuitive notions of probability).

Rational reconstructions could have a purely descriptive impetus. The emphasis is on reconstruction as "again"-construction ("re-" as "again"), making the object "more equal to itself" by extracting essential elements and reformulating and restructuring them. The main criteria for adequacy in this case are clarity, exactness and similarity to the original. One step further is a rational reconstruction with a prescriptive impetus. The emphasis is on "new"-construction ("re-" as "new"). The original material is taken as a starting point, but based on critical examination (on the basis of external, formal criteria such as logic), it is corrected. Besides clarity and exactness, one has in this case the criterion of consistency, which replaces the criterion of similarity. One could regard the Six Sigma method as an attempt to reconstruct the know-how needed to conduct a quality improvement project. Its validation would amount to a verification of:

- Similarity (To what extent does the methodological aspects of the Six Sigma programme correspond with the tacit knowledge of experienced project leaders?);
- Exactness (To what extent do definitions and classifications give unambiguous demarcations of concepts?);
- Clarity (How clearly organised is the exposition of the Six Sigma method?).

Would we regard Six Sigma method as a reconstruction with a prescriptive impetus, we would compromise the similarity criterion to the favour of consistency, i.e. to what extent is Six Sigma method free of internal contradictions?

Although elements of this approach are important, also this approach does not give us the whole picture, mostly because it makes the know-how of experienced project leaders the prime referee of the validity of the Six Sigma method. This may be suitable for other examples of reconstruction research (such as linguistics and law), but prescriptions are a means to an end, and empirical records of the extent to which they attain their intended ends are perhaps even more important referees of their validity.

### 1.2.3 Grounding research

Grounding research – the third option considered – is an investigation into the rationality of prescriptions, or in general, of actions. Actions are called rational if they are criticisable and can be grounded (Habermas, 1981, pp.25ff.). Rational actions embody certain presumed knowledge, and therefore imply a validity claim. For example, if a person performs a certain action with a specific purpose in mind, he implicitly claims the effectiveness of the chosen action in attaining the purpose. Or if a person makes a statement about certain matters of fact, he claims the truth of his statement. The rationality in these actions consists of their claimed effectiveness or truth. To ground an action is to show that these claims are warranted, i.e., that the knowledge on which they are based is true. Different types of actions raise different validity claims ("effectiveness", "truth"), and should, consequently, be grounded differently, depending on the precise manner in which the action relates to the knowledge that underpins it (Habermas, 1981, p.67). One of the reasons why the rationality of actions matters, is that their criticisability makes it possible to improve them. Thus, grounding is closely related to learning (Habermas, 1981, pp.38–39).

In order to ground the Six Sigma method (which, as we noted, can be seen as a system of prescriptions), we have to formulate the validity claims that it makes, and next, verify that these claims are warranted. The basic form of a prescription is:

*Given a certain situation,*
*then take action X in order to attain a certain goal Y.* (1)

The validity claim that a prescription makes is "usefulness". This claim is composed of two claims:

*The goal Y is legitimate*, (2)

*Cause (action) X results in effect (goal) Y.* (3)

To ground (or validate the usefulness of) a prescription of the form (1), one would have to do the following things:

1. *Rational reconstruction*. Bring the prescription in the form (1),
2. *Value grounding*. Validate the legitimacy of goal Y (2),
3. *Empirical grounding*. Validate the explanatory argument (3) by providing empirical evidence that confirms the stated X-Y relation (2),
4. *Theoretical grounding*. Validate the explanatory argument (3) by providing another statement or theory, which is valid and which implies (3),

5. *Specification of applicability*. And, finally, specify the situations in which it is applicable.

The analysis above was much influenced by a similar analysis by Lind and Goldkuhl (2002) who study the grounding of methods for business change. The analysis specifies the various aspects of a complete grounding study of the Six Sigma method. Below, these aspects are elaborated, thus establishing a research plan. The results of a part of this research plan will be presented in this thesis.

### 1.2.4 Research design for this thesis

**Rational reconstruction**

The Six Sigma method is formulated in unscientific language (ranging from imprecise and incoherent to meaningless and silly). For example, the demarcation of the phases Measure, Analyse, Improve and Control in Harry (1997, p.21.19) is inconsistent with the steps that these phases are comprised of (p.21.22). Definitions of concepts such as CTQ (Harry, 1997, p.12.20) do not meet scientific standards of precision. Moreover, while most accounts of the methodology agree on the DMAIC phase structure, descriptions of the steps that these phases are comprised of and the tools that are prescribed for them diverge. The first phase of this research project, therefore, aims to distill from the imprecise and loose accounts available a precise, consistent and articulate description of Six Sigma's method. The resulting reconstruction of the Six Sigma method will be presented in chapter 2.

**Value grounding**

Prescriptions are legitimatised by their goal. Are Six Sigma method's goal and its associated values valid? Sometimes it is stated that the goal of Six Sigma projects is to bring each process on the Six Sigma level of quality (3.4 parts per million defects) (Harry, 1997, pp.2.11-18). From an economical point of view this claim is in this general form untenable, and it is questionable whether the majority of Six Sigma projects really aim at this objective (let alone whether it is possible to prove that such a low defect rate is attained, given the enormous sample sizes that are required to do so). Other descriptions of the goal of Six Sigma projects are quality improvement, breakthrough, variation reduction, and defect reduction. In turn, these goals are legitimatised by concepts as Costs of Poor Quality (Breyfogle, 1999, chapter 1). The adequacy of this paradigm should be studied, and alternative paradigms (borrowed from, for instance, economics and strategic management rather than quality management) should be explored. This matter is partly covered in chapter 2, in which the value propositions provided in the literature are reconstructed. More research on this point is needed, though, especially aimed at integrating these accounts with theory in business economics.

### Theoretical grounding

The effectiveness of prescriptions can be validated by explaining from (an external) theory why they work. For improvement strategies this is done in De Mast (2002), in which the effectiveness of the Six Sigma method is explained by showing that it follows scientific method for empirical inquiry.

### Empirical grounding

Empirical grounding takes the form of survey research, in which the effectiveness of the Six Sigma method is estimated from empirical data, possibly as function of various factors. An example of the type of surveys that is meant is Easton and Jarrell (1988), who study the effectiveness of TQM.

### Specification of applicability

The analyses announced above should provide indications about the applicability of the Six Sigma method. They should identify factors which affect the effectiveness of the Six Sigma method, or which could even make it completely ineffective (an impossibility to collect measurements, to mention an example). However, this issue should also be limited to methodological conditions (organisational conditions that should be met in order to conduct an improvement project successfully – for instance the question to what industry LSS is applied – should be studied elsewhere).

## 1.3 Literature on the Six Sigma method

This section presents a first inventory of scientific literature on Six Sigma, especially with the research approach described above in mind. We considered articles that have been published in four scientific journals in the field of industrial statistics:

- Quality Engineering,
- Quality and Reliability Engineering International,
- Journal of Quality Technology,
- International Journal of Quality and Reliability Management.

In addition, two books were considered in the overview: Harry (1997) and Breyfogle (1999). The objective of the inventory is to assess to what extent the elements of the research design sketched in the previous section are covered in literature.

### 1.3.1 Overview of relevant papers

It was noted earlier that Harry's (1997) exposition of the Six Sigma method does not meet scientific standards of consistency and precision. Breyfogle's (1999) exposition has similar shortcomings, and we did not find better descriptions, underscoring the observation that an explication or *rational reconstruction* of the Six Sigma method is needed.

The issue of *value grounding* is addressed by Harry (1997), who focuses on the hidden factory model to validate Six Sigma goals (p.14.10). Defects have an effect on the amount of rework, which in turn affects costs, cycle times and required inventory levels (p.17.16). The Six Sigma method reduces these, Harry reasons (pp.2.12–13). Other accounts of Six Sigma in terms of organisational goals are given by respectively Wasserman and Lindland (1996) and Bisgaard and Freiesleben (2001).

The former authors use the cost-of-quality framework to advocate the use of the Six Sigma method. They argue that there is an essential trade-off between the cost-of-control versus cost-of-lack-of-control. The optimal quality level (in terms of conformance) is at the level for which the sum of cost-of-control and cost-of-lack-of-control is lowest. This optimum shifts to a higher value as customer expectations rise. As a consequence, in view of ever increasing customer expectations, organisations are forced to provide higher and higher quality (conformance) levels. This justifies the deployment of the Six Sigma method.

Bisgaard and Freiesleben (2001) show that defect elimination and prevention can create financial results (high return on investment). The conclusion is that "(1) quality improvement is an investment not a cost and (2) any financial benefit of improving operational efficiency, the stated goal of Six Sigma, goes directly to the bottom line and often provides an exceptionally high rate of return." Moreover, because reducing defects is an internal affair, it is on principle easier to reduce cost than to increase sales. It appears that the intent of Bisgaard and Freiesleben is to give an illustration, rather than a scientific understanding of the validity of the goals of Six Sigma.

An integrated account of the functionality and purpose of Six Sigma lacks. All three accounts frame the benefits of Six Sigma in accountancy terms (costs) and focus on the Six Sigma method as a method for quality improvement. The costs paradigm seems valid, but is one-sided. The functionality of the Six Sigma method should also be studied from other perspectives, such as business strategy, process innovation, the use of knowledge in organisations, and others. The limitation of the Six Sigma method as an approach for quality improvement is overly restrictive, because many projects focus on cost reduction, cycle time reduction or yield improvement. These can only be subsumed under quality by stretching the meaning of that term. This type of conceptual erosion is scientifically speaking undesirable.

*Theoretical grounding* of the Six Sigma method has been done by the De Mast (2003; 2004), who shows that the Six Sigma method follows scientific method for empirical inquiry. The author also identifies a number of anomalies, in which the Six Sigma method deviates from standard research methodology for no apparent logical reason. The lack of emphasis of the iterative nature of empirical research, and the underexposure of the elaboration phase in the Six Sigma method serve as examples.

The literature is poor when it comes to *empirical grounding* of the Six Sigma method. Hahn, Doganaksoy and Hoerl (2000) mention three famous showcases of billion dollar savings due to Six Sigma (Motorola, AlliedSignal, General Electric), but this is anecdotal evidence. An example of serious empirical grounding (having a scientifically acceptable research design) of quality improvement methodologies is the research by Easton and Jarrell (1988). These authors have investigated the impact of TQM on financial performance. Although TQM is different from the Six Sigma method, their methodology may be useful for evaluation of this approach.

Hardly any attempt has been made to show in which situations, under what conditions, and for what purposes the application of the Six Sigma method is successful. A possible reason for a lack of this type of research is the lack of agreement of what Six Sigma is. Opinions about the conditions under which the Six Sigma method applies diverge. Goh (2002) claims that the Six Sigma method does not apply to knowledge based environments, such as scientific research. Others (Hahn, Doganaksoy and Hoerl, 2000) see tremendous opportunities for Six Sigma in virtually any context. Along the same lines Sanders and Hild (2000) contend that Six Sigma is applicable to any business process: “The concepts of measuring process performance, making decisions via data, increasing efficiency, and improving quality are obviously much needed and logically applicable in the administrative and business areas of organizations.” However, all these viewpoints are based on personal experience instead of systematic empirical research.

### 1.3.2 Conclusions of literature review

One can conclude that the Six Sigma method has not been grounded sufficiently in current literature: the extent to which the questions of the research plan are addressed ranges from poorly to not at all. Specifically, we draw the following conclusions:

1. Expositions of the Six Sigma method fail to meet standards of consistency and precision.
2. There have been some attempts at value and empirical grounding of the Six Sigma method, but these attempts are insufficient from a scientific point of view. Legitimisation of the goals of application of the Six Sigma method is too one-

sidedly focused on costs. Empirical grounding relies solely on personal experiences of practitioners, not on serious empirical research.

3. Theoretical grounding of the Six Sigma method has been done to some extent. The conclusion is that the Six Sigma method largely follows standard research methodology. Directions for improvement have been identified.

## 1.4 Objective of the thesis

The objective of this thesis is to scientifically ground the methodological aspects of the Six Sigma programme. Methodologies such as Six Sigma consist of four classes of elements, which are listed and discussed below:

- *Business context*. At the background of the Six Sigma programme is a philosophy that presents a business strategy. This philosophy provides the motivation for implementing the programme by specifying which benefits it is claimed to have, and – of more importance here – the type of objectives that can be pursued with the methodology.
- *Stepwise strategy*. The Six Sigma method gives a stepwise procedure for tackling projects. Harry (1997), for instance, proposes 12 steps that are grouped in four phases. Steps define end terms (the deliverable of the step) and prescribe in which format they should be documented. For example, the end term of Harry's step 4 is that the process's performance is estimated; this result should be reported in the form of a capability index Z.
- *Tools and techniques*. The Six Sigma method offers a wide range of procedures that are intended to assist the project leader in attaining intermediate results. Some of these tools and techniques are linked to particular steps of the strategy (for instance the gauge R&R technique proposed for Harry's step 3, "Validate measurement system"), others are more general (for instance statistical estimation). Some tools and techniques are statistical, others are non-statistical.
- *Concepts and classifications*. In order to communicate the elements above, the Six Sigma method offers concepts (such as the hidden factory and CTQ) and classifications (the phases Measure, Analyse, Improve, Control; the distinction between vital Xs and trivial Xs).

The thesis aims to ground the four aspects of the Six Sigma method outlined above, and, in addition, make contributions to any of these classes of elements.

## 1.5 Motivation for the thesis

In this section three issues relevant to the motivation for this this will be examined:

- The motivation for researching Six Sigma. It is examined why researching Six Sigma is relevant to both practitioners and to the scientific community.
- The motivation for executing the research within the mathematics discipline. To this end the relation between research paradigms in industrial statistics and the research paradigm in mathematics will be analysed.
- The validity of this research as scientific research. The relevance and quality of this thesis as valid scientific research will be discussed.

### 1.5.1 Motivation for research on Six Sigma

Scientific research on Six Sigma should aim at two objectives:

- It should provide *understanding* of Six Sigma. This is primarily relevant to the scientific community.
- It should enable *effective use* of Six Sigma by the practitioner.

In order for the practitioner to be able to make *effective use* of Six Sigma, he needs to know (i) when to apply, and (ii) how to apply the method. The first condition to meet these objectives is the existence of a crystal clear description of the Six Sigma method (rational reconstruction, see section 1.2.4). Otherwise it is hard to apply the method in the right way. Furthermore, research should provide guidelines to indicate for what kind of objectives Six Sigma is applicable, so that the practitioner knows when to apply Six Sigma (specification of applicability, section 1.2.4).

A first step to *understanding* of Six Sigma, the stake of the scientific community, is also a crystal clear formulation of Six Sigma. Otherwise Six Sigma can not be investigated in the first place. Secondly, understanding Six Sigma means that one can explain its effectiveness. This comes down to providing a good grounding of Six Sigma: One explains from a theoretical perspective why Six Sigma works (theoretical grounding, section 1.2.4) and tests empirically Six Sigma's effectiveness (empirical grounding, section 1.2.4).

Part of the reason for doing research on Six Sigma is that it is considered as the de facto standard for quality improvement in the business world. As such it has been one of the most successful and large scale applications of statistical methods. Thus it is an important vehicle for the statistical sciences for getting their methods applied. Tens, perhaps, hundreds of thousands of BBs and GBs are trained in advanced level statistical techniques, owing to Six Sigma. And thanks to Six Sigma, statistics has found its way to board rooms and is having an impact on businesses.

In spite of its enormous impact in the business environment, Six Sigma has not been well researched. There is an extensive literature on the subject, but this literature lacks the accuracy and critical attitude of scientific research. A similar observation has been made by Linderman, Schroeder, Zaheer and Choo (2003): “While Six Sigma has made a big impact on industry, the academic community lags behind in understanding of Six Sigma” (cf. Stephens, 2003, p.28). So, we can conclude that the objectives of scientific research of Six Sigma just identified – providing a crystal clear description and good grounding of Six Sigma – have not been met so far. Therefore scientific research into the validity and applicability of Six Sigma is important.

### 1.5.2 Industrial statistics and mathematical research

It will be clear that the envisaged type of research is not purely mathematical research. Because of this fact, this research might raise questions, and therefore, it is important to clarify the relationship between industrial statistics and mathematics. Industrial statistics could be described as (De Mast and Does, 2006):

> “The discipline which develops quantitative methods and paradigms for inquiry and routine decision making in industry” (De Mast and Does, p.273, 2006)

From reading the industrial statistical journals (such as Technometrics, Journal of Quality Technology, Quality and Reliability Engineering International and Quality Engineering) one could get the impression that industrial statistics is a specialism within mathematics. Statistical inference is, however, certainly not a form of mathematical reasoning: in the latter, theorems are derived by deduction from axioms; in the former, conclusions are arrived at by inductive reasoning. Mathematics enters where statisticians study an empirical system by advancing a model for it (see Mayo, 1996, chapter 5). The internal logic of the model (with all its standard machinery of reasoning, such as hypothesis testing and confidence interval estimation) is based on mathematical axioms and deductions. But the definition of the system under study is an empirical matter, not a mathematical one (i.e., empirical reality is the guiding principle here, not mathematical axioms), and the translation of inferences for the model to conclusions about the empirical system requires extra-mathematical (inductive) reasoning.

We can conclude that mathematics is only a part of industrial statistics and that research in industrial statistics should not be restricted to mathematical research. In fact, this holds for statistics in general, and even for probability: “Probability is no more a branch of mathematics than is physics, although it owes a great debt to mathematics for its formulation and development” (Fine, 1988). Statistics *uses* a lot of mathematics, but is in itself not a form of mathematics.

If research in industrial statistics is not restricted to methods used in mathematics, it needs other methods in addition. In section 1.2 rational reconstruction and grounding research were introduced as an important part of the methodology suitable for researching Six Sigma. The use of these methods in industrial statistics is not new, however. If we turn to the history of this field we see a tradition in which mathematical research is complemented with rational reconstruction and grounding research, although this is generally not acknowledged explicitly and the terms rational reconstruction are virtually never mentioned. We give two examples to illustrate the omnipresent occurrence of rational reconstruction and grounding in statistical research.

Shewhart is a first example (see De Mast and Does, 2006). He introduced the control chart which is used to determine whether the process has changed or is stable. Before the existence of the control chart operators already intervened in the process if they thought something had changed, but the distinction between an assignable cause and chance cause was intuitive at best. As a consequence, operators tended to intervene even when nothing really had changed. Their intuitive, imperfect notion of the distinction between assignable causes and chance causes was provided a precise, mathematical formulation by Shewhart. So he made explicit intuitive and inarticulate notions that already existed. This is the essence of rational reconstruction.

Similarly, concepts used in acceptance sampling, such as acceptable quality level (AQL) and limiting quality level (LQL) and tools such as the operating characteristic (OC-) curve provide a precise and consistent framework for analyzing sampling schemes. In sampling inspection 100% inspection is replaced by checking only a sample of a batch of products. Intuitively it is easy to understand that accepting or rejecting batches of products based on inspection of only a sample instead of a whole batch creates risks. The sample could give too optimistic or too pessimistic an impression of the batch. The consumer's risk (the case that we have too optimistic an impression of the batch) is made mathematically precise with the concept of LQL, whereas the producer's risk (the case that we have too pessimistic an impression of the batch) is made precise with the concept of AQL. In combination with the OC-curve they provide a framework that can be used to analyse the effect of a chosen sampling plan on both the consumer's and producer's risk. Intuitive, inarticulate understanding of the trade-off between consumer's and producer's risk existed before the development of the acceptance sampling framework, but this framework provided this intuitive and inarticulate understanding with a mathematically precise and consistent fundament. Again, this is a clear case of rational reconstruction. The definition of the concepts of AQL, LQL and OC-curve to replace intuitive notions is rational reconstruction (although the further development of the framework uses a lot of mathematics).

### 1.5.3 Scientific research

We have observed that the type of research pursued here is not just mathematical research, but much broader than that. This makes it harder to judge whether it qualifies as valid scientific research, because the standard criteria used for purely mathematical research are not sufficient. The promotional regulation of the University of Amsterdam provides guidelines as to what constitutes good research in the form of a number of criteria. These are:

1. Development and clear expression of a problem statement;
2. Organisation, analysis, and processing the subject of the thesis;
3. Originality and creativity in the treatment of the subject of the thesis;
4. Purity of the method used in the analysis;
5. Critical judgement of existing theories and existing opinions;
6. Balanced structure, appropriate style, and clear phrasing of the thesis.

The list of criteria makes it clear that different kinds of research qualify as scientific. On the one hand research exists that tackles a well-defined, clearly delineated (academic) problem with existing methods. On the other hand there is research – like the current research – in which a real-life, but typically not well-defined problem is analysed. One tries to properly define and analyse this fuzzy problem with scientific precision and objectivity.

To illustrate the difference we give an example. Vermaat and Does (2006) provide an illustration of research in which a well-defined problem is tackled with existing methods. They improve, using a semi-Bayesian approach, the traditional Shewhart control chart for a special case in which the performance is not satisfying (non-normality and a sample size in the range of 250 to 1000). Quite the opposite is the research by De Mast (2003; 2004), which provides an example of research in which a real-life, not well-defined problem is analysed. He compares three approaches to quality improvement, namely the Six Sigma programme, the Shainin System, and Taguchi's methods. To this end first a methodological framework is created, then these methods are reconstructed from literature and compared with the help of the methodological framework.

This thesis was written from a strong conviction that both type of problems – the well-defined allowing advanced analysis and refinement, and the ill-defined requiring scholarly inclination – are valid subjects for a PhD-thesis. In various disciplines, a tendency has been noted under academics to shy away from the second type of problems (cf. Bennis and O'Toole, 2005), even to the extent that appreciation of research is heavily biased towards the first type of problems. A paragraph by the influential economist Galbraith (1991) merits quoting at length:

> "The central assumption of classical economics (...) lends itself admirably to technical and mathematical refinement. This, in turn, is tested not by its representation of the real world but by its internal logic and the theoretical

> and mathematical competence that is brought to bear in analysis and exposition. From this closed intellectual exercise, which is fascinating to its participants, intruders and critics are excluded, often by their own choice, as being technically unqualified. And, a more significant matter, so is the reality of economic life, which, alas, is not, in its varied disorder, suitable for mathematical replication." (Galbraith, 1991, p.285)

Also in statistics there is, in the conviction of the author, a tendency to appraise research based on its level of mathematical refinement and elegance rather than its usefulness for or validity in bearing on real data analysis.

Inquiry of real-life issues in both economics and industrial statistics does not lend itself to be fully reduced to mathematical conundrums. Moreover, only judging these disciplines by criteria of internal logic and theoretical and mathematical competence would compromise other important criteria for good research such as usefulness and practical relevance. Therefore it is important to entertain both kinds of research: tackling well-defined problems with existing methods, *and* real-life, but fuzzy, not well-defined problems analysed with scientific precision and objectivity. As indicated before, the current research belongs to the latter type.

## 1.6 Outline of the thesis

In the next chapter the rational reconstruction of the Six Sigma method is carried out. This analysis results in a precisely formulated account of the methodology of the Six Sigma method (its stepwise strategy, tools and techniques), its business context, and its terminology (concepts and classifications). The chapter also discusses the integration of Six Sigma and Lean Thinking. It will be argued that Lean and Six Sigma are separate approaches to process improvement with complementary strengths. When combined as Lean Six Sigma this approach provides a unified framework for systematically developing process and product improvements in service and industry.

The final two chapters of the thesis make contributions to Six Sigma's body of tools and techniques. The third chapter of this thesis focuses on a analysis model that is at the heart of the first steps of Six Sigma's DMAIC method, the CTQ flowdown. The CTQ flowdown is a model for developing clear project definitions and for clarification of the business rationale of an improvement project. This chapter provides a theoretical grounding of the CTQ flowdown, but also provides practitioners with a prescriptive template. Our model allows us to define a number of generic categories of Lean Six Sigma projects in financial services and healthcare. Moreover, this chapter contributes to the theoretical grounding of LSS, by validating with the help of by external scientific theories the effectiveness of a part of the stepwise strategy of LSS.

The last part of the thesis dealt with the so-called gauge repeatability and reproducibility study, which is the standard method for assessing a measurement system's precision. It plays an important role in the Measure phase of Lean Six Sigma projects, but problems occur when one deals with destructive measurements. An approach to deal with some of these situations is developed in chapter 4.

# 2 A rational reconstruction of Six Sigma and Lean Six Sigma

In the previous chapter it was argued that the first step to scientifically ground the Six Sigma method consists of a rational reconstruction of four classes of elements forming the methodological part of the Six Sigma approach, namely:

1. Business context,
2. Stepwise strategy,
3. Tools and techniques,
4. Concepts and classifications.

The purpose of this chapter is to develop a consistent and crystallised exposition of these four classes of elements. It is the intention to present existing accounts of Six Sigma's methodology as clear as possible, thus aiming for a descriptive (rather than prescriptive) reconstruction. It is not the intention to evaluate these accounts against external criteria, such as theoretical frameworks in the literature on quality management or methodology. A comparable study is Reed, Lemak and Mero (2000), who distill from existing literature a set of core principles of total quality management (TQM). The material that the reconstruction starts from consists of accounts of the four elements mentioned above – business context, stepwise strategy, tools and techniques, concepts and classifications – in the scientific and non-scientific literature. Specifically, we consider articles that have been published in seven journals relevant to industrial statistics:

- Quality Engineering
- Quality Progress
- Quality and Reliability Engineering International
- Journal of Quality Technology
- International Journal of Quality and Reliability Management
- The American Statistician
- International Journal of Six Sigma and Competitive Advantage

In addition, nine books were studied in this research: Harry (1997), Breyfogle (1999), Harry and Schroeder (2000), Pande, Neuman and Cavanagh (2000), Eckes (2001), Pyzdek (2001), Creveling, Slutsky and Antis (2003), Park (2003), and Stephens (2003).

The subsequent sections present the reconstruction of the business context, stepwise strategy, and tools and techniques. Relevant concepts and classifications are reviewed and defined when they are needed. In the second part of this chapter the synthesis between Lean and Six Sigma is expounded.

The first part of this chapter is largely based on De Koning and De Mast (2006), the second part on a series of articles on Lean Six Sigma (De Koning, Verver, Van den Heuvel, Bisgaard, and Does, 2006; Van den Heuvel, Does, and De Koning, 2006; De Koning, Does, and Bisgaard, 2007).

## 2.1 Reconstruction of the business context

The business context of Six Sigma refers to the method's purpose. In the studied literature, the usefulness of Six Sigma is argued from three perspectives:

- Showcases, arguing Six Sigma's usefulness from anecdotal evidence of successful applications.
- The hidden factory and cost of poor quality models, which argue Six Sigma's usefulness from its power to improve a company's cost structure by improving quality.
- Strategic benefits associated with improved quality and customer satisfaction, notably, market share increase and reduced price sensitivity.

### Showcases

The Six Sigma literature abounds in showcases, with Motorola, AlliedSignal, and General Electric being the most spectacular ones (see Harry, 1997; Breyfogle, 1999; Hahn, Hill, Hoerl and Zinkgraf, 1999; Pande, Neuman and Cavanagh, 2001). Showcases argue the usefulness of Six Sigma from benefits – mostly monetary – claimed by companies that implemented the programme. To give an example, Hahn, Hill, Hoerl and Zinkgraf (1999) remark that "The Six Sigma initiative was at least one key factor in Motorola winning the coveted 1988 Malcolm Baldrige Award for Quality, and produced reported savings of over $940 million in three years."

### Hidden factory and cost of poor quality models

Cost of poor quality (COPQ) is "any cost that would not have been expended if quality were perfect" (Pyzdek, 2001, p.163). In the Six Sigma literature, COPQ is usually divided in four categories: prevention, appraisal, external, and internal failure costs (Breyfogle, 1999, p.4). The COPQ concept is used to establish a relation between conformance quality and production costs. The main idea is that conformance quality improvement reduces costs associated with internal or external failure (called cost of lack of control by Wasserman and Lindland, 1996), and of appraisal costs.

The hidden factory model makes the same argument: hidden factory refers to all extra activities needed because of nonconformance (Harry, 1997, p.14.10). Nonconformance results in a larger hidden factory, which brings about higher costs, higher cycle times, higher inventory levels, lower reliability, et cetera (see, for example, Harry, 1997, pp.15.5 and 17.4). Improving conformance quality by deployment of the Six Sigma programme reduces costs, and this benefit goes directly to the bottom line (Bisgaard and Freiesleben, 2001). What adds to the importance of focusing on conformance quality is that cost of poor quality contains substantial hidden components (Harry, 1997, p.17.3), which are often ignored or forgotten. Furthermore, the ever increasing complexity of products and processes leverages the impact of nonconformance onto production cost (Bisgaard and Freiesleben, 2001). Thus, the usefulness of Six Sigma is argued from its power to tackle quality problems effectively, which is claimed to improve a company's cost structure.

#### Strategic benefits associated with quality and customer satisfaction

Improved quality, it is argued, results in more value and thus satisfaction for customers (Creveling, Slutsky and Antis, 2003, p.31). This advantage could be turned to account, according to the Six Sigma literature, either in the form of increased market share, or in the form of higher profit margins (Harry, 1998).

### The concept of quality

The term *quality* plays an important role in the descriptions above, and in fact, Six Sigma is usually described as a quality improvement strategy. This section reconstructs what various authors have in mind when they use the term.

Creveling, Slutsky and Antis (2003, p.31) describe quality as a total of product and service characteristics, such as performance, features, reliability, conformance, durability, serviceability, aesthetics and perceived quality. In line with traditional notions of quality (e.g., quality as "fitness for use"), the customer is taken as the criterion for quality: "Quality [is] performance to standard expected by the customer" (Harry, 1997, p.3.6). *Customer* sometimes refers to the end-user, but most authors stretch the meaning of the term to include entities in the producing company: "Many teams make the mistake of assuming that the customer is the external entity that pays the bill" (Eckes, 2001, p.50), and: "Customer [is] anyone internal or external to the organization who

comes in contact with the product or output of my work" (Harry, 1997, p.3.6). A further generalisation of the term quality is introduced by Harry and Schroeder (2000, p.6): "The Six Sigma Breakthrough Strategy broadens the definition of quality to include economic value and practical utility to both the company and the customer. We say that quality is a state in which value entitlement is realised for the customer and provider in every aspect of the business relationship."
What do Six Sigma authors mean when they relate Six Sigma's benefits to quality improvement? Looking at the third perspective mentioned above (strategic benefits associated with quality and customer satisfaction), it is clear that quality is used to describe properties of products (including services). It is also clear that *customer* refers to the paying customer. It is proposed to discern this notion as *product quality*, and to define:

- *Definition*: product quality refers to product characteristics and the extent to which they meet customer (meaning: end-user) demands. Product characteristics that together make up product quality are: performance, features, reliability, conformance, durability, serviceability, aesthetics and perceived quality.

Regarding the second perspective above (the hidden factory and cost of poor quality models), *quality* and *quality improvement* refer to properties of processes, rather than properties of individual products. In its most limited scope, *quality* is used as synonymous to *process capability*:

- *Definition*: process capability refers to the extent to which a process makes products which are free from defects.

The sigma metric of quality (e.g. De Mast (2007)) is a measure of process quality in this sense. But references to cycle time, yield, and other indicators of "economic value" (in the definition of Harry and Schroeder cited above) suggest a broader definition:

- *Definition*: process quality reflects the demands of internal customers, and comes down to effectiveness (the extent to which a process provides required features) and efficiency (being effective at low cost). Dimensions of process quality include defect rates, but as well cycle time, yield and production costs not related to defects.

It is concluded that the Six Sigma literature argues the usefulness of the method from its power to improve product quality (which is claimed to result in strategic advantages such as increased market share or reduced price sensitivity) or improve process quality (which is claimed to improve a company's cost structure), both of which are illustrated from showcases. Figure 2.1 conceptualises these lines of argumentation.

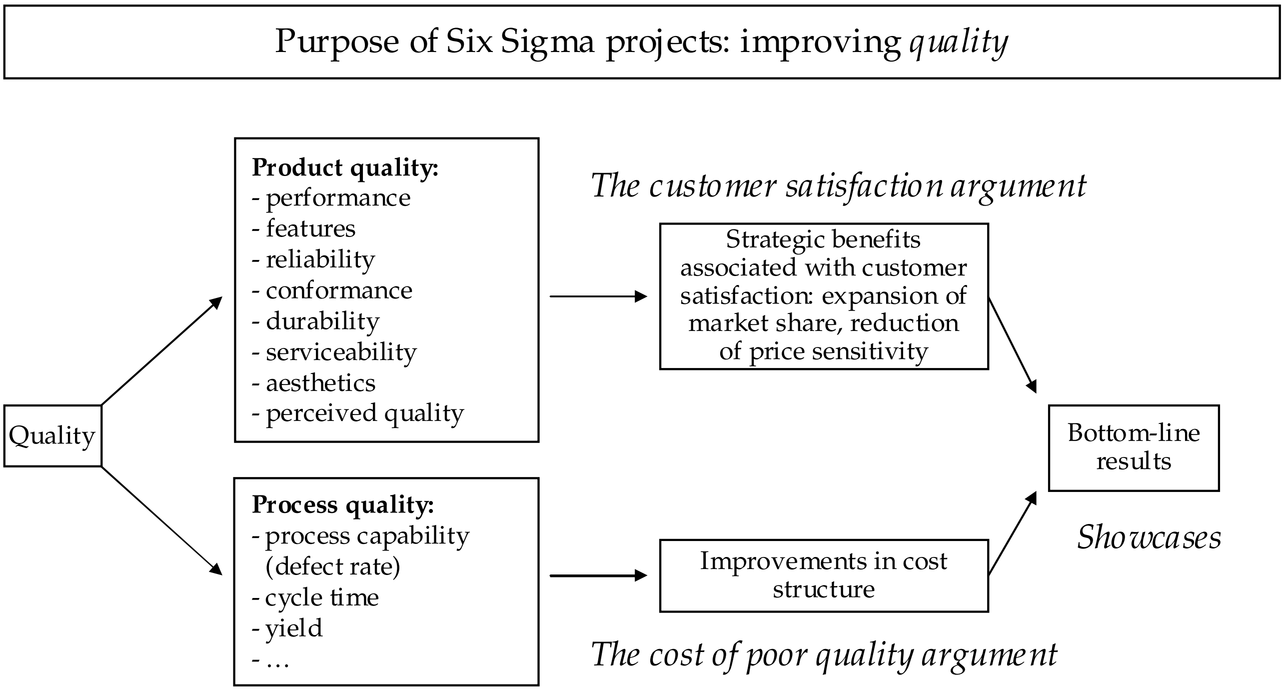

Figure 2.1 *Rational reconstruction of Six Sigma's business context.*

## 2.2 Reconstruction of the strategy and step plan

Six Sigma operationalises its strategy with the help of two types of concepts:

- Steps, which either specify the actions a project leader has to perform (for instance: do a process capability analysis), or the intermediate result a project leader has to achieve (for instance: establish the capability of the process), or a combination of both.
- Phases, which group together a number of steps.

Before reconstructing Six Sigma's strategy, several concepts that play important roles in the methodology are studied.

### The concepts of CTQ and influence factor

The particular subject of a project is made measurable in the form of one or more quality characteristics, which most Six Sigma authors (Harry, 1997; Hahn, Doganaksoy and Hoerl, 2000; Pande, Neuman and Cavanagh, 2000; Rasis Gitlow, and Popovich, 2002; Snee, 2004) call critical to quality characteristics or CTQs. Other terms used to denote the same concept are key process output variables (KPOVs) (Breyfogle, 1999), and Ys (Hahn, Hill, Hoerl and Zinkgraf, 1999).

Six Sigma projects aim to achieve improvement by identifying factors that influence the relevant CTQs (see later in this section). These influence factors, and especially the "vital few", are referred to as Xs, root causes (Hahn, Hill, Hoerl and Zinkgraf, 1999; Pande, Neuman and Cavanagh, 2000; Eckes, 2001; Rasis, Gitlow, and Popovich, 2002; Snee, 2004), key (input) process variables (KPIVs) (Breyfogle, 1999; Hahn, Doganaksoy and Hoerl, 2000), leverage variables or independent variables (Harry, 1997). We define:

- *Definition*: CTQs are dimensions of product and process quality (as defined in the previous section). In particular: CTQs are those quality dimensions on which a Six Sigma project aims to achieve improvement.
- *Definition*: Influence factors are factors that causally affect the CTQ. The *vital few* influence factors consist of the group of factors whose effects dominate the effects of all other factors (the *trivial many*).

**Phases: DMAIC**

The Six Sigma method entails a four phase procedure consisting of the phases Measure (M), Analyse (A), Improve (I) and Control (C); especially in more recent accounts, a Define (D) phase is added before the Measure phase. This MAIC or DMAIC structure is adopted by all authors taken into consideration, except Pyzdek (2001). The basis of the reconstruction of the functionality of these phases is formed by descriptions and definitions taken from the following sources:

1. Harry (1997, p.21.7);
2. Breyfogle (1999);
3. Hahn, Hill, Hoerl and Zinkgraf (1999);
4. Hahn, Doganaksoy and Hoerl (2000);
5. Pande, Neuman and Cavanagh (2000, p.239; p.251; p.276, p.337);
6. Rasis, Gitlow and Popovich (2002).

Without listing all descriptions and definitions found in these sources, Figure 2.2 presents a limited number of typical descriptions of each phase's functionality and their source. Based on this material, we constructed definitions of each phase's functionality, which are also presented in Figure 2.2 and discussed below. Since rational reconstructions aim to define the communalities in the various accounts that are used as a source, it is likely that individual accounts deviate from the resulting account. The listing below highlights serious deviations.

Although some descriptions for the Define and Measure phases in the above mentioned sources are clearer than others, there are no serious inconsistencies. The following two definitions are proposed:

| ***Define*** |
|---|
| Establishment of the rationale for a Six Sigma project. 6) |
| Define the problem to be solved, including customer impact and potential benefits. 4) |
| *Generic: Problem selection and benefit analysis.* |

| ***Measure*** |
|---|
| Identify the critical-to-quality characteristics (CTQs) of the product or service. Verify measurement capability. Baseline the current defect rate and set goals for improvement. 4) |
| This phase is concerned with selecting one or more product characteristics; i.e., dependent variables, mapping the respective process, making the necessary measurements, recording the results on process "control cards," and estimating the short- and long-term process capability. 1) |
| *Generic: Translation of the problem into a measurable form, and measurement of the current situation.* |

| ***Analyse*** |
|---|
| Understand root causes of why defects occur; identify key process variables that cause defects. 4) |
| undertaken to identify the common factors of successful performance; i.e., what factors explain best-in-class performance. 1) |
| Analyze the preliminary data [collected in the Measure phase] to document current performance (baseline process capability), and to begin identifying root causes of defects (i.e., the “X’s”, or independent variables) and their impact, and act accordingly. 3) |
| *Generic: Identification of influence factors and causes that determine the CTQ’s behaviour.* |

| ***Improve*** |
|---|
| Determine how to intervene in the process to significantly reduce the defect levels. 3) |
| Generating, selecting, and implementing solutions. 5) |
| *Generic: Design and implementation of adjustments to the process to improve the performance of the CTQs.* |

| ***Control*** |
|---|
| Implement ongoing measures and actions to sustain improvement. 5) |
| improvements are sustained, even though significant resources may no longer be focused on the problem. 3) |
| *Generic: Adjustment of the process management and control system in order that improvements are sustainable.* |

Figure 2.2 *Rational reconstruction of Six Sigma’s phase structure; notes refer to the numbered sources listed in the text .*

- Define phase: *Problem selection and benefit analysis.*
- Measure phase: *Translation of the problem into a measurable form, and measurement of the current situation.*

The majority of authors is followed in defining the functionality of the Analyse phase as:

- Analyse phase: *Identification of influence factors and causes that determine the CTQ's behaviour.*

The notable deviation is Hahn, Hill, Hoerl and Zinkgraf (1999), who describe the Analyse phase as: "Analyze the preliminary data [collected in the Measure phase] to document current performance (baseline process capability), and to begin identifying root causes of defects (i.e. the "X's", or independent variables) and their impact and act accordingly." This description implies that besides the identification of causes, also the establishment of the baseline process capability, as well as the implementation of corrective actions are among the functionalities of the Analyse phase in the view of these authors (they are part of, respectively, the Measure and the Improve phase according to the other authors).

The following definition of the functionality of the Improve phase captures the ideas of most authors:

- Improve phase: *Design and implementation of adjustments to the process to improve the performance of the CTQs.*

All authors mention the *design* of improvement actions as functionality of this phase, but the inclusion of their *implementation* in this phase is not shared by all authors. Finally, the definition of the functionality of the Control phase is:

- Control phase: *Adjustment of the process management and control system in order that improvements are sustainable.*

### Steps

The functionality of each phase describes its goal. The steps that each phase consists of specify intermediate results and actions. An overview of the steps that various authors provide is given in Figure 2.3. The figure is based on the following references:

- Harry (1997), p.21.33 for the Define steps, p.22.2 for the other steps. The numbers 1 through 12 indicate Harry's numbering of steps.
- Breyfogle (1999, pp.18-20). The numbers 1a through 21 indicate Breyfogle's numbering. Not all steps of Breyfogle's stepwise strategy are included. Steps 2 and 4 are omitted, because they are related to the organisational context of Six Sigma.

Steps 14, 15, 17 and 18 are omitted, because they refer to specific tools instead of functional steps.

- Hahn, Doganaksoy and Hoerl (2000).
- Pande, Neuman and Cavanagh (2000). These authors place the DMAIC method and its steps in an encompassing roadmap for implementation of Six Sigma in a company (pp.67-79). As a consequence, many actions have been performed before a DMAIC project starts, and many steps in the Define and Measure phase are reiterations or refinements of these earlier actions. For this reason, Figure 2.3 lists both the steps prescribed in the preliminary steps of the roadmap (in italics and bracketed, based on pp.206-7 and 218) and steps listed under the Define and Measure phase (p.39, but see as well pp.239, 256, 259, 271, 276-281, 337).
- Eckes (2001, pp.44, 50-55, 59, 71-79, 93-109, 131-137, 173, 205).
- Rasis, Gitlow and Popovich (2002)

Figure 2.3 collates stepwise strategies proposed by various authors. Shading indicates the authors' allocation of steps to phases. As much as possible, steps with equivalent functionalities are listed in the same row. Our rational reconstruction of the steps of Six Sigma's method has taken the form of the rightmost column, headed *Generic*. It was formed by extracting for each row the communalities from the steps proposed by the selected authors, and formulating these communalities in a more generic terminology. Figure 2.3 shows that there is considerable agreement among authors about the steps that should be given to project leaders as guidelines for their projects, although most authors omit one or a few steps. Consequently, the generic steps can be considered an adequate reconstruction of Six Sigma's stepwise strategy. Nevertheless, deviations can be noted in the form of omissions, additions, and differences in order. We discuss the most salient ones.

### Omitted steps

Many authors omit one or more steps. Especially about the steps in the Define phase there is less unanimity. In subsequent phases, step M5 (Define objectives) is listed by only half of the authors. Pande, Neuman and Cavanagh as well as Eckes omit the quantification of the relation between influence factors and CTQs (I1). They see the quality problem as a consequence of one or a few root causes. Probably as a consequence of this, the emphasis is less on estimation of a transfer function, but more on the identification of the root cause - once it is tracked down, improvement is seen as straightforward. Step I3 (Conduct pilot test of improvement actions) is listed by only two authors. In the Control phase only Harry; Breyfogle; and Pande, Neuman and Cavanagh propose to assess the capability of the improved process (C1).

| *Phase* | *Harry* | *Breyfogle* | *Eckes* | *Pande, Neuman and Cavangh* | *Rasis, Gitlow and Popovich* | *Hahn, Doganaksoy and Hoerl* | *Generic* |
|---|---|---|---|---|---|---|---|
| *D* | | 7. Create a flowchart/process map. | Create the high-level process map. | *(Identify "core" business processes.*<br>*Create high-level core process map.)* | Map processes. | Define problem to be solved | **D1.** Identify and map relevant processes. |
| | Identify customer. | | Define customers | *(Define process outputs & key customers)* | | | **D2.** Identify targeted stakeholder. |
| | Define needs and specify deliverables.<br>Identify CTQs, map process and link CTQs. | 1a. Identify critical customer requirements from a high level project measurement point of view. | Determine needs of customers.<br>Define requirements on needs. | Identify the problem and define requirements. *(Gather customer data and develop Voice of Customer strategy. Develop performance standards and requirements statements. Analyze and prioritize requirements; evaluate per business strategy.)* | Identify issues or concerns relevant to customers. Identify CTQs. | | **D3.** Determine and prioritize customer needs and requirements. |
| | | 1c. Implement a balanced scorecard considering COPQ and RTY metrics.<br>3. Describe business impact. Address financial measurement issues. | Make a business case. | Set goal. | Prepare a business case. | Determine potential benefits and determine customer impact. | **D4.** Make a business case for the project. |
| *M* | 1. Select CTQ characteristic. | 1b. Identify key process output variables (KPOV's) that will be used for project metrics. | Identify measures of:<br>- input (supplier effectiveness)<br>- process measures (your efficiency)<br>- output measures (your effectiveness) | *(Select what to measure.)* | Study and understand CTQs. | Identify the critical-to-quality characteristics (CTQs) of the product or service. | **M1**. Select one or more CTQs. |
| | 2. Define performance standards. | | Make operational definitions. | *(Develop operational definitions.)* | Develop operational definitions for each CTQ variable. | | **M2.** Determine operational definitions for CTQs and requirements. |
| | 3. Validate measurement system. | 10. Conduct a measurement systems analysis. Consider a variance component analysis. | | *(Test measurement accuracy and value.)* | Perform a GRR study for each CTQ. | Verify measurement capability. | **M3.** Validate measurement system of the CTQs. |

| *Phase* | *Harry* | *Breyfogle* | *Eckes* | *Pande, Neuman and Cavangh* | *Rasis, Gitlow and Popovich* | *Hahn, Doganaksoy and Hoerl* | *Generic* |
|---|---|---|---|---|---|---|---|
| *A* | 4. Establish product capability. | 5. Start compiling project metrics in a time series format. Utilize a sampling frequency that reflects “long-term” variability. Create run charts and control charts of KPOV's.<br>6. Determine "long-term" process capability / performance of KPOV's. Quantify nonconformance proportion. Determine baseline performance. | Baseline sigma level of process and determine variation types. | Validate problem.<br>*(Develop baseline defect measures.)* | Establish baseline capabilities for each CTQ. | Baseline the current defect rate. | **M4**. Assess the current process capability. |
| | 5. Define performance objectives. | | | Refine problem / goal. | | Set goals for improvement. | **M5.** Define objectives. |
| | | | | Measure key steps / inputs. | Determine key measures for upstream suppliers, inputs and processes and collect baseline data for those measures. | | |
| | 6. Identify variation sources. | 8 and 9. Create a fishbone diagram to identify variables that can affect the process output;<br>Create a cause and effect matrix assessing the strength of relationships thought to exist between KPIV's and KPOV's. | Brainstorm all the possible ideas that could explain the Y. | Develop causal hypotheses. | Identify upstream Xs for the CTQs. | Understand the root causes of why defects occur and identify key process variables that cause defects. | **A1.** Identify potential influence factors. |
| | | | | | Operationally define, perform a GRR analysis for and baseline each X. Control the Xs for each CTQ. | | |
| | 7. Screen potential causes. | 11. Rank importance of key process influence factors (KPIV's) using a Pareto chart. | Cull down the large number of ideas to a more manageable number.<br>Reduce the causes down to the vital few. | Identify "vital few" root causes.<br>Validate hypothesis. | Identify the major noise variables for each CTQ.<br>Understand the effect of the Xs on each CTQ.<br>Determine the "vital few" Xs for each CTQ. | | **A2.** Select the vital few influence factors. |
| | | 12. Prepare a focused FMEA. Assess current control plans. | | | | | |

| *Phase* | *Harry* | *Breyfogle* | *Eckes* | *Pande, Neuman and Cavangh* | *Rasis, Gitlow and Popovich* | *Hahn, Doganaksoy and Hoerl* | *Generic* |
|---|---|---|---|---|---|---|---|
| *I* | 8. Discover variable relationship. | 13 and 16. Collect data for assessing the KPIV/KPOV relationships that are thought to exist. | | | Understand the relationship between CTQ's and high risk Xs / major noise variables. | Quantify influences of key process variables on the CTQs. | **I1.** Quantify relationship between Xs and CTQs. |
| | 9. Establish operating tolerances. | 19. Determine optimum operating windows of KPIV's from DOE's and other tools. | Generate and implement solutions that either eliminate the root cause, soften or dampen the effects of the root cause, or neutralize root causation effects. | Develop ideas to remove root causes. | Generate actions needed to implement the optimal levels of vital few Xs that optimize spread, center and shape of CTQs. Develop action plans. | Identify acceptable limits of the key process variables and modify the process to stay within these limits, thereby reducing defect levels in the CTQs. | **I2.** Design actions to modify the process or settings of influence factors in such a way that the CTQs are optimized. |
| | | | | Test solutions. | Conduct pilot tests of actions. | | **I3.** Conduct pilot test of improvement actions. |
| *C* | 10. Validate measurement system (of Xs). | | | | | | |
| | 11. Determine process capability. | 21. Verify process improvements, stability, and capability / performance using demonstration runs. | | Standardize solution / measure results. | | | **C1.** Determine the new process capability. |
| | 12. Implement process controls. | 20. Update control plan. Implement control charts to timely identify special cause excursions of KPIV's. | Implement process controls to hold the gains. | Establish standard measures to maintain performance.<br>Correct problems as needed. | Lock-in improvements by developing, documenting, and implementing process control plans for all high risk Xs and CTQs. | Ensure that the modified process now keeps the key process variables within acceptable limits in order to maintain the gains long term. | **C2.** Implement control plans. |

Figure 2.3 *Rational reconstruction of Six Sigma's stepwise strategy.*

### Added steps

Rasis, Gitlow and Popovich add a step between the Measure and Analyse phase in which key measures for upstream suppliers, inputs and processes are determined and baseline data for those measures are collected. A second addition is a step placed after the identification of possible influence factors in which these are operationally defined, baselined and a measurement system analysis is done. Harry also adds the validation of the measurement system of the Xs as an extra step, but only after the Improve phase. Both additions make sense, in view of the fact that similar actions are done for the CTQs. Because most authors do not include these steps, they were not incorporated in the generic steps. Finally, Breyfogle suggests to assess current control plans at the end of the Analyse phase. It is not abundantly clear to what end one should this.

### Differences in ordering

Breyfogle's step plan is the only one with an order that is very distinctive from the generic steps. At odds with other authors, he places the validation of the measurement system (his step 10; generic step M3) after the identification of influence factors. Moreover the creation of a flowchart or process map (his step 7) takes place between the Measure and Analyse phase. Other accounts place process mapping early in the Define phase (D1).

## Steps and phases combined

Steps provide an operationalisation of the functionality of the phases. This section comments briefly on the consistency of the stated functionality of each phase and the steps that it consists of, also addressing some additional methodological prescriptions that individual authors make.

The steps D1-D4 that the Define phase consists of agree with its functionality. The same holds for the steps that the Measure phase consists of, except that step M5 (Define objectives) is not implied in the phase's functionality. It is preserved in the reconstruction because one could argue that this step comes down to a verification and possible adjustment (based on the assessed current capability) of the business case that was established in the Define phase (step D5). Another anomaly is Harry (1997), who lists his steps 4 and 5 (which correspond to generic steps M4 and M5) under the Analyse phase, which seems at odds with even his own description of the Measure and Analyse phase (p.21.19).

The steps A1 and A2 agree with the stated functionality of the Analyse phase, and a similar conclusion holds for steps I1, I2 and I3 of the Improve phase. Most authors imply that step I2 (Design actions to modify the process or settings of influence factors in such a way that the CTQs are optimised) is based on quantified relations between influence factors and CTQs (so called transfer functions). Together with step I1 (Quantify the relationship between Xs and CTQs) this shows that Six Sigma prescribes that improvement actions should be derived from discovered causal relationships between influence factors and CTQs. In the formulation of step I2 in the corresponding steps of Harry (1997), Breyfogle (1999), and Hahn, Doganaksoy and Hoerl (2000) improvement actions are limited to the design of suitable tolerance limits, but it is questionable whether this restriction is really the authors' intention.

Comparing steps C1 and C2 to the stated functionality of the Control phase, it appears that C1 (Determine the new process capability) does not relate directly to the Control phase's functionality (Adjustment of the process management and control system in order that improvements are sustainable). In view of the fact that C1 is logical in its place, we revise the formulation of the Control phase's functionality:

*Empirical verification of the project's results and adjustment of the process management and control system in order that improvements are sustainable.*

This section is concluded by addressing additional and deviating methodological prescriptions that are raised by various authors.

Pande, Neuman and Cavanagh (2000) place the DMAIC method and its steps in an encompassing roadmap for implementation of Six Sigma in a company (pp.67-79). The five steps are:

1. Identify core processes and key customers.
2. Define customer requirements.
3. Measure current performance.
4. Prioritise, analyse, and implement improvements.
5. Expand and integrate the Six Sigma system.

Steps 1., 2., and 3. are done for the whole company and help select improvement projects. A "voice of the customer system" is built, which measures performance on a wide range of characteristics. The fourth step consists of Six Sigma projects and encompasses the DMAIC phases.

Also Harry (1997) places improvement projects (the inner MAIC loop) in an encompassing roadmap (the outer MAIC loop; see pp.21.18-23). The outer loop, performed by management and technical leaders, encompasses selection and execution of the phases Measure (product benchmarking), Analyse (process baseline analysis), Improve (the improvement projects, following the inner MAIC loop), and Control (audit and review).

Some authors mention an extra methodological rule: improvement and/or analysis has an iterative nature (Pande, Neuman and Cavanagh, 2000, p.239, call this the "back-and-forth nature of process improvement"). This means that several iterations of the Improvement phase might be needed (Hahn, Hill, Hoerl and Zinkgraf, 1999). Along the same lines Harry (1998, p.62) argues that "... it may be necessary to revisit one or more of the preceding phases." One might even have to reconsider the project's initial goals (Pande, Neuman and Cavanagh, 2000, p.239).

## 2.3 Reconstruction of Six Sigma's toolbox

Besides a business context and a strategy, Six Sigma provides a collection of tools. This section gives an overview of tools per DMAIC step. Tools come in various forms, such as models, analysis templates, and procedures. They intend to assist the project leader to obtain intermediate results within steps. This section gives an overview of the tools that are prescribed for each of the DMAIC phases. The following sources are used:

1. Harry (1997, pp.21.37-8; p.22.4-7) (For applications in service quality, Harry lists tools assigned to particular phases. For general projects, tools are listed without reference to particular phases; the assignment to phases below was done by us.);
2. Breyfogle (1999);
3. Hahn, Hill, Hoerl and Zinkgraf (1999);
4. Pande, Neuman and Cavanagh (2000, pp.168; p.181; pp.192-3; p.209; p.212-7; p.218; pp.257-69; pp.277-81; p.343; p.346; p.351; p.356-73) (Most tools are assigned to phases; when the link was absent, the assignment was done by us.);
5. Eckes (2001, pp.52-3; p.73; pp.114-48; p.175; pp.210-12);
6. Hoerl (2001);
7. Rasis, Gitlow, and Popovich (2002).

Upon studying Figure 2.4, one could conclude that Six Sigma's toolkit draws heavily from the field of statistical quality control (SQC, or industrial statistics and quality engineering). One finds virtually all the standard techniques that are described in the standard textbooks in that field, such as Duncan (1986) and Montgomery (1997), except for acceptance sampling, which plays a very modest role (if any at all) in Six Sigma. Besides the statistical SQC tools, Six Sigma's toolkit features the simple problem-solving and process analysis tools whose use was widely promoted by the Japanese (see Ishikawa, 1982): process maps, cause and effect matrix, pareto chart, five why's, et cetera.

The toolbox is supplemented with techniques borrowed from marketing: focus groups, customer interviews, survey studies, and the like (cf. the tools listed under the Define

| *Phase* | *Tool* | *Functionality* |
|---|---|---|
| General | Check sheet[1,2] | Data analysis |
| General | Data collection plan, form, sheet[1,5] | Data analysis |
| General | Bar chart[1,5,6] | Data analysis |
| General | Pie chart[1,5] | Data analysis |
| General | Box plot[2] | Data analysis |
| General | Line chart[1,5,6] | Data analysis |
| General | Histogram[1,2,5,6] | Data analysis |
| General | Sampling[1,2,4,5,6] | Data analysis |
| General | Descriptive statistics[1,2] | Data analysis |
| Define | Process mapping, flowchart, SIPOC model[1,2,3,4,5,6,7] | Identify and map relevant processes |
| Define | Customer interview[4,5] | Determine and prioritize customer needs and requirements |
| Define | Survey[1,4,5] | Determine and prioritize customer needs and requirements |
| Define | Focus group[4,5] | Determine and prioritize customer needs and requirements |
| Define | Customer observation[4,5] | Determine and prioritize customer needs and requirements |
| Define | Customer complaint system[4,5] | Determine and prioritize customer needs and requirements |
| Define | Voice of the customer analysis[7] | Identify concerns important to customers |
| Define | Kano's model[4,7] | Determine and prioritize customer needs and requirements; classification of customer requirements into dissatisfiers, satisfiers, and delighters |
| Define | Quality function deployment[2,3,4,6,7] | Adjust the on-line quality control system; keep track of processed products |
| Define | CTQ tree, tree diagram, CTQ flowdown[1,4,5] | Determine and prioritize customer needs and requirements |
| Define | Affinity diagram[2,4,5,6] | Determine and prioritize customer needs and requirements |
| Define | Interrelationship diagraph[2,6] | Determine and prioritize customer needs and requirements; identification and classification of needs and requirements |
| Measure | Pareto chart[1,2,4,5] | Select one or more CTQs |
| Measure | Failure modes and effects analysis[1,2,6,7] | Select one or more CTQs |
| Measure | Unit, defect and opportunity[1,4,5] | Determine operational definitions for CTQs and requirements |
| Measure | Measurement system analysis, Gauge R&R study[1,2,3,4,6,7] | Validate measurement system of the CTQs |
| Measure | Control chart[1,2,4,5,6,7] | Process capability analysis |
| Measure | Process capability analysis[1,2,3,5,6,7] | Assess the current process capability |
| Measure | Capability index[1,2,5] | Process capability analysis |
| Measure | Probability plot[2,7] | Process capability analysis |
| Measure | Benchmarking[1,2,4,5] | Adjust the on-line quality control system; keep track of processed products |
| Analyse | Cause and effect or fishbone diagram[1,2,4,5] | Identify potential influence factors |
| Analyse | Brainstorming[1,2,4,5] | Identify potential influence factors |
| Analyse | Process map, flowchart[4,5] | Identify potential influence factors |
| Analyse | Value stream map[4,5] | Identify potential influence factors; identify process inefficiencies |
| Analyse | Data mining[7] | Identify potential influence factors |

| ***Phase*** | ***Tool*** | ***Functionality*** |
|---|---|---|
| Analyse | Screening experimental design[2,7] | Identify potential influence factors |
| Analyse | Transmission of variance analysis[2] | Identify potential influence factors |
| Analyse | Five why's[1,5] | Adjust the on-line quality control system; keep track of processed products |
| Analyse | Exploratory data analysis tools[1,2,4,5] | Identify potential influence factors |
| Analyse | Cause and effect matrix[1,2] | Select the vital few influence factors; keep track of influence factors |
| Analyse | Statistical significance tests (chi-square test, t-test, (M)ANOVA, hypothesis testing, confidence intervals, regression analysis)[1,2,3,4,5,6] | Select the vital few influence factors |
| Analyse | Design of experiments[1,2,3,4,5,6,7] | Select the vital few influence factors |
| Analyse | Logical cause analysis[4] | Select the vital few influence factors |
| Analyse | Bootstrapping[2] | Select the vital few influence factors; establishment of confidence intervals on estimates |
| Improve | Statistical model building[1,2,3,4,5,6] | Quantify relationship between influence factors and CTQs |
| Improve | Design and analysis of experiments[1,2,3,4,5,6,7] | Quantify relationship between influence factors and CTQs |
| Improve | Response surface methodology[1,2,3,6] | Quantify relationship between influence factors and CTQs |
| Improve | Tolerance design[1] | Design improvement actions; determination of specification levels for influence factors |
| Improve | Robust design[2] | Design improvement actions |
| Improve | Benchmarking[1,2,4,5] | Design improvement actions |
| Improve | Brainstorming[1,2,4,5] | Design improvement actions |
| Improve | Affinity diagram[2,4,5] | Design improvement actions |
| Improve | Application of Must and Want criteria[5] | Adjust the on-line quality control system; keep track of processed products |
| Control | Statistical significance test[1,2,3,4,5,6] | Determine the new process capability; demonstrate improvement |
| Control | Process capability analysis[1,2,3,5,6] | Determine the new process capability; demonstrate improvement |
| Control | Mistake proofing, Poka Yoke[2,3,4,6,7] | Adjust the on-line quality control system |
| Control | Control plans[3,4,5,6,7] | Adjust the on-line quality control system |
| Control | Process scorecard[4] | Adjust the on-line quality control system |
| Control | Statistical process control[1,2,4,5,6] | Adjust the on-line quality control system |
| Control | Control chart[1,2,4,5,6,7] | Adjust the on-line quality control system |
| Control | Pre-control chart[2] | Adjust the on-line quality control system |
| Control | Gantt chart, schedule[5] | Adjust the on-line quality control system; keep track of processed products |
| Control | Checklist[5] | Adjust the on-line quality control system |
| Control | Audit[5] | Adjust the on-line quality control system |
| Control | Failure modes and effects analysis[1,2,4,6,7] | Adjust the on-line quality control system |
| Control | Risk management[7] | Adjust the on-line quality control system |
| Control | Lean manufacturing[2] | Adjust the on-line quality control system; streamline processes; functionality within Six Sigma not clear |
| Control | Reliability engineering[2] | Adjust the on-line quality control system; functionality within Six Sigma not clear |

Figure 2.4 *Rational reconstruction of Six Sigma's toolbox; notes refer to the numbered sources listed above.*

phase). For some tools, such as reliability engineering and lean manufacturing (listed by Breyfogle, 1999, only), the functionality within Six Sigma is not clear. Lean manufacturing and reliability engineering seem a bit odd in the Six Sigma toolbox, being complete approaches in themselves, rather than tools. The assimilation of lean manufacturing in Six Sigma is discussed in section 2.5.

## 2.4 Discussion of Six Sigma's reconstruction

The reconstruction in this chapter is purely descriptive. That is, it structures the accounts that the Six Sigma literature itself provides, without evaluating them against theoretical frameworks beyond the Six Sigma literature. A partial prescriptive reconstruction is given by, for example, De Mast (2003), which focuses on the stepwise strategy.

The reconstruction that this chapter provides is intended to serve as a basis for scientific studies. We mention three applications of the results of this paper in scientific research:

1. Compare Six Sigma with and position it with respect to other approaches.
2. Study the method's applicability (under what conditions and for what type of problems does the method work?). For example: is the same method suitable for both the manufacturing and service industry?
3. Analyse whether Six Sigma can be integrated with other approaches for quality improvement, such as TQM or Lean.

In the last sections of this chapter part of application (3), the analysis of an integration of Six Sigma and Lean, will be executed.

The main result of the study consists of a structured account of the Six Sigma method, as provided by Figure 2.1, 2.2, 2.3 and 2.4. Furthermore, the reconstruction allows us to draw a number of conclusions about Six Sigma, which characterise the method.

1. Project selection is customer-focused (as opposed to being driven by technology, experts, or perception), and starts from an inventory of customer needs. Typically, the term *customer* here refers to either the end-user (projects focusing on product quality) or the company (projects focusing on process quality). Support for this conclusion is provided by generic steps D2 (Identify targeted stakeholder) and D3 (Determine and prioritise customer needs and requirements), and the inclusion of tools for analysing the voice of the customer (such as customer interviews and focus groups).
2. The method prescribes that problems and issues be parameterised. Problems and issues are translated into the form of variables and requirements, thus providing an unambiguous and operational definition of the problem under study.

Cf. steps M1 (Select one or more CTQs) and M2 (Determine operational definitions for CTQs and requirements).

3. Emphasis is on quantification: variables are preferably numeric, and the magnitude of problems or the effects of influence factors should be quantified. This enables prioritisation and optimisation of interaction effects and trade-offs, as embodied in techniques like the Pareto analysis and response surface methodology.
4. Relationships among variables are modelled: strategic goals (whether customer demands or the company's strategic focal points) are related to CTQs. The CTQs' behaviour in turn is related to influence factors that causally affect it. Thus, improvement actions are based on understanding of relationships among factors and on the discovery of causal mechanisms. Generic step I1 (Quantify relationship between Xs and CTQs), as well as tools such as the CTQ flowdown, quality function deployment, and the many statistical modelling tools like regression analysis all support this conclusion.
5. Ideas are tested to empirical reality. One of Six Sigma's maxims reads "Show me the data". During projects, this means that a data-based problem diagnosis precedes attempts at solving the problem, that the hypothesised effects of influence factors are experimentally studied, and that improvement actions are tested in practice before they are accepted. More in general, one could say that Six Sigma emphasises empirical research and analysis, not as a substitute, but as an indispensable supplement to expert knowledge. See, for instance, steps such as M4 (Assess the current process capability), A2 (Select the vital few influence factors), and I3 (Conduct pilot test of improvement actions), and tools such as the capability analysis, design of experiments, and statistical significance tests.
6. Six Sigma does not offer standard cures, but a method for gaining understanding of the causal mechanisms underlying a problem. Two directions could be discerned in the type of improvements that Six Sigma prescribes. On the one hand is the view put forward by Harry (1997), Breyfogle (1999), Hahn, Doganaksoy and Hoerl (2000), and Rasis, Gitlow and Popovich (2002), who advise the project leader to find a transfer function that quantifies the effect of influence factors onto the CTQ (step I1). Influence factors are described as variables, rather than disturbances or events. Improvement actions exploit the knowledge of this relationship, and could take the form of optimisation of process settings, the economical design of tolerances, or pointed countermeasures against noise variables. On the other hand is the view put forward by Pande, Neuman and Cavanagh (2000) and Eckes (2001), who are less focused on finding a transfer function. Their description of improvement actions is more general, for instance, "remove root causes."

7. Tools and techniques are advanced, considering that they are taught to non-statisticians (compared to, e.g., Ishikawa's (1982) seven tools). But they do in general not reach the level of courses for professional quality engineers or industrial statisticians (see Hoerl, 2001). Tools and techniques are drawn from various disciplines, but especially SQC and marketing.

## 2.5 Lean Six Sigma: the assimilation of Lean Thinking

Six Sigma is characterised in the previous section as similar to good medical practice used since the time of Hippocrates: first relevant information is gathered followed by careful diagnosis. After the diagnosis is completed, a treatment is proposed and implemented. Finally, checks are applied to see if the treatment is effective. Six Sigma offers a structured, analytic and logically sound approach to problem solving. Six Sigma is, however, often perceived as overly complex, which is a weakness. In case of simple problems with obvious and easy-to-implement solutions, the rigorous adherence to the Six Sigma problem-solving process may be considered "overkill" and inefficient (George, 2003). Furthermore, Six Sigma typically does not resort to standard solutions to common problems (see the conclusion (6) of the previous section). Thirdly, the tools and techniques used, such as Design for Experiments, are considered difficult to master by some practitioners.

In this section an alternative approach to quality improvement, Lean Thinking, is introduced. It will be argued that Lean Thinking and Six Sigma have complementary strengths. The combined approach Lean Six Sigma provides an effective framework for producing quality improvements in service and industry.

The proliferation of Lean Thinking or Lean Manufacturing in the Western World was facilitated by the publication of Womack, Jones and Roos (1990). They made a case for Lean Manufacturing by showing that the Japanese manufacturers in automotive outperformed American and Western European manufacturers dramatically. Partly because of this book Lean Manufacturing became generally accepted in manufacturing in the Western world in the eighties and nineties. More recently it is also applied in service environments (Womack and Jones, 2003).

It is not straightforward to characterise Lean (as it is often abbreviated) in a compact and comprehensive way, because it consists of a patchwork of diverse tools and techniques. This diversity and lack of coherence can be traced to Lean's development. It has grown in production processes, focusing on concrete problems. Most production processes suffer from diverse impediments that give rise to inefficiencies. Typical impediments are long changeover times, capacity bottlenecks, and quality defects. Lean consists of a variety of practical, down-to-earth tools to solve or compensate for these impediments. These tools and solutions are highly industry specific (see Zipkin, 1991).

Despite the diversity of the tools and techniques, there is a common denominator in all Lean applications: Lean applications aim to optimise the efficiency of processes (see Wren, 2005, or De Mast and Does, 2006). The typical strategy is to start mapping and modelling processing times, throughput times, and queue times, and mapping redundancies and inefficiencies in processes. After mapping these, standard improvement models are applied to remove redundancies and inefficiencies in order to decrease processing times, throughput times, and queue times. The most important improvement models are the following:

- *Line balancing*. Balancing and fine-tuning the processing capacity of each of the process steps in order to prevent both overcapacity and under-capacity.
- *5S-method*. An approach to make and keep the workspace well-organised and clean. This reduces inefficiencies due to poor organisation.
- *Single Minute Exchange of Dies (SMED) or Rapid changeovers*. Optimizing the utilisation of production resources by reducing downtime of the production resource.
- *Visual management*. Making the work flow and work pace visible to the employees, for instance in the form of dashboards. This provides employees with feedback on their performance and thus helps them to improve their performance.
- *Cellular production*. Collocating process steps and rearranging the workspace to optimise it with respect to efficiency.
- *Pull systems*. A system in which the production or service delivery process only starts after a customer order. The aim is to reduce inventory levels and overproduction.
- *One piece flow*. Processing work items one-by-one instead of as a batch, which helps to reduce inventory levels and throughput time.
- *Critical path analysis*. Analysis of interdependence of process steps with the aim to improve their mutual coordination to reduce the total throughput time of the process.
- *Complexity reduction*. Complexity is the number of different products and services and the number of processes. Complexity reduction reduces these numbers with the aim to improve efficiency.

Details of these and other improvement models and the analysis tools used in the Lean approach can be found in literature (Shingo, 1989; Standard and Davis, 1999).

Lean provides several analysis tools and a number of standard improvement models, but lacks an organisation structure and a stepwise strategy. Six Sigma, on the other hand, offers fewer standard improvement models, but provides a method consisting of phases and steps for problem solving and an organisation structure. The ideal solution is to combine the two approaches. Many practitioners have done so tacitly for quite some time. We now outline an integrated framework for Lean Six Sigma consisting of the following elements:

1. Organisation structure: The organisational infrastructure is based on Six Sigma. This means Lean Six Sigma uses a project organisation consisting of black belts (BBs), green belts (GBs), and champions. Moreover the Lean Six Sigma initiative is managed as a programme, and the project training and training programme are also copied from the Six Sigma approach (see section 1.1).
2. Methodology: The stepwise strategy for projects of Six Sigma is used, containing the DMAIC steps and phases. Lean analysis tools and standard improvement models are embedded in this project approach, which offers an analysis of the project goals (Define and Measure phase), a diagnosis of the current process (Measure phase), and a good anchoring of solutions (Control phase). In chapter 3 we will focus on a analysis model, the CTQ flowdown, that helps to analyse the project goals in the first steps of the Lean Six Sigma method.
3. Tools and Techniques: In Lean Six Sigma the toolboxes of both Six Sigma (see section 2.3) and Lean (see earlier this section) are combined. Lean typically offers simple tools without much mathematical refinement. These tools are easy to apply and effective to solve commonly encountered problems in processes. The tools and techniques are incorporated in the stepwise strategy and help the BBs and GBs to attain intermediate results.
4. Concepts and classifications: The concepts and classifications of both approaches are combined. From Six Sigma terms such as CTQ, influence factors are taken, whereas Lean provides concepts such as takt time, critical path, and waste.

In figure 2.5 the Lean Six Sigma method, including some of its tools, is shown.

## 2.6 Conclusions

1. Six Sigma's methodology is a system of prescriptions; it consists of four classes of elements, namely a description of the type of purposes for which it applies, a stepwise strategy, a collection of tools, and concepts and classifications.
2. A comparison of various descriptions of the method demonstrates that these descriptions have enough communalities to consider them as variations of a single method, and therefore to allow a meaningful reconstruction of their shared essence.
3. Six Sigma's approach to process improvement is heavily based on the theory of empirical inquiry, as well for the method it prescribes (modelling of the causal structure that underlies a problem), as for its approach (empirical study of hypotheses), and for its tools (statistical tools for empirical research).

4. Six Sigma offers procedures for the study and analysis of problems, rather than standard cures.
5. Lean Thinking and Six Sigma have complementary strengths and weaknesses.
6. Synthesizing Lean Thinking and Six Sigma leads to an integrated programme combining the best of both worlds. Lean Six Sigma incorporates the organisation structure and the method of Six Sigma. It merges the Six Sigma and Lean toolbox, concepts and classifications.

# 3 The CTQ flowdown as a conceptual model of project objectives

## 3.1 Introduction

This chapter is about a tool that plays a prominent role in the Define phase of Lean Six Sigma (LSS) projects, namely, the CTQ flowdown. It relates high level strategic focal points to project objectives. In their turn project objectives are linked to, and decomposed into CTQs, which are made operational in the form of measurements. The completed CTQ flowdown provides a conceptual model of project objectives and their rationale.

The CTQ flowdown has evolved in practice without a good underpinning or good description. The first goal of this chapter is to give a thorough account of the technique, providing a useful prescriptive template for users, and an outline of its theoretical background for those interested in deeper understanding.

The second aim of the chapter is to use the developed theory and resulting techniques to arrive at a number of generic models ("templates") for definitions of LSS projects. The chapter only presents the templates applicable to finance, but also in healthcare generic models have been constructed (see Does, Vermaat, De Koning, Bisgaard, and Van den Heuvel, 2006).

The first part of this chapter is based on De Koning and De Mast (2007), the second part on De Koning, De Mast, Does, Vermaat, and Simons (2007).

## 3.2 Lean Six Sigma project definitions

The project selection process results in a definition of the project's objectives. These project definitions come in different levels of precision and completeness. The dimensions on which the project should aim for improvement are sometimes defined in highly tangible and specific form, for example in terms of metrics or performance indicators. In other cases they are framed in more abstract or looser terms, lack an operational form, or are defined from a customer rather than from a process control perspective (cf. functional versus technical requirements in product design). Sometimes the project definition is so vague as to only define the process that should be improved, without indicating on which dimensions improvement will be measured. Besides lack of clarity about the project's objectives, also the rationale and assumptions on which the project definition is based often lack explication.

LSS prescribes that the first step of a project is to explicate the objectives of the project in the form of measurable indicators. The translation of a more or less specific project definition into one or a few measurable indicators (called CTQs) is done in the Define and Measure phase (see section 2.3). A commonly used tool to go from a project definition to these specific and measurable CTQs is the CTQ flowdown. It aims to make explicit and structure the rationale underlying the project. It shows how CTQs relate to higher level concepts such as performance indicators and strategic focal points. Downward it shows how CTQs relate to measurements.

As an example of a CTQ flowdown, we study a process of an insurance company that processes insurance claims of customers. The process is concisely described by the SIPOC chart of Figure 3.1, which specifies the process's inputs and suppliers, as well as its outputs and customers. In addition the main steps of the process are outlined.

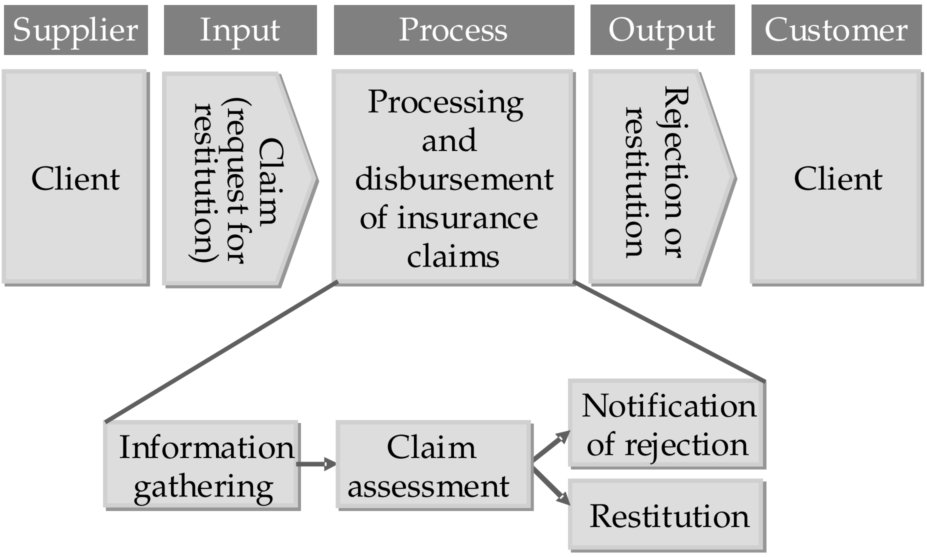

Figure 3.1 *SIPOC chart of processing insurance claims.*

Figure 3.2 shows the CTQ flowdown of a project executed at this company. The company's strategic focal points are customer satisfaction and operational cost. The project objective related to operational cost is reduction of workforce, which amounts to a reduction of the processing time per claim. This one-dimensional CTQ can be decomposed into the constituents cycle times per process step and additional processing time due to complications. Customer satisfaction is translated into the project objective of improving the service quality of the process. Service quality is determined by the total throughput time per claim and the accuracy at which the claim is processed. Throughput time is broken down into total waiting time and total processing time; accuracy is decomposed into complications due to a variety of reasons. The CTQ flowdown is used to describe the sketched relations between strategic focal points, project objectives, one-dimensional CTQs, and their constituents.

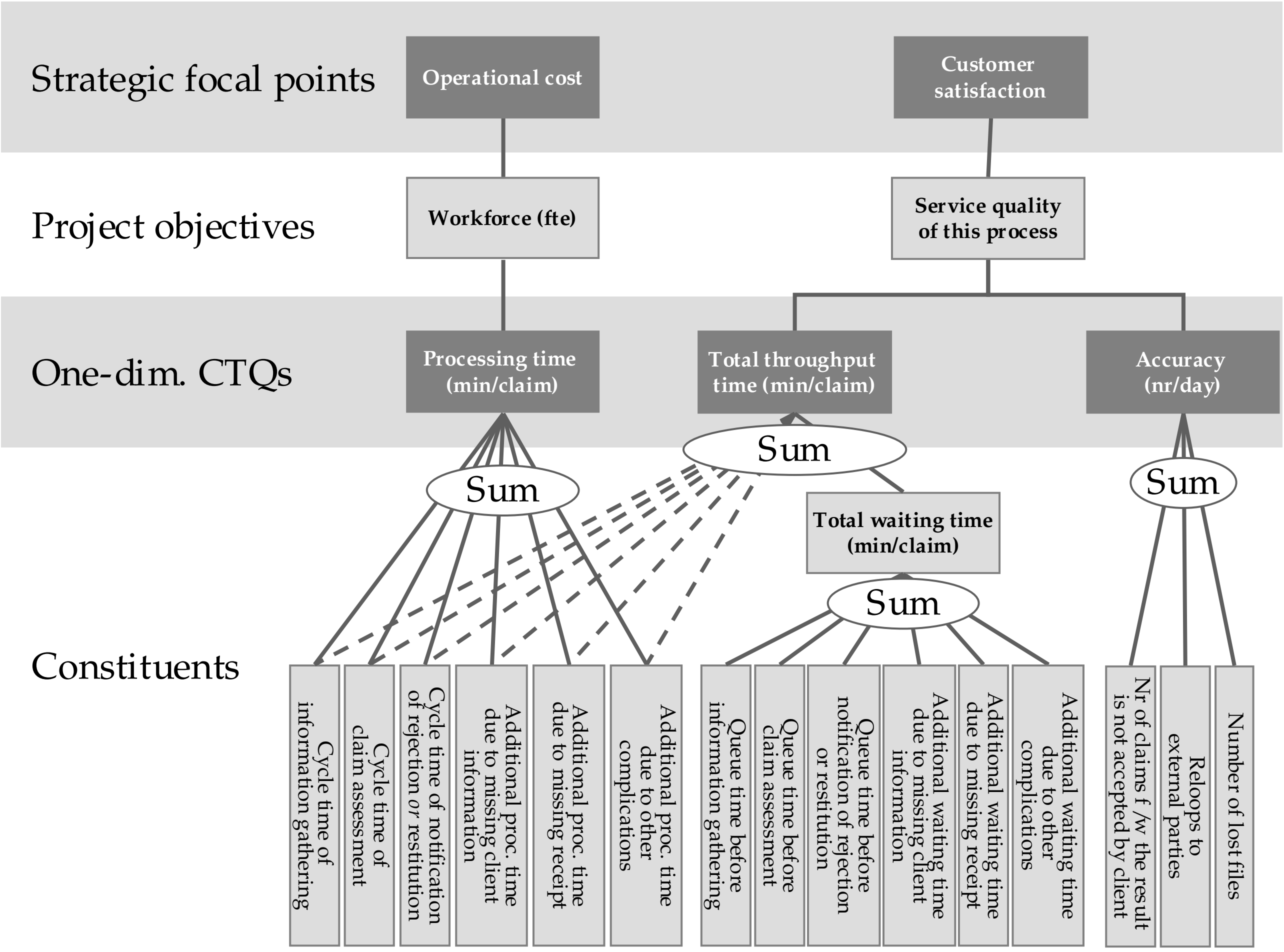

Figure 3.2 *Example of a CTQ flowdown for processing insurance claims.*

The CTQ flowdown is mainly described in training material, for example as part of Six Sigma training curricula. It is commonly used in practice, but like many tools developed through practical usage, it lacks a precise formulation and theoretical grounding. The purpose of this chapter is to provide a clear formulation of the CTQ flowdown by discerning the elements that it consists of and elaborating these into a grammar. Further, the chapter aims to provide theoretical grounding of the CTQ flowdown by linking the elements that it consists of to relevant literature.

The purpose of diagrams such as in Figures 3.1 and 3.2 is that they provide a conceptual model for the context and objectives of a quality improvement project. They structure thinking and communication processes by defining terminology and a frame of reference. A conceptual model is understood to be a network or graph consisting of concepts (nodes) and relationships (linking the nodes) (see Thagard, 1992). The CTQ flowdown combines several forms of conceptual modelling, which can be found in different contexts in the literature. This chapter is built around the notion that CTQ flowdowns generally consist of various canonical layers of nodes and canonical relationships as specified in Figure 3.3.

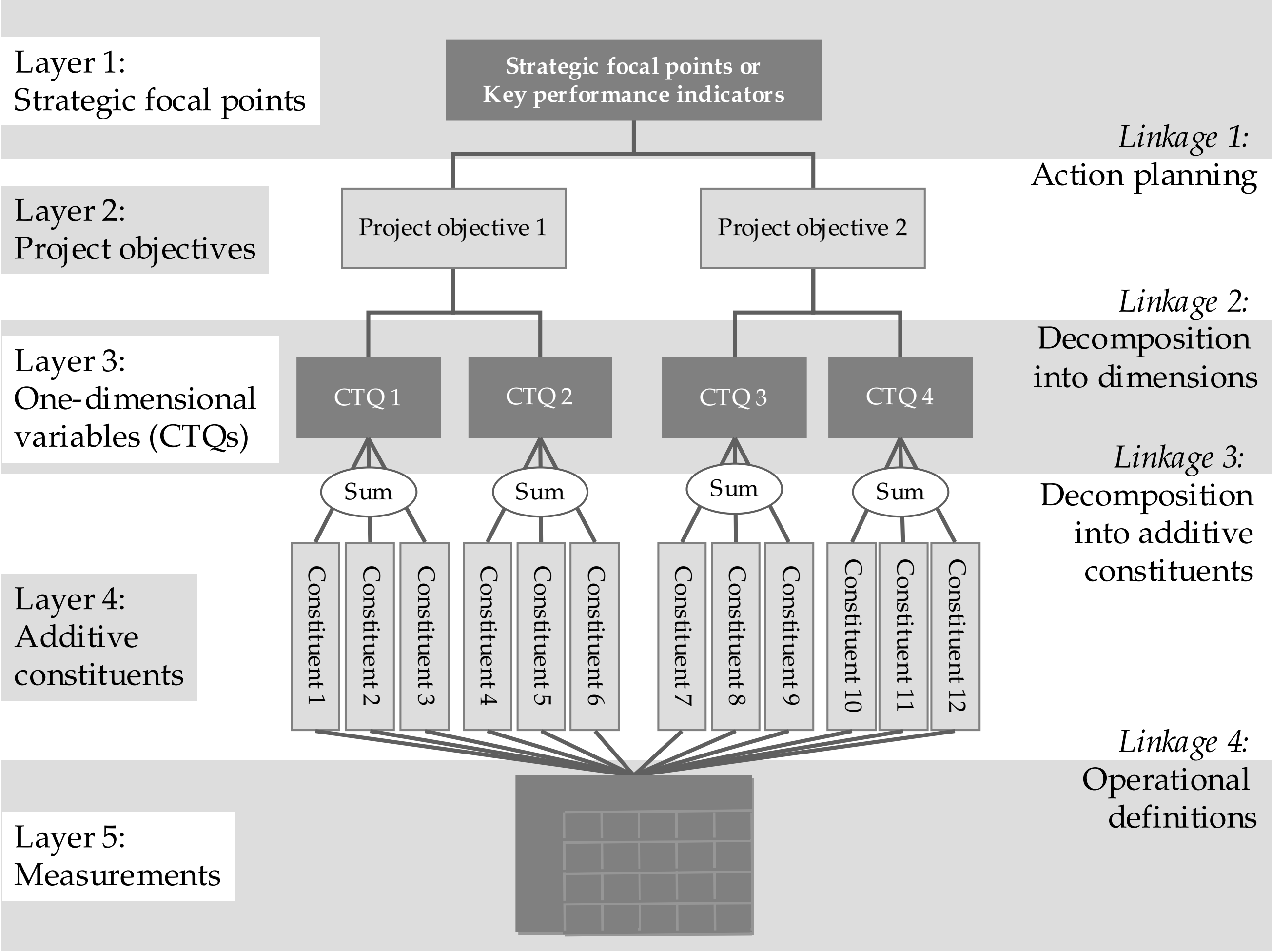

Figure 3.3 *Template of canonical layers of the CTQ flowdown.*

In this chapter the concepts of the canonical layers are defined and the relationships between them will be clarified. Current practice in the use of the CTQ flowdown is taken as a starting point, but is improved upon with the help of scientific literature. The elaboration of the CTQ flowdown will provide a prescriptive template for practitioners, which will help them to effectively deploy and execute quality improvement projects.

The CTQ flowdown is related to other tools which model relationships among concepts in a diagrammatic form. Kaplan and Norton's (2001) Balanced Scorecard links a business strategy to measurable performance indicators in the context of performance control. Cause and effect diagrams (Gano, 2001; Doggett, 2005) relate symptoms to their causes, while Goldratt's Current Reality Tree (Doggett, 2005) models more complex cause and effect structures.

The structure of the chapter's first few subsections follows the template of canonical layers shown in Figure 3.3.

## 3.3 From strategic focal points to measurements

### 3.3.1 Layers 1 and 2: strategic focal points and project objectives

In this subsection we discuss the first canonical layer which consists of strategic focal points (or key performance indicators, or other concepts that specify the objectives on the level of a business), and we show how it is linked to the second layer, the one consisting of project objectives. The link is established by a type of relationship called *action planning*.

Strategic focal points guide and focus action at the level of a business and characterise its strategy. For example, a company pursuing a strategy of cost leadership could try to increase the efficiency of its operations. Improvement of operating efficiency is to this company one of the strategic focal points. To translate strategy into actions, projects are defined. Project objectives delineate an improvement project and serve as a yardstick of project success. A company that wants to improve operating efficiency as a part of its strategy to attain cost leadership can select project objectives such as "reduce the number of defective products" or "increase machine utilisation" depending on the business it operates in. As another example, consider the insurance company introduced in the previous section (see Figures 3.1 and 3.2). Two of the company's strategic focal points are customer satisfaction and operational cost. The project objective related to customer satisfaction is the improvement of the service quality of the insurance claims processing. Similarly, the strategic focal point operational cost can be translated into the project objective of reduction of the workforce.

The translation of strategic focal points to project objectives is called "action planning" by Mintzberg (1994). According to Mintzberg the goal of action planning is before-the-act specification of behavior. In particular strategic focal points are translated into improvement programmes which in turn initiate and coordinate improvement projects. Action planning as first part of a project may seem a bit after-the-fact. After all, the project objectives have already been established in the project selection process. Why then, as a first step of the project itself, should the BB reconsider the selection of objectives? The issue here is that although in some cases project selection has been such a disciplined and structured process that clearly defined project objectives are available and their relations to strategic focal points are explicit, in many cases objectives are not delineated so clearly, and their relationship to strategic focal points is not clearly articulated. In the latter case, the BB has to reconstruct after-the-fact:

- *The rationale of the project*. How do objectives link to the bigger scope and strategic focal points in particular?
- *A precise definition of the objectives of the project*. This gives a clear and articulated account of the "what" and "why" of a project, which helps communicate the motivation for and exact goal of the project.

### 3.3.2 Layers 2 and 3: project objectives and one-dimensional variables

Project objectives could be stated in terms of aggregate concepts. If that is the case, they should be decomposed into their constituting dimensions. This decomposition of a project objective into one-dimensional variables is represented in our structure in the link between the second and third layer.

An example of an aggregated project objective could be the statement that a project seeks to increase the quality of a service. The concept of service quality aggregates a number of aspects, such as fitness for use, timeliness, professionalism and courtesy (Parasuraman, Zeithaml and Berry (1985) distinguish seven "service quality dimensions"). Similarly, a project objective to enhance the ease with which semi-manufactures can be processed in further process steps, could be decomposed into various geometrical dimensions, and perhaps dimensions such as brittleness and crookedness. Finally, service quality of insurance claims processing can be decomposed into the throughput time of processing a claim and the accuracy at which a claim is processed (see Figure 3.2).

By decomposing an aggregate project objective into its composing dimensions, the objective is made more precise. Further, this translation into one-dimensional variables is a first step towards making the project objective measurable. The link between the aggregate concept (the project objective) and the individual dimensions it consists of is called a "part-of" relationship. An aggregate decomposed into its constituting dimensions is called a Cartesian product structure in the theory of semantic networks (Hoare, 1972; Smith and Smith, 1977; Thagard, 1992).

### 3.3.3 Layers 3 and 4: one-dimensional variables and their additive constituents

Often, the one-dimensional variables that the project objective is decomposed into can be viewed of as a sum of lower level constituents. In our structure, the decomposition of one-dimensional variables into their constituents takes place in the transition from the third to the fourth layer.

The insurance company example introduced earlier illustrates this idea. The throughput time of processing a claim is decomposed into the sum of processing time and waiting time. Processing time and waiting time in turn may be decomposed in respectively the cycle times and queue times of the individual process steps and the additional processing and waiting time due to a variety of complications. Likewise, the accuracy can be decomposed into the number of lost files, the number of reloops to external parties, and the number of claims for which the result is not accepted by the client (see Figure 3.2). Notice that the word 'Sum' is used in the CTQ flowdown to indicate that lower level constituents terms sum to the one-dimensional CTQ one level higher. Moreover, note that the link between total throughput time and the processing times is indicated by dashed lines. Processing time will typically be only a minor constituent of throughput time, and therefore, as a first approximation, throughput time links only to waiting times.

By breaking down total sums into their constituents the dominant contributants can be discerned from the trivial many (Pareto principle). This helps to focus the project and reduces the scope.

### 3.3.4 Layer 5: measurements

In this subsection we discuss how CTQs, broken down into their constituents, are linked to the fifth canonical layer, consisting of measurements. The gap between CTQs and the realm of measurements is bridged by operational definitions. Operational definitions make CTQs measurable by specifying a measurement procedure. Choosing a measurement procedure means among other things that one has to choose *per what* one will measure, i.e. per what entity a datum will be collected. If, for instance, yield is the CTQ, one should indicate whether one measures it per day, shift or hour (i.e., is one datum a daily yield, a yield per shift, or hourly yield?). The methodological name for the entity per which measurements are collected is (experimental) unit. The collection of all units for which we aim to make conclusions is called the *population*.

By giving an operational definition for one or a few CTQs, the BB defines a template for data collection (a *measurement plan*). A measurement plan has the structure of a datamatrix. The rows of the datamatrix correspond to units (for each unit there is a CTQ value and thus a row in the datamatrix or dataset) and the columns correspond to CTQs.

As an example we turn to the operational definition of the CTQs "waiting time", and "processing time", defined for the processing of insurance claims (Figures 3.1 and 3.2). The unit for these CTQs is a single claim. All claims together form the population on which the measurements are defined. To measure the waiting time and processing time each claim gets a time stamp when it enters and leaves each process step for the first time. The difference between the start and end time of a process step is the cycle time of a claim; the difference between the end time of one process step and start time of the next is the queue time between these steps. The resulting measurement plan is shown in Figure 3.4. The dataset does not contain the raw data, but data already transformed into CTQ measurements. For instance, for the cycle times this means that the end and start data are simply subtracted, but often this involves more complicated manipulations.

| Unit | Processing time | | | | | | Waiting time | | | | | | Accuracy | | | Attributes | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Case nr | Cycle time of information gathering | Cycle time of claim assessment | Cycle time of notification of rejection *or* restitution | Additional processing time due to missing client information | Additional processing time due missing receipt | Additional processing time due to other complications | Queue time before information gathering | Queue time before claim assessment | Queue time before notification of rejection *or* restitution | Additional waiting time due to missing client information | Additional waiting time due to missing receipt | Additional waiting time due to other complications | Number of lost files | Reloops to external parties | Number claims for which the result is not accepted by the client | Employee-id | Claimed amount | District | Type of claim |
| Case 1 | | | | | | | | | | | | | | | | | | | |
| Case 2 | | | | | | | | | | | | | | | | | | | |
| Case 3 | | | | | | | | | | | | | | | | | | | |
| Case 4 | | | | | | | | | | | | | | | | | | | |
| Case 5 | | | | | | | | | | | | | | | | | | | |
| Case 6 | | | | | | | | | | | | | | | | | | | |
| Case 7 | | | | | | | | | | | | | | | | | | | |

Figure 3.4 *Measurement plan.*

Note that, although CTQ measurements typically are numerical, they can also be categorical. If, for example, we measure whether a product conforms to customer requirements, the measurements can adopt two possible values: conforming or nonconforming. Because numerical measurements contain more information these are preferable over categorical measurements.

Operational definitions were a hallmark of a philosophy called operationism (Bridgman, 1927). Making the meaning of concepts more specific by giving an operational definition makes sure that the statements in which they occur are testable and that they lend themselves to use in explanations and predictions. The extreme stance of operationism, that a concept is synonymous with the corresponding set of operations, is generally regarded untenable (it implies that two different measurement procedures to measure cycle time would define two different concepts), but the importance of operational definitions as the link between concepts and empirical data is generally acknowledged.

## 3.4 Conceptual modelling of the causal structure of problems

A characteristic of LSS is the principle that improving a process requires understanding of how it works. To understand a process means that one is able to relate the behavior of the CTQs to so-called influence factors. This is symbolised by the equation $CTQ = f(X_1, X_2, ..., X_n)$. Without understanding of the mechanics of a process, solutions to the problem will be cosmetic, i.e. one is just fighting symptoms. In LSS the BB typically starts searching for influence factors after the definition phase. The relations between these influence factors and CTQs are shown graphically in diagrams, whose form is much similar to the CTQ flowdown. Instances of such diagrams are the cause-and-effect diagram, and the Current Reality Tree (Gano, 2001; Doggett, 2005).

To illustrate the idea, we can examine the measurement plan of Figure 3.4. Apart from units and measurements (of CTQs) also *attributes*, such as employee-id, and type of claim are included. Although these attributes are not shown in the CTQ flowdown – they are not part of the project definition proper – they are taken along to get as much detail out of the measurements as possible. These attributes are all potential influence factors, so taking them along helps diagnosis and to find the solution later in the project. For instance, if some employees work more efficient than others, this will show up in the data: We will then see differences in cycle times between employees. The solution could possibly be to reduce these differences by adopting the work practices of the best employee.

## 3.5 Application of the CTQ flowdown model to create generic definitions for projects in finance

Even though LSS offers a standardised project approach, some LSS projects fail. As we noted in the first section, an important cause for project failure is not having the project clearly defined (see also Partington, 1996). The project definition should specify the deliverables of the project and indicate the expected benefits resulting from project completion (Lynch, Berolono and Cloutier, 2003). Ideally, the project definition process is guided by an effective system of operational and financial metrics. Especially where such systems are lacking or insufficient, project definitions come in different levels of precision and completeness. In these cases they tend to be vague and lack a business rationale, sometimes even to the extent that the project definition only defines the process that should be improved, without indicating on which dimensions improvement will be measured (see section 3.2). As a result it is not uncommon that as the project develops, BBs and champions have diverging views of what constitutes a successful project and what should be delivered by the BB. Wasted effort, missed deadlines and even preliminary project termination may be the consequence.

The purpose of the remainder of this chapter is to facilitate the process of defining LSS projects in financial services. Our strategy is based on the observation that many projects have similar goals and comparable project definitions, and that the CTQ flowdown provides an effective model for categorising these projects by type. Thus it is possible to provide a number of standard project definitions (‘generic templates’). This classification of LSS project definitions groups projects with a common *project goal*. In other respects, for instance the type of improvement actions, the projects might differ. Project leaders in financial services can use these templates as example and guide in the definition phase of their own projects. This helps them to formulate crystal clear project definitions which have explicitly stated goals and a solid business rationale. Our second concern is to check to what extent these templates form a classification (taxonomy) of LSS projects in finance.

The idea of providing templates forming a classification of LSS project definition categories has been applied earlier to healthcare (Does, Vermaat, De Koning, Bisgaard, and Van den Heuvel, 2006). Based on a large sample of LSS healthcare projects Does et.al. (2006) established six generic categories of LSS project definitions. These categories are distinct in their structure and aimed at improving one or few generic strategic goals.

In the next section of this chapter the research methodology to construct the templates and a possible classification of LSS project definitions will be explained. In the subsequent section we present the resulting project definition categories and the associated standardised templates. Then, we will analyze the validity of the results, pinpoint limitations of the current research, and suggest directions for further research. The final section provides conclusions.

## 3.6 Research methodology to construct generic project definitions

Classification studies are common in disciplines such as medicine and biology. In psychiatry, for instance, one started classifying psychiatric disorders, resulting in the Diagnostic and Statistical Manual (American Psychiatric Association, 2000). The literature does not describe in great detail the process of constructing a taxonomy, but criteria for judging taxonomy once constructed are provided. Chrisman, Hofer and Boulton (1988) mention the following criteria:

- Categories should be *mutually exclusive*. Each project belongs only to one category;
- Categories should be *internally homogeneous*. Projects belonging to one category are more similar to one another than projects belonging to another category;
- Categories should be *collectively exhaustive*. Each project should belong to at least one category;
- Categories should be described clearly.

In this research, which partly aims to classify LSS project definitions into a few categories, these criteria will be used to judge the quality of the constructed taxonomy.

The starting point for the construction of a taxonomy was a dataset consisting of the descriptions of 65 LSS projects carried out in five different financial services institutions. This sample of 65 LSS projects represents a cross section, which varies along key dimensions, such as type of department (back-office, staff, or front-office), type of organisation (both insurance companies and banks), scope (both BB and GB projects), and size (ranging from 20,000 to approximately 3,000,000 euros' worth of savings).

Part of the description of each project was a project definition, which included at least:

- A business case, specifying the business rationale for the project;
- A (macro level) process description;
- The selected CTQs;
- A description of the measurement procedure for each CTQ (operational definition).

Still, the information available per project varied: the descriptions did not have a uniform format, and the terminology used varied widely. Therefore the descriptions of the project definitions were not yet in a format suitable for classification, and had to be structured first. We decided to work with a standard structure to capture the project definitions. This structure consists of two elements, namely the CTQ flowdown and operational definitions per CTQ (both described in previous sections; see Figure 3.3).

For each project the CTQ flowdown and operational definitions were reconstructed. The resulting project definitions were compared and it was judged whether they were similar or not. This resulted in an (initial) classification of projects. From the common denominator of the projects that were considered similar the CTQ flowdowns for each of the LSS project definition categories were constructed. Then, after having constructed the LSS project definition categories, it was verified whether the resulting taxonomy complied with the previously mentioned criteria. Finally, after the CTQ flowdowns per project definition category were finalised, operational definitions were made: It was checked for the projects in each category how the CTQs were operationalised and, from this, generic operational definitions were constructed.

## 3.7 Templates for generic Lean Six Sigma projects in finance

Our analysis resulted in six generic project definitions. Two of the categories focus on reducing operational cost, two of the categories focus on increasing revenue, one category focuses on reducing operational losses, and the last category aims to improve business decision making. In this section we give an in-depth discussion of each of the six generic project definition categories, but first we give an overview of the categories:

1. Decreasing operational cost by improving processing efficiency;
2. Decreasing operational cost by using cheaper channels;
3. Increasing revenue by increasing customer satisfaction;
4. Increasing revenue by servicing more customers;
5. Decreasing operational losses;
6. Improving business decision making;

Since many projects combine in their definition the goals related to the categories 1 and 3, we distinguish a seventh that we label "7: Increasing customer satisfaction and improving processing efficiency".

Based on Figure 3.5 we can see that the category 7 (increasing customer satisfaction and improving processing efficiency) accounts for 31% of all projects, followed by category 3 (improving revenue by increasing customer satisfaction) accounting for 20%, category 1 (decreasing operational cost by improving processing efficiency) accounting for 17%, and category 4 (improving revenue by servicing more customers) accounting for 9%. Cumulatively these four project definition categories account for almost 80% of the projects we encountered. The three smallest categories, category 6 (improving business decision making), category 2 (decreasing operational cost by using cheaper channels) and category 5 (decreasing operational losses) account for another 17%. Only 6% of the projects in the sample could not be classified in the proposed taxonomy (indicated in Figure 3.5 as "other"). Some of them, however, are hybrids, combinations of the other categories. Because these are stand-alone cases, they did not justify adding new categories.

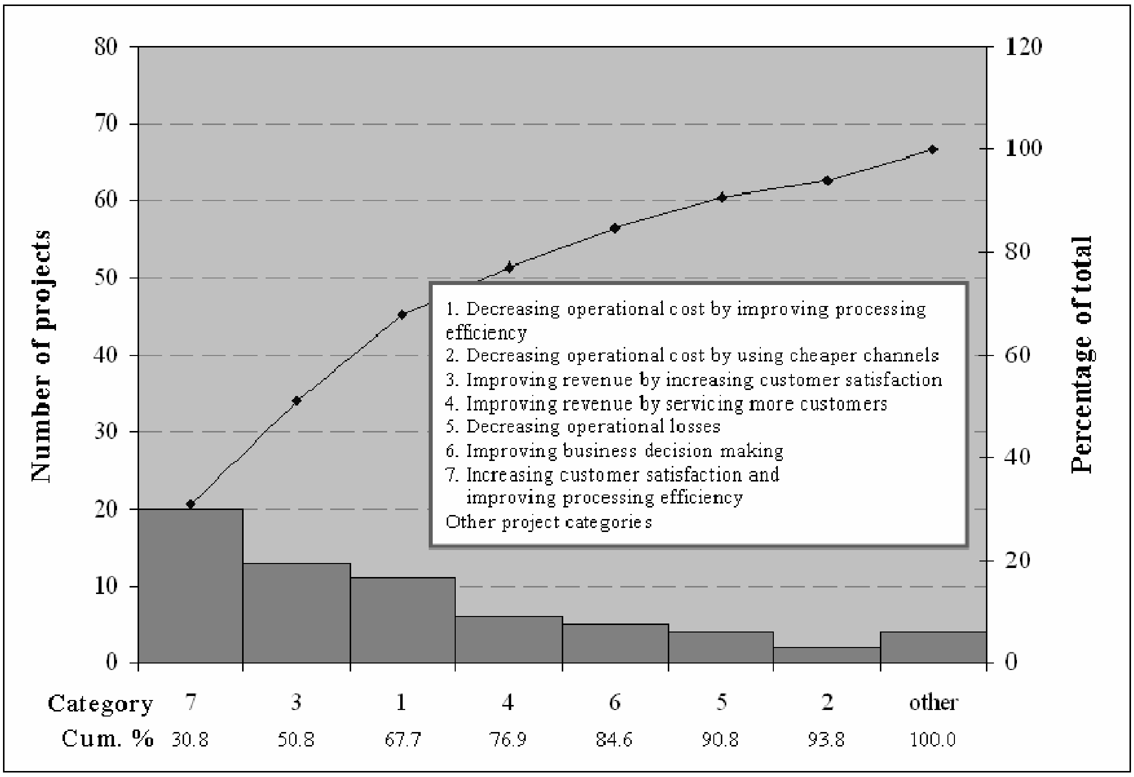

Figure 3.5 *Pareto chart of LSS financial services project definitions.*

As indicated above, our research has shown that the majority of financial services projects can be classified as one of six generic categories. We present the six generic project definition categories in terms of the CTQ flowdown and operational definition of the CTQs. For each generic category we also provide an example.

### Project category 1: Decreasing operational cost by improving processing efficiency

Projects in this category strategically focus on operational cost, which is in a large part driven by personnel cost. Personnel cost itself is determined by headcount (i.e. the number of full time equivalents (FTEs)) and the average cost per FTE. The LSS projects belonging to this category in the sample focused only on headcount (not on average cost per FTE), which in turn is composed of:

- Total processing time per task, which is divided into
  - net processing time (PT) per task, and
  - additional PT per task due to rework;
- Work volume;
- Number of productive hours per employee per day or per week.

In Figure 3.6 these relations are shown and the four CTQs of this category are indicated. Projects in this category have all or some of the CTQs shown in Figure 3.6 as

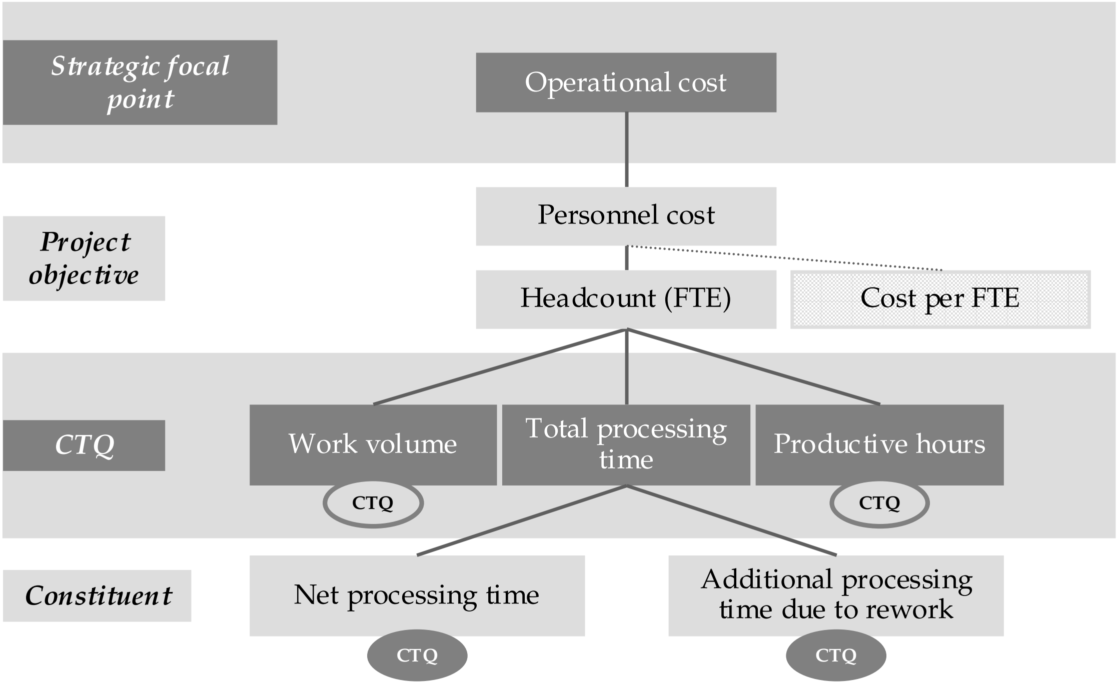

Figure 3.6 *CTQ flowdown for projects decreasing operational cost by improving processing efficiency.*

project goals. The total PT is sometimes split up in the PTs per process step. This provides more information to diagnose the problem in the analyze phase. Rework can originate internally, but also externally. If for instance a client complains or asks for additional information, this typically causes an additional processing loop. In the latter case, an additional benefit of reducing the amount of rework, one of the project aims, may be an increase in customer satisfaction. The operational definitions needed to measure the CTQs are shown in Figure 3.7.

Figure 3.7 shows that the operational definition of a CTQ consists of three elements. First one specifies per which entity the CTQ is measured. This entity is called the experimental unit. Net PT and additional PT due to rework are measured per job (a request, file, complaint, etc.), whereas work volume is typically measured per day or per week. Secondly, one specifies a measurement procedure for the CTQ. PT is commonly measured with the help of time stamps. A travel sheet is attached to a file on which employees can time stamp the start and end time of the processing of the file, in case of rework loops more than once. Alternatively one can use job tracking systems to measure PT or the ‘Day in the life of’ (DILO) method. The DILO method prescribes that employees record on regular time intervals (for instance every ten minutes) during the day the activities they are engaged in. Finally, the operational definition includes a goal for the CTQ. In case of PT and additional PT due to rework we aim to make it as short as possible

| CTQ | Net processing time / Additional processing time due to rework | Work volume | Productive hours |
|---|---|---|---|
| Unit | Per job (request, file, complaint, payment, etc.) | Per day, per week | Per day, per week |
| Measurement procedure | Track a sample of jobs (time stamps), job tracking system | ERP, job tracking system | Time sheets |
| Goal | As short as possible | As little non-value adding work as possible | As close to target as possible |

Figure 3.7 *Operational definitions for projects decreasing operational cost by improving processing efficiency.*

**Example 1**

In a department, which frames offers for loans and processes contracts resulting from these offers, an LSS project was executed aimed at reducing both the net PT and the amount of rework of processing offers and contracts. Analysis revealed that the offers and contracts were processed batch wise, partly because the different process steps were physically far removed from one another. These problems was solved by integrating the process flow and locating the different process steps at one location, making single piece flow feasible. Although the net PT was only reduced from 12 to 10 minutes on average, the percentage of rework loops was improved more drastically, taken down from its original value of 8.0% to 2.4%.

### Project category 2: Decreasing operational cost by using cheaper channels

In financial institutions many processes have been automated to some extent, with the aim to improve the service level, but mostly to reduce cost. In some cases almost all processing is done automatically and the small manual component is just exception handling, while in other cases the automated channel processes only a small percentage of the total work volume. Projects of this category focus on cost reduction by trying to increase the (relative) amount of the work volume that is handled automatically. Shifting work to an automatic channel enables to reduce headcount and hence personnel cost. Processing the work items automatically also imposes costs, but these are almost independent of volume. In sum the total operating cost is reduced. In

Figure 3.8 this line of reasoning is shown, including the resulting CTQ. Note that the typical CTQs of the former category, processing time and number of productive hours, are only taken along as boundary conditions; they should not deteriorate as a result of the project, but their improvement is not included as project objective.

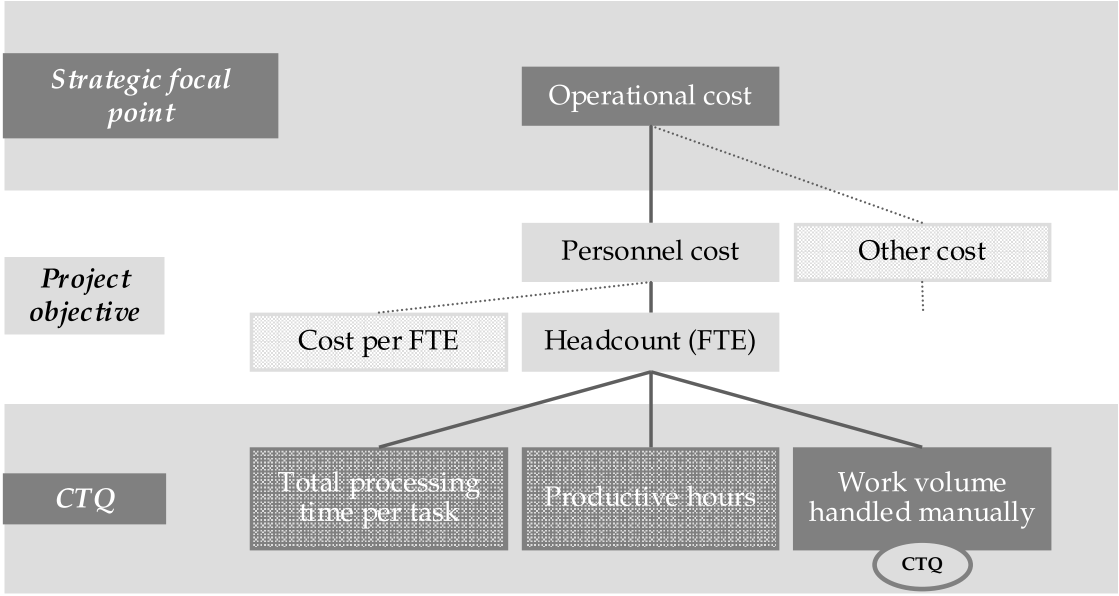

Figure 3.8 *CTQ flowdown for projects decreasing operational cost by using cheaper channels.*

Projects in this category focus either on the total work volume handled manually, or the relative amount of work volume handled manually. The operational definition of this CTQ is shown in Figure 3.9. The information needed for the measurement of the work volume processed manually is typically available in ERP or job tracking systems.

| CTQ | Work volume handled manually (percentage or volume) |
|---|---|
| Unit | Per day, per week, per month |
| Measurement procedure | ERP, job tracking system |
| Goal | As small as possible |

Figure 3.9 *Operational definitions for projects decreasing operational cost by using cheaper channels.*

**Example 2**

Customers of a Dutch bank can ask the bank to provide a copy of a bank current account statement. This can be done in two ways. In some cases customers call the bank and are guided through a menu structure which works fully electronically and results in the customer getting the required copy without human intervention. In other cases customers call the service desk of the bank and get the copy by asking a call center employee. The CTQ of the project was the percentage using the first, fully automated channel. The project succeeded in increasing this percentage from 56% up to 64%, reducing operational cost by 260,000 euros. Partly this was done by making the menu structure more customer-friendly and simple. The menu structure was changed, because analysis showed that about 40% of customers that tried the automatic channel broke off prematurely and ended up calling the call centre.

**Project category 3: Improving revenue by increasing customer satisfaction**

Customer satisfaction is seen as driver of revenue, either because it affects market share, or because it reduces price sensitivity. To this end, projects seek to improve service delivery processes in order to improve service quality. Following the studied projects, service quality can be decomposed into the following underlying dimensions:

- External iterations;
- Throughput time, which can be decomposed into
  - Net waiting time (WT);
  - Additional WT due to (internal and external) rework;
  - Net processing time (PT);
  - Additional PT due to (internal and external) rework;
- Perceived quality.

These relations are shown in Figure 3.10. Projects belonging to this category are aimed at increasing customer satisfaction and the selected CTQs are therefore the ones indicated in Figure 3.10.

To relate net WT, additional WT due to rework, net PT, additional PT due to rework, number of external iterations, and perceived quality to specific measurements, we need operational definitions. They are provided in Figure 3.11.

External iterations and the underlying components of throughput time are quite straightforward to measure. External iterations (or errors) are either measured automatically in case an ERP or other logging system is in place. Otherwise it is best to sample a number of jobs and measure the percentage that contains errors or is iterated. In the case of throughput time sometimes number of rework loops instead of additional processing and waiting time due to rework is measured.

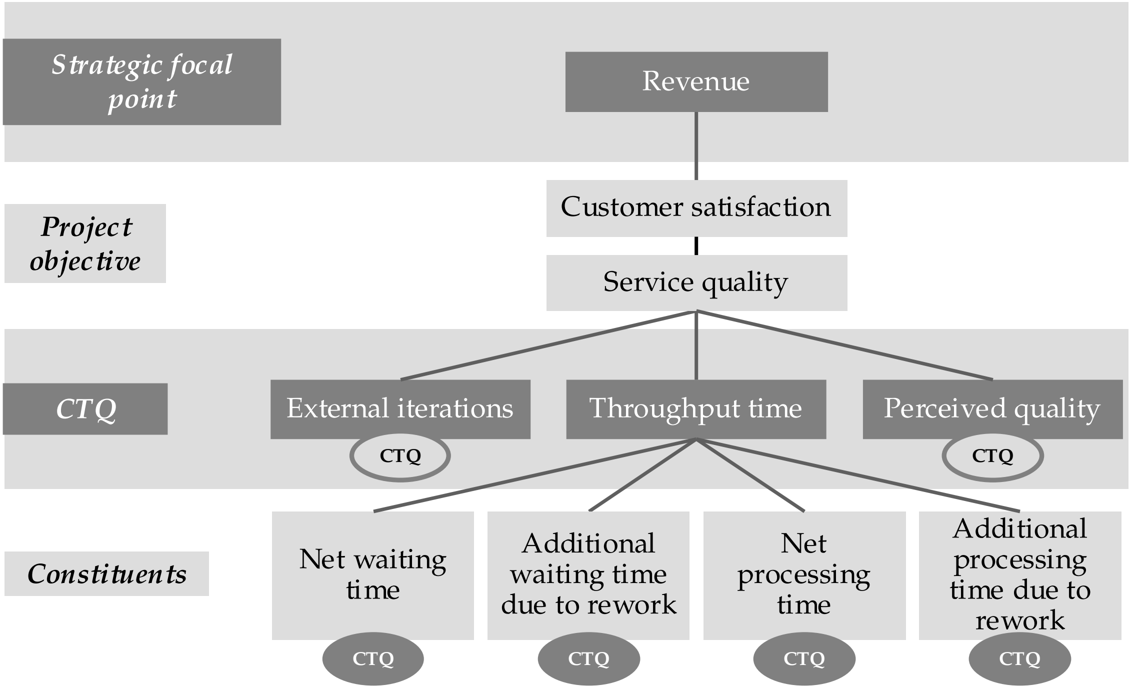

Figure 3.10 *CTQ flowdown for projects improving revenue by increasing customer satisfaction.*

On the measurement of the CTQ “perceived quality” less agreement exists. In some of the LSS projects in the sample “perceived quality” is measured by asking directly to the customer (using a customer survey), whereas in others it is hypothesised to be related to underlying dimensions. These underlying dimensions are highly project specific. To give some examples:

- Completeness of answers to customer questions and the correct routing of service calls (in a call centre of a service desk);
- Clearness of the response to the customer, tone of voice of the agent, and the difference between the client’s expectation and the actual solution offered by the agent in reclaim handling.

We conclude that the details of the measurement procedure for “perceived quality” are hard to capture in a generic operational definition. This view is concurred by the findings of Parasuraman, Zeithaml, and Berry (1985). In an effort to define the concept of service quality, they find that, although there is a set of fairly general dimensions of service quality, it is highly situation specific which dimension is most important. Note that in Figure 3.11 only one possible operational definition of perceived quality is shown.

| CTQ | PT / WT / Additional PT due to rework / Additional WT due to rework | External iterations | Perceived quality |
|---|---|---|---|
| Unit | Per job (a request, file, complaint, payment, etc.) | Per day, per week | Per customer |
| Measurement procedure | Track a sample of jobs (time stamps), job tracking system | Counting based on a sample of jobs, or from an ERP or other logging system | Quality rating based on a survey of customers |
| Goal | As short as possible | As small as possible | As good as possible |

Figure 3.11 *Operational definitions for projects improving revenue by increasing customer satisfaction.*

**Example 3**

In one of the projects in the sample the processing of mutations in life insurance policies was improved. A customer survey showed that customers were particularly dissatisfied with the throughput time of this process, less so with other quality issues, such as errors, and communication. Therefore, total throughput time was chosen as CTQ. The project scope was confined to the two most current products. Measurements showed that the throughput time was longer than the agreed service level (being 45 days) in more than 30% of the cases. Furthermore, analysis showed that waiting time contributed more than 99% to the total throughput time. The solution was to change from batch processing to single piece flow and to apply critical path techniques, changing the sequence of some of the process steps.

### Project category 4: Improving revenue by servicing more customers

Projects in the previous category aim to improve revenue by satisfying customers more. Typically this is done by improving service quality through improvement of back-office processes. Projects in category 4 try to improve revenue as well, albeit differently. The focus here is to sell more products or services by improving sales processes. Improving revenue can be done by either getting more revenue per client or increasing the number of clients. The former can be done by cross selling or asking a higher price. The latter, increasing the number of clients, can be effectuated by:

- Increasing the number of new clients, by

  - identifying more prospect clients, or
  - improving the conversion rate from prospect to contract, which is called the hit or conversion rate.
- Keeping the existing customer base intact by increasing the retention rate.

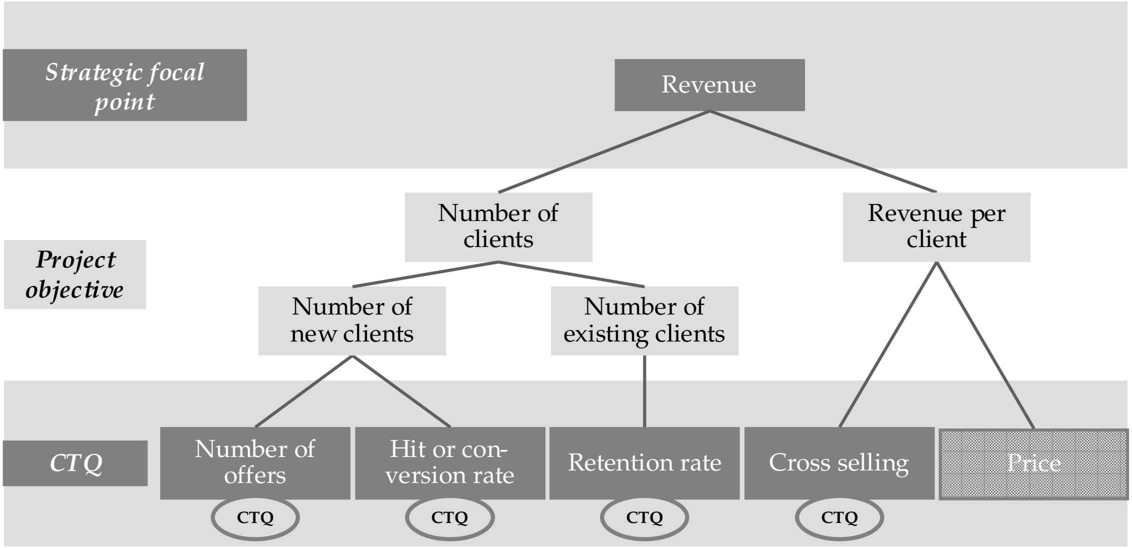

Figure 3.12 *CTQ flowdown for projects improving revenue by servicing more customers.*

These relations and the four CTQs are represented in Figure 3.12. Note that price is not included as CTQ, because its improvement is outside the scope of typical Lean Six Sigma projects.

The operational definitions of these CTQs are shown in Figure 3.13. CRM database is the abbreviation for customer relations management database. In the definition of cross selling the goal reads "as large as possible".

**Example 4**

In one of the projects in this category the goal was to improve the conversion rate from offers for loans to accepted offers for loans. The improvement effort focused on three elements:

1. Making the conversion ratio more visible to employees (visual management);
2. Providing incentives for managers by incorporating the conversion rate for their department in their performance contract;
3. Employees were given more freedom with respect to negotiating loan conditions, giving them more room to operate commercially.

| CTQ | Number of offers | Hit or conversion rate | Retention rate | Cross selling |
|---|---|---|---|---|
| Unit | Per week, per month | Per week, per month | Per week, per month | Per customer |
| Measurement procedure | Via the CRM database | Via the CRM database | Via the CRM database | Via the CRM database |
| Goal | As much as possible | A large as possible | A large as possible | As large as possible |

Figure 3.13 *Operational definitions for projects improving revenue by servicing more customers.*

### Project category 5: Decreasing operational losses

A specific kind of operational cost or lost revenue is called an operational loss. Operational losses can have a large variety of causes, such as fraud, accidents, product flaws, natural disasters, etc.. In financial services institutions an operational loss can for instance be caused by making errors in offers to clients, proposing a lower than intended provision or interest rate. This would result in lower revenue. In other instances, in case of penalties for example, additional costs are incurred. In both cases the total financial impact is determined by:

- The number of operational losses;
- The average value of the operational losses.

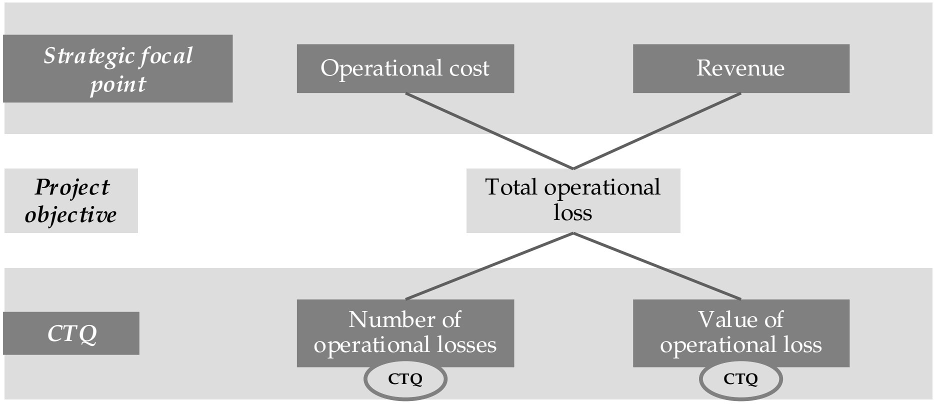

Figure 3.14 *CTQ flowdown for projects decreasing operational losses.*

In the projects in this category these are selected as CTQs, as shown in Figure 3.14. Operational definitions are described in Figure 3.15. The data needed for measurement are extracted from an operational loss database.

| CTQ | Number of operational losses | Value of operational loss |
|---|---|---|
| Unit | Per week, per month | Per operational loss |
| Measurement procedure | Via the operational loss database | Via the operational loss database |
| Goal | As small as possible | As small as possible |

Figure 3.15 *CTQ flowdown for projects decreasing operational losses.*

**Example 5**

In the processing of direct payments, mistakes sometimes lead to an operational loss. In one of the LSS projects these operational losses were measured in number and value. One of the causes for operational losses was a lack of information about the number of operational losses and their causes. To tackle this, several improvement actions were initiated:

1. A feedback loop was installed, providing employees with information on operational losses. This helps the employee not to repeat the same mistake;
2. The management information system in which operational losses are reported was changed. A major improvement was to include an overview of underlying causes for operational losses;
3. Procedure compliance was reviewed. Moreover, a quarterly quality workshop was introduced to raise awareness about operational losses.

**Project category 6: Improving business decision making**

Companies need accountancy reports for legal purposes, but also as input for business decision making, such as decisions about investments, new services to launch, markets to serve, and the like. In order to make the right decisions top management needs information, which is up to date, accurate and relevant. Three principal accountancy statements that serve to provide this information are:

- The income statement (also called profit and loss account, or just P&L);
- The balance sheet, showing assets, liabilities and equity;
- The cash flow statement (showing the starting cash position, and cash flows from operations, investments, and financing activities).

Several projects in the sample dealt with the quality of information of the first of these accountancy statements, the income statement. CTQs taken along are:

- Throughput time of open (credit and debit) book entries, i.e. how long does it take to allocate book entries to a business unit;
- Allocation of (credit and debit) book entries, i.e. what percentage or volume of book entries have been allocated to a business unit;
- Accuracy of allocation of (credit and debit) book entries, i.e. what percentage or volume of book entries have been allocated to the right business unit.

The first two CTQs relate to availability of management information, the last one to the accuracy of the management information. The rationale for these CTQs is shown in Figure 3.16.

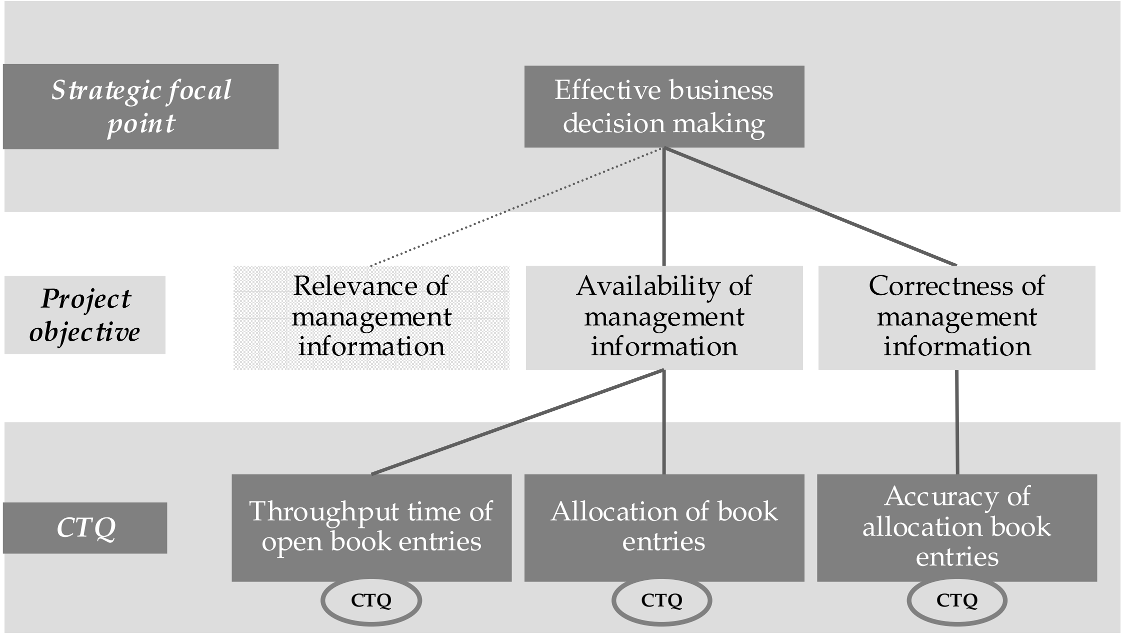

Figure 3.16 *CTQ flowdown for projects improving business decision making.*

The operational definitions of the CTQs are shown in Figure 3.17.

**Example 6**

In one of the projects in this category the goal was to explain the deficit in one of the profit and loss accounts and to allocate so called open entries, i.e. entries that were not

| CTQ | Throughput time of open book entries | Allocation of book entries | Accuracy of allocation of book entries |
|---|---|---|---|
| Unit | Per open book entry | Per week, per month | Per week, per month |
| Measurement procedure | Via the system, database | Via the system, database | Counting based on a sample of book entries |
| Goal | As short as possible | A large as possible | As large as possible |

Figure 3.17 *Operational definitions for projects improving business decision making.*

allocated to one of the business units of the company. A second goal was to prevent the emergence of a deficit in the future. One of the solutions was to implement a different account structure and a different management information system. Moreover, the BB introduced a control tool to monitor debit and credit entries on a daily basis. Consequently the management information is both more accurate and timely. A beneficial spin-off of the project was that in explaining the deficit it turned out that a business partner could be after-charged for approximately 300,000 euros.

**Project category 7: Increasing customer satisfaction and improving processing efficiency**

Some projects are aimed at improving both customer satisfaction and improving processing efficiency. Projects of this category are a combination of the project categories 1 and 3, as can be seen from Figure 3.18.
Evidently, the CTQs are also a combination of the ones of project category 1 and 3. Their operational definitions can be found in Figure 3.7 and Figure 3.11.

**Example 7**

While issuing a new insurance policy, correct and complete information is critically important: If during the pre-processing one finds that information is missing, a request for information is sent out to the client. The processing of the new insurance policy is pending until the required information is retrieved. The key aspects of information requests that drive this process's performance are:

- Additional processing time (PT) due an information request (rework);
- Additional waiting time (WT) due to an information request (rework);
- The number of information requests per application.

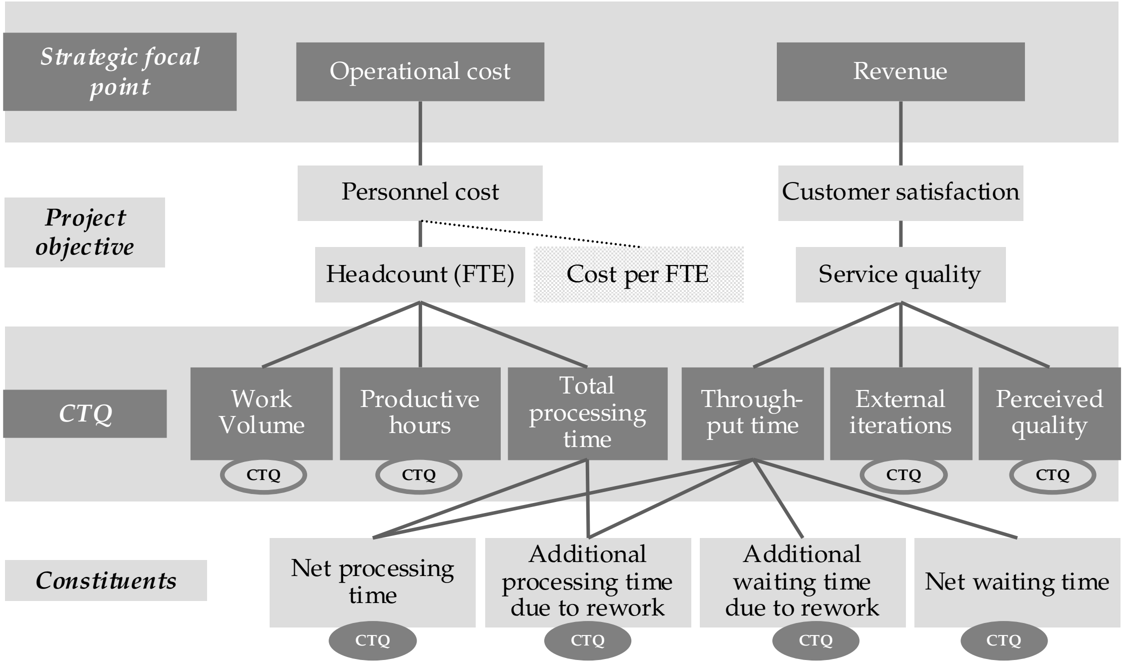

Figure 3.18 *CTQ flowdown for projects increasing customer satisfaction and improving processing efficiency.*

Seeing the additional PT due to an information request as given, the BB pursued reduction of the number of information requests per application, and reduction of the WT per information request. The improvement action was to design a standardised process and to communicate more clearly to the customer what information is needed. The basic principles of this newly designed process are the following:

- The frequency of communications with the customer and the communication channel used for this are standardised;
  - A communication frequency of once in ten days is compulsory. The number of communications with the customer is limited to a maximum of three times. If after three information request the information is still not delivered, no more effort is done to get it;
  - Only written communication with the customer is allowed.
- A standardised template for the written communication with the client is provided to the employees. Use of this template is compulsory.
- A checklist per type of insurance for the required information is provided to the employee.

The average number of information requests per application has dropped from 5.5 to 2.6. This has resulted in estimated savings of 330,000 euros annually. The average WT per information request has increased a little from 3.9 to 4.8 days, because employees are only allowed to send a new information request after 10 days. However the number of information requests has dropped, so the average total WT per application due to information requests has dropped from 21.5 to 12.3 days.

### Remaining projects

Four projects in the sample could not be classified in the existing taxonomy of project definition categories. Two of them strongly resembled category 7 projects, but they have one extra element. In both projects as an additional measure it was checked how much time it takes between doing the actual work and the moment the customer pays. If this takes a long time, the cost of capital is increased, because the lacking money has to be borrowed. The other two projects that could not be classified were actually hybrids, combining categories 7 and 5.

## 3.8 Validity of the classification

One of the most important reasons for LSS project failure is the lack of a clear project definition. In the current chapter we developed a set of generic LSS project definition categories and corresponding standardised templates for projects in finance. As pointed out in section 3.6, the quality of the taxonomy is determined by the extent to which the categories are described clearly, are mutually exclusive, are internally homogeneous, and are collectively exhaustive.

Use of a clear conceptual structure was guaranteed because we used existing structures developed in the first part of the chapter to capture the project definitions. The project definition categories were structured with the help of a CTQ flowdown, including operational definitions for the CTQs. This allowed us to use a unified conceptual framework, containing concepts such as strategic focal point, project objective and CTQ.

The categories are not collectively exhaustive, but the large majority of projects can be allocated to one of the categories. The ones that could not be classified (6% of the projects in our sample) have a hybrid structure, being a combination of other categories. Even with a refined taxonomy it does not seem likely that all LSS projects can be classified.

Finally, the categories 1 to 6 are mutually exclusive. The internal homogeneity is limited to the level of project goals and their rational, and the operationalisation of these goals. Projects within a category differ in the types of analyses that are conducted and the improvement actions in which they result.

The issues above address the study's internal validity. External validity addresses the question to what extent this taxonomy is also useful and applicable outside the sample used. In the case of this study external validity is determined by the extent to which the sample is representative. Judged from this angle the current set-up has some limitations:

- Projects are sampled from one country only, the Netherlands;
- All the project leaders are trained with the same "dialect" of LSS, in the sense that they are all trained by the same institution;
- The sample contains relatively few projects carried out outside a back office environment.

On the other hand:

- The sample represents five companies;
- The sample reflects a good mix of BB and GB projects;
- The projects in the sample varied widely in terms of scope.

Future research, that is needed to refine and corroborate the proposed classification of generic categories, is at least threefold. First of all, cases originating from a larger sample of countries, different training environments, and from different kind of departments within financial services (trading, sales, marketing) could help to test and expand the current taxonomy. Other categories or exceptions may emerge in this way. Moreover, theoretical grounding is needed to see if the strategic focal points covered (by the generic categories) give a reasonably complete enumeration of drivers of business results in the financial services industry.
Thirdly, further research could be directed to study the depth of the generality, i.e. at what level do LSS project definitions within one category start to diverge. Finally, it is interesting to check what the similarities and dissimilarities are between the generic categories of different lines of business.

## 3.9 Conclusions

The definition phase of LSS projects results in project definitions that come in different levels of precision and completeness. Sometimes clarity about the project's objectives lacks, sometimes the rationale and assumptions on which the project definition is based are not clear. It is argued in this chapter that the CTQ flowdown helps to structure the project definition and make explicit the rationale underlying the project. Strategic focal points, project objectives, one dimensional CTQs, and constituent parts of CTQs are placed in a diagram linking them to together. Moreover, these concepts are related to measurements. The elaborate study of the CTQ flowdown allows us to draw a number of conclusions about this tool. Linking concepts has the following purposes:

1. The CTQ flowdown makes explicit the business economic rationale for the project. Project objectives are linked to strategic focal points, which enables programme management and the champion to check the rationale of the project before it is started.
2. The project is placed in its larger scope. This makes explicit the choices as to which aspects will be excluded from the project. These aspects might be suitable topics for later projects, but are excluded from the current project for feasibility reasons. In the example of processing insurance claims (see Figures 3.1 and 3.2) one could, for instance, only focus the project on reducing *throughput time* and leave the improvement of *accuracy* to another project. Placing the project in its larger scope is also crucial for a second reason: focusing. Modelling the larger business economic scope enables one to identify the real drivers of performance, and thus allows the verification that the project is tackling one of the vital few issues (as opposed to one of the trivial many); this also contributes to making explicit the rationale.
3. The CTQ flowdown makes explicit the assumptions on which the project is based. For example: the decomposition of *service quality* into *throughput time* and *accuracy* is based on the assumption that accuracy and throughput time drive service quality (see Figure 3.2). Making this assumption explicit enables a debate or consideration of its legitimacy.
4. The CTQ flowdown serves as a communication tool by providing a common frame of reference and a common language. This ensures that the BB, the champion and programme management all go in the same direction and pursue the same goals.

Making concepts measurable, the second feature of the CTQ flowdown, has the following purposes:

1. Effective problem-solving activities and improvements are based on well defined, crystal clear problem definitions. The first four canonical layers, the ones consisting of strategic focal points, project objectives one dimensional CTQs, and their constituent parts, are all formulated in abstract terminology and too abstract and intangible to base improvement on. Operational definitions make them well defined and crystal clear.
2. Measurements focus the project on the most important (sub-) problems by quantifying the relative magnitude of various aspects of the problem to be solved. This helps to focus on the vital few dominant problems (*Pareto principle*). For example: Only after we have measured the processing times of the several process steps (see Figure 3.2) it will become clear which process step has the largest con-

tribution to the total processing time and, consequently, on which process step we have to focus improvement efforts.

3. Many problems (quality versus cost, speed versus defects) are trade-off problems. The matter, then, is not "either/or", but "how much of one and how much of the other?" Quantification sheds light on the trade-off nature of problems, and makes it possible to solve problems optimally.

The analysis and research on generic LSS project definition categories for finance allows us to draw several conclusions as well:

1. Providing standardised templates for project definitions facilitates making crystal clear project definitions, having a solid business rationale.
2. The majority of LSS projects in financial services can be classified in one of six generic categories.
3. All of these generic project definition categories have a clear rationale from a business point of view. Most are directly related to drivers of operational cost, whereas some are related to revenue and effective business decision making.

Apart from these conclusions some limitations can be pinpointed:

1. The relative size of each category still has to be determined. The sample size and representativeness of the sample are insufficient to determine this precisely enough.
2. The project definition categories need to be validated in other circumstances and contexts as well. The current sample contains projects carried out in the Netherlands, with BBs and GBs trained by one training institution.
3. The project definition categories need to be validated theoretically, to check whether all important strategic focal points relevant to financial services institutions are covered. Literature on management information systems and balanced scorecards provides an interesting angle.

This chapter provides a theoretical grounding of the CTQ flowdown. Although it mainly developed through practical usage, the first sections of this chapter relate elements of the CTQ flowdown to theories about business strategy / action planning, semantic networks / Cartesian product structures, and scientific method / operational definitions (see the references given earlier in this chapter). Apart from this the practical value of the CTQ flowdown model is shown.

# 4 An experimental set-up for destructive gauge R&R assuming patterned object variation

## 4.1 Introduction

Lean Six Sigma (LSS) projects usually depend heavily on reliable data for making conclusions. The reliability of the data, therefore, is an important issue, which is often underexposed in improvement projects, but the importance of which is rightly emphasised in LSS. The reliability of data is the subject of measurement system analysis (MSA; see Van Wieringen, 2003). Besides the validity and calibration of measurements, the precision of a measurement system is an important aspect of MSA. Precision is also referred to as random measurement error or measurement spread. In the engineering sciences it is customary to distinguish two components of measurement spread, namely, repeatability (equipment imprecision) and reproducibility (additional spread due to the way the equipment is operated).

For measurements on a numerical scale, the standard method in the engineering sciences for the assessment of a measurement system's precision, is a gauge repeatability and reproducibility (R&R) study (Burdick, Borror and Montgomery, 2003). In their regular form, they are a standard and important element in the LSS-toolbox (see Figure 2.4). The standard layout of such a study is presented in Table 4.1.

| | Operator | | | | | |
|---|---|---|---|---|---|---|
| Objects | 1 | | 2 | | 3 | |
| 1 | $y_{111}$ | $y_{112}$ | $y_{121}$ | $y_{122}$ | $y_{131}$ | $y_{132}$ |
| 2 | $y_{211}$ | $y_{212}$ | $y_{221}$ | $y_{222}$ | $y_{231}$ | $y_{232}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| 10 | $y_{10,1,1}$ | $y_{10,1,2}$ | $y_{10,2,1}$ | $y_{10,2,2}$ | $y_{10,3,1}$ | $y_{10,3,2}$ |

*Table 4.1: Standard layout of gauge R&R study*

Each object out of a sample of objects is measured multiple times by a number of operators. Variation within rows is measurement spread. We denote the data by $y_{ijk}$, where $i = 1, \ldots, p$ indexes objects, $j = 1, \ldots, q$ indexes operators, and $k = 1, \ldots, r$ indexes replications. The data are modelled as:

$$y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + \epsilon_{ijk}.$$

Here $\mu$ denotes the overall average, $a_i \sim N(0, \sigma_1^2)$ are random object effects, $b_j \sim N(0, \sigma_2^2)$ are random operator effects, and $(ab)_{ij} \sim N(0, \sigma_3^2)$ represent the object-operator interaction. The $\epsilon_{ijk} \sim N(0, \sigma_0^2)$ are the error terms. All $a_i$, $b_j$, $(ab)_{ij}$ and $\epsilon_{ijk}$ are assumed stochastically independent. One is typically interested in the repeatability $\sigma_0^2$, the reproducibility $\sigma_2^2 + \sigma_3^2$, and the total measurement spread $\sigma_m = \sqrt{\sigma_2^2 + \sigma_3^2 + \sigma_0^2}$.
The standard approach exploits replications to estimate measurement spread. For some measurements it is not feasible to obtain replications, for example because objects are destroyed when they are measured, or because the object being measured changes over time. Such measurements are called *destructive*. De Mast and Trip (2005) give a precise definition of the problem of gauge R&R studies for destructive measurements. This problem has been a persistent problem in quality engineering. Although there is no structural solution to it, there is a number of approaches that work in some cases. De Mast and Trip (2005) give an overview of seven such approaches.
One of these approaches works with an experimental layout in which the rows do not contain multiple measurements of the same object, but instead contain measurements of $qr$ different objects. This experimental layout necessarily confounds measurement spread with object-to-object variation. The usual estimators for measurement spread now estimate $\sqrt{\sigma_m^2 + \sigma_1^2}$ instead of $\sigma_m$. If object-to-object variation within rows is not negligible this approach gives an overestimation of the measurement spread. Although this is commonly the case, the approach is still useful, since the bias is on the safe side: if the estimated measurement spread is acceptable, then the true measurement spread is as well. As suggested in De Mast and Trip (2005), this approach could be improved if the object-to-object spread within rows is not just noise, but has a pattern. De Mast and Trip give an example in which part of the object-to-object variation within rows can be attributed to a linear trend. The idea is to fit a model for this systematic part of the within rows object variation and correct the data for it. This approach leads to a smaller overestimation of measurement spread.
The approach outlined above — labelled *patterned object variation* by De Mast and Trip (2005) — requires a more advanced experimental analysis than standard gauge studies. Moreover although a few software packages such as Minitab allow to analyse the sort of statistical model that this set-up requires, they act like a black box, i.e. the estimation procedure is not clear. The research reported in this chapter aims to provide a worked-out set-up and describes in full the analysis procedures needed to estimate

the variance components that constitute the measurement spread.
The remainder of this chapter is organised as follows. Section 4.2 describes the experimental design and the accompanying data model. In section 4.3 the estimation methods, maximum likelihood (ML) and restricted maximum likelihood (REML) estimation, are briefly explained. After this it is discussed how the standard errors of parameter estimators can be computed, using bootstrapping and using asymptotic results. Moreover, it is explained how to replicate the design. Section 4.6 illustrates the procedure by providing an example in which the temperature of a food product is measured. Technical details are gathered in the appendix.

## 4.2 Experimental design and statistical model

The situation taken into consideration is inspired by an example given by De Mast and Trip (2005) in which the strength of biscuits is measured. A slowly increasing pressure is exerted onto the biscuits. The pressure at which they break is the measured strength. This is an example where the patterned object variation approach can be used. The measurements are destructive as the biscuits are destroyed after the measurement. Consequently, the standard gauge R&R setup cannot be executed, since a biscuit can be measured only once. In our approach, replication is replaced with measuring multiple biscuits collected in a sample. The patterned object variation approach assumes that the between biscuits within sample spread is not white noise, but can in part be modelled by a parametric model. In this example, a substantial part of the biscuit-to-biscuit variation can be attributed to fixed differences between positions at the oven belt. In the envisaged set-up, the gauge R&R study consists of taking on $p$ time instants a sample of $qr$ biscuits from $s$ different (fixed) positions in the oven. Each of $q$ operators measures $r$ of these biscuits within each sample. In Table 4.2 this experimental design for six samples ($p = 6$), three operators ($q = 3$), with two objects per operator within a sample ($r = 2$), taken at six ($s = 6$) possible positions, is shown. The design is based on a $6$ x $6$ Latin square design.

| | Operator | | | | | |
|---|---|---|---|---|---|---|
| Sample | 1 | | 2 | | 3 | |
| 1 | 1 | 4 | 3 | 5 | 2 | 6 |
| 2 | 2 | 1 | 5 | 3 | 6 | 4 |
| 3 | 3 | 5 | 4 | 6 | 1 | 2 |
| 4 | 4 | 3 | 6 | 2 | 5 | 1 |
| 5 | 6 | 2 | 1 | 4 | 3 | 5 |
| 6 | 5 | 6 | 2 | 1 | 4 | 3 |

*Table 4.2: Latin square design* (entries indicate oven belt positions $k \in \{1, \ldots, 6\}$)

The measurement on the $l$th biscuit in sample $i$ by operator $j$ is denoted by $y_{ijl}$. We consider the following model:

$$y_{ijl} = \mu + a_i + b_j + (ab)_{ij} + \gamma_{k(ijl)} + \epsilon_{ijl}. \tag{4.1}$$

The fixed differences among positions $k = 1, 2, ..., 6$ are modelled in the $\gamma_{k(ijl)}$. (Notice that, given the design, the indices $i, j$, and $l$ fully determine the position index $k$.) The advantage of this design and model is that the systematic part of the between-objects-within-sample variation is attributed to these $\gamma_{k(ijl)}$. Had we not included this term in the model, the object-to-object variation within a sample would have been completely confounded with the measurement variation. In the proposed approach, only the random part of between-objects variation is confounded with the measurement variation, resulting in a smaller overestimation of the measurement variance. We assume that the sequences $\{a_i\}$, $\{b_j\}$, $\{(ab)_{ij}\}$, and $\{\epsilon_{ijl}\}$ are mutually independent and normally distributed:

$$\begin{array}{ccc} \{a_i\} & \text{i.i.d.} & N(0, \sigma_1^2), \\ \{b_j\} & \text{i.i.d.} & N(0, \sigma_2^2), \\ \{(ab)_{ij}\} & \text{i.i.d.} & N(0, \sigma_3^2), \\ \{\epsilon_{ijl}\} & \text{i.i.d.} & N(0, \sigma_0^2), \end{array} \tag{4.2}$$

(in which i.i.d. denotes independent and identically distributed). To enable maximum likelihood estimation we derive the joint distribution of the observations. Since equation (4.1) defines a mixed linear model we can write it as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{u}. \tag{4.3}$$

Here $\mathbf{y}$ represents the vector of all observations, which is

$$\mathbf{y} = (y_{111}, y_{112}, \ldots, y_{11r}, y_{121}, \ldots, y_{12r}, \ldots, y_{1q1}, \ldots, y_{1qr}, y_{211}, \ldots, y_{pqr})'.$$

The matrix $\mathbf{X}$ is the $pqr \times (qr + 1)$ $(= 36 \times 7$ in this case) design matrix associated with the fixed effects. Its first column is associated with $\mu$, while the six remaining columns represent the position in the oven belt that each biscuit was taken from. Based on the design specified in Table 4.2, $\mathbf{X}$ has the following form:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

The vector $\boldsymbol{\gamma} = (\mu, \gamma_1, \gamma_2, \ldots, \gamma_s)'$ contains the fixed effects. To ensure identifiability we assume that $\sum_{i=1}^{s} \gamma_i = 0$. Furthermore, $\mathbf{u}$ is the vector of random effects and has the form $(\mathbf{u_0}, \mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3})'$, with:

$$\mathbf{u_0} = (\epsilon_{111}, \epsilon_{112}, \ldots, \epsilon_{11r}, \epsilon_{121}, \ldots, \epsilon_{12r}, \ldots \epsilon_{1q1}, \ldots, \epsilon_{1qr}, \epsilon_{211}, \ldots, \epsilon_{pqr}),$$

$$\mathbf{u_1} = (a_1, a_2, \ldots, a_p),$$

$$\mathbf{u_2} = (b_1, b_2, \ldots, b_q),$$

and

$$\mathbf{u_3} = (ab_{11}, ab_{12}, \ldots, ab_{1q}, ab_{21}, \ldots, ab_{pq}).$$

The matrix $\mathbf{Z}$ is the design matrix associated with the random effects. We can partition $\mathbf{Z}$ in a similar fashion as the vector $\mathbf{u}$ as:

$$\mathbf{Z} = [\mathbf{Z_0 Z_1 Z_2 Z_3}] \quad \text{such that} \quad \mathbf{Zu} = \mathbf{Z_0 u_0} + \mathbf{Z_1 u_1} + \mathbf{Z_2 u_2} + \mathbf{Z_3 u_3}.$$

If we denote the Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ by $\mathbf{A} \otimes \mathbf{B}$ then the $\mathbf{Z}$-matrices take on the following form.

$$\begin{aligned} \mathbf{Z_0} &= \mathbf{I}_{pqr} \\ \mathbf{Z_1} &= \mathbf{I}_p \otimes \{_c\mathbf{I}_{qr}\} \\ \mathbf{Z_2} &= \{_c\mathbf{I}_p\} \otimes \mathbf{I}_q \otimes \{_c\mathbf{I}_r\} \\ \mathbf{Z_3} &= \mathbf{I}_{pq} \otimes \{_c\mathbf{I}_r\} \end{aligned} \tag{4.4}$$

in which $\mathbf{I}_v$ is the identity matrix of dimension $v$, and $\{_c\mathbf{I}_w\}$ is a column vector of dimension $w$ containing only ones.

By equation (4.2) the vector $\mathbf{y}$ is multivariate normally distributed. Its covariance matrix $\mathbf{V}$ is as follows

$$\begin{aligned} \mathbf{V} &= \mathbf{Z}(\text{Cov}(\mathbf{u}))\mathbf{Z}' = \sum_{i=0}^{3} \sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i' \\ &= \mathbf{I}_p \otimes \{\sigma_1^2(\mathbf{J}_q \otimes \mathbf{J}_r) + \mathbf{I}_q \otimes (\sigma_3^2 \mathbf{J}_r + \sigma_0^2 \mathbf{I}_r)\} + \sigma_2^2 \mathbf{J}_p \otimes \mathbf{I}_q \otimes \mathbf{J}_r, \end{aligned}$$

where $\mathbf{J}_n$ is a $n$ x $n$ matrix of all ones. The matrix formulation of the covariance structure is needed for the computational algorithm in the next section. For the interested reader we also show a more intuitively appealing version of the covariance structure:

$$\mathrm{Cov}(y_{ijl}, y_{i'j'l'}) = \begin{cases} \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_0^2 & i = i', j = j', l = l' \\ \sigma_1^2 + \sigma_2^2 + \sigma_3^2 & i = i', j = j', l \neq l' \\ \sigma_1^2 & i = i', j \neq j' \\ \sigma_2^2 & i \neq i', j = j' \\ 0 & i \neq i', j \neq j'. \end{cases}$$

In the remaining sections denote the parameter vector by $\boldsymbol{\theta}$ where $\boldsymbol{\theta} = (\boldsymbol{\gamma}, \boldsymbol{\sigma}^2)$.

## 4.3 Estimation

The traditional way to estimate variance components is the ANOVA method. Unfortunately, ANOVA has some well established drawbacks, especially in the case of unbalanced data. The most important are (see Searle, Casella, and McCulloch, 1992, pp.35–39):

1. The possibility of negative estimates for variance components, which are not realistic from a practical viewpoint.
2. The lack of uniqueness of the choice of sums of squares. For unbalanced data one can use for instance Henderson's methods I, II, and III, which are all using different sets of sums of squares. On top of this, criteria for deciding which choice for sums of squares is optimal are lacking. Therefore the choice for a particular set of sum of squares is arbitrary.

Maximum likelihood (ML) estimation is a viable alternative for estimating variance components. Negative estimates are impossible when using ML estimation and the problem of the arbitrary nature of ANOVA is solved as well. An additional benefit of ML estimation is that the resulting estimators are asymptotically efficient. Detailed explanation of its application to variance component estimation can be found in Searle, Casella, and McCulloch (1992). The ML estimator is defined as the maximiser of the log-likelihood, which is given by:

$$l(\boldsymbol{\gamma}, \boldsymbol{\sigma}^2) = -\frac{1}{2}N\log 2\pi - \frac{1}{2}N\log|\mathbf{V}(\boldsymbol{\sigma}^2)| - \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\gamma})'\mathbf{V}(\boldsymbol{\sigma}^2)^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\gamma}),$$

where $N = pqr$ is the total number of observations. A slight drawback of ML estimation for mixed models (such as the model under study) is that estimators for variance components are dependent of the fixed effects. This problem is circumvented by restricted maximum likelihood (REML) estimation, where the fixed parameters are treated as nuisance parameters (see McCullagh and Nelder, 1989, chapter 8). To get rid of the fixed effects, REML applies the ML procedure to a linear combination $\mathbf{K}'\mathbf{y}$ of the data instead of the data $\mathbf{y}$ themselves. The matrix $\mathbf{K}$ is chosen such that:

$$\mathbf{K}'\mathbf{X} = \mathbf{0}, \tag{4.5}$$

in which $\mathbf{X}$ is the design matrix associated with the fixed effects. Multiplying equation (4.3) to the left by $\mathbf{K}'$ then results in:

$$\mathbf{K}'\mathbf{y} = \mathbf{K}'\mathbf{Z}\mathbf{u}.$$

Evidently, this equation does not contain the fixed effects anymore. Therefore the estimators of the random effects are made independent of the fixed effects. As a trade-off, it is no longer possible to estimate the fixed effects. In this chapter both ML and REML are used. We construct $\mathbf{K}$ as follows. Let $(\mathbf{X}'\mathbf{X})^-$ denote a generalised inverse of $(\mathbf{X}'\mathbf{X})$, which is defined as any matrix $\mathbf{G}$, that satisfies $\mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X} = \mathbf{X}'\mathbf{X}$. Defining the matrix M by:

$$\mathbf{M} = \mathbf{I}_{pqr} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}',$$

we get $\mathbf{M}\mathbf{X} = \mathbf{0}$ and $\mathbf{M}$ has rank $(p-1)qr$. It is customary to delete the last $qr$ rows of $\mathbf{M}$, so that the remaining matrix, denoted by $\mathbf{K}'$, is nonsingular.

The estimators can be computed by the EM-algorithm. Both the implementation details and a proof of the convergence of this algorithm are in the appendix.

## 4.4 Standard errors of the estimators

### 4.4.1 Standard errors of parameter estimators by bootstrapping

In order to determine the standard errors of the ML estimators we use a parametric bootstrap procedure. This procedure runs as follows:

- Estimate the model parameter vector $\boldsymbol{\theta}$, denoted by $\tilde{\boldsymbol{\theta}}$ $(= (\tilde{\gamma}, \tilde{\boldsymbol{\sigma}^2}))$, by using the EM algorithm;
- Take a large number $B$ and let index $m$ run from $1$ to $B$;
- Generate $B$ realisations $\mathbf{y}_m$ of the model, using a multivariate normal distribution $N_N(\tilde{\gamma}, \mathbf{V}(\tilde{\boldsymbol{\sigma}^2}))$;
- Estimate $\tilde{\boldsymbol{\theta}}_m$, based on $\mathbf{y}_m$.

This generates $B$ estimates $\tilde{\boldsymbol{\theta}}_m$. An estimator of the standard error per component of parameter vector $\tilde{\boldsymbol{\theta}}$ can be obtained by computing the standard deviation over the set of $B$ estimates.

Determining the standard errors of the REML estimators goes analogously. Since the fixed effects do not play a role whatsoever, the bootstrap samples can be generated from $N_N(\mathbf{0}, \mathbf{V}(\tilde{\boldsymbol{\sigma}^2}))$.

### 4.4.2 Standard errors of parameter estimators - asymptotic approach

One of the attractive features of ML estimators is that the large sample, or asymptotic (as $N = pqr \to \infty$) covariance matrix of the estimators is always available. It is the inverse of the Fisher information matrix $I(\gamma, \boldsymbol{\sigma})$:

$$\text{Cov}(\tilde{\gamma}, \tilde{\boldsymbol{\sigma}}) \backsimeq I^{-1}(\gamma, \boldsymbol{\sigma}).$$

Searle, Casella, and McCulloch (1992, pp.239–240) provide explicit expressions for $\text{Cov}(\tilde{\gamma})$ and $\text{Cov}(\tilde{\boldsymbol{\sigma}}^2)$, which are:

$$\text{Cov}(\tilde{\gamma}) \backsimeq (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$$

and

$$\text{Cov}(\tilde{\boldsymbol{\sigma}}^2) \backsimeq 2[\{{}_m\text{tr}(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j')\}_{i,j=0}^{3}]^{-1}. \tag{4.6}$$

Here the $\{{}_m(a)_{ij}\}_{i,j=0}^{3}$ notation is used to denote a matrix with elements $a_{ij}$, in which both indices run from $0$ up to $3$. For the REML estimates one has similar equations to estimate the variance (see Searle, Casella, and McCulloch, 1992, pp.252–253). Explicitly they read:

$$\text{Cov}(\tilde{\boldsymbol{\sigma}}^2_{REML}) \backsimeq 2[\{{}_m\text{tr}(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i')\}_{i,j=0}^{3}]^{-1}. \tag{4.7}$$

The asymptotic covariance matrix depends on the "true" parameters $\boldsymbol{\sigma}^2$. Since these are unknown we can estimate the asymptotic covariance by replacing $\mathbf{V}$ by $\tilde{\mathbf{V}} = \mathbf{V}(\widetilde{\boldsymbol{\sigma}^2})$ and $\mathbf{P}$ by $\tilde{\mathbf{P}} = \mathbf{P}(\widetilde{\boldsymbol{\sigma}^2})$ in (4.6) and (4.7) respectively.

## 4.5 Replication of the experimental design

In some cases the standard errors of the parameter estimates are too large for the estimators to be useful in judging the precision of the measurement system. In that case it is possible to replicate the design depicted in Table 4.2. This can be done in two possible ways (see Montgomery (2005), pp.140–141):

1. Adding a similar Latin square by doubling the number of samples ($p \to 2p$), but keeping the number of operators ($q \to q$), number of objects per sample ($r \to r$) and fixed positions ($s \to s$) the same.
2. Adding a similar Latin square by doubling the number of samples ($p \to 2p$) and the number of operators ($q \to 2q$) at the same time, but keeping the number of objects per sample ($r \to r$) and fixed positions ($s \to s$) the same. The first $p$ samples are measured by the first $q$ operators; the last $p$ samples are measured by the last $q$ operators.

The estimation procedure remains the same, as well as the implementation scheme (using the EM algorithm). Moreover, the estimation of sampling error does not change either. What does change are the elements of the statistical model shown in equation (4.1). For instance in the replicated design the length of data vector $\mathbf{y}$ doubles. Other elements of the statistical model also change, which is covered in detail in the appendix.

In order to allow more freedom in the set-up of the experimental design one can also use incomplete Latin squares or otherwise modified Latin squares (with added or deleted rows or columns) (see Cochran and Cox (1950, chapter 10)). The analysis is analogous, although here the elements of the statistical model shown in equation (1) change a little as well.

## 4.6 Illustration: Measuring the core temperature of a food product

### 4.6.1 Set-up of the experiment

A certain food product is baked until its core reaches a temperature of about $80°$C. The core temperature is measured by inserting a digital thermometer into the product. Because heat is not distributed perfectly homogeneously over the product, and the operators insert the thermometer by feel (aiming for the core), it is likely that the random measurement error is substantial. Further, the operators use different types of thermometers, thus adding to the variability.

To estimate random measurement error we could not do a standard gauge R&R study. Each product could be measured multiple times, but since the product cools down quite rapidly (about $1.0°$C per minute) these repeated measurements would confound measurement spread with variation in the product's true core temperature. We used the theory presented in this chapter to set-up the following experiment.

We selected six specimens of the food product. Each specimen was to be measured twice by each of three operators, according to the design in Table 4.2. The two times three measurements were to be done with 60 seconds in between successive measurements. We assigned a type of thermometer to each operator, thus confounding operator-to-operator variation and between-thermometer variation.

### 4.6.2 Data

The results of the experiment are shown in Table 4.3. Entries are measured core temperatures (°C). Numbers between brackets indicate at which time (seconds) the food products are measured.

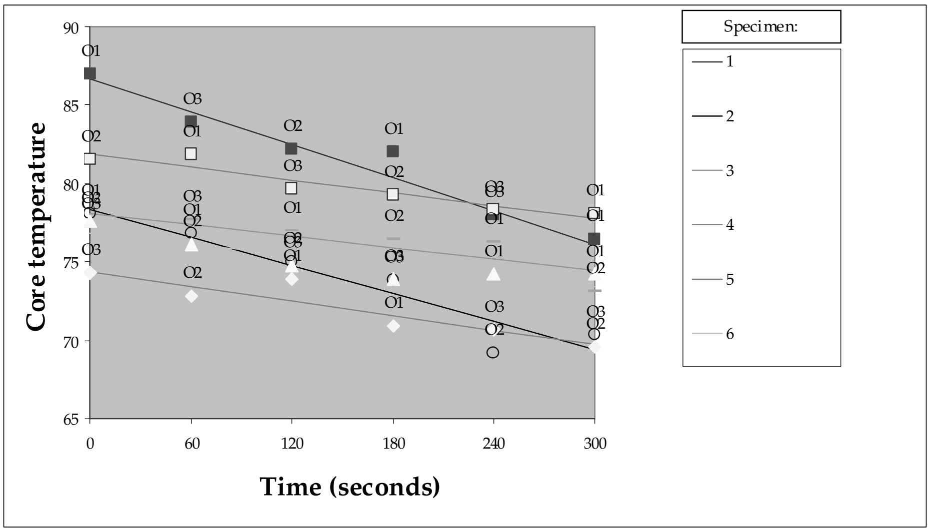

Figure 4.1 *Scatter plot showing core temperature versus time.*

| | Operator | | | | | |
|---|---|---|---|---|---|---|
| Specimen | 1 | | 2 | | 3 | |
| 1 | 87.0(0) | 82.0(180) | 82.2(120) | 77.0(240) | 83.9(60) | 76.4(300) |
| 2 | 76.8(60) | 78.1(0) | 69.2(240) | 75.0(120) | 70.4(300) | 73.8(180) |
| 3 | 77.0(120) | 76.3(240) | 76.4(180) | 73.1(300) | 77.2(0) | 77.7(60) |
| 4 | 70.9(180) | 73.9(120) | 69.6(300) | 72.8(60) | 70.7(240) | 74.3(0) |
| 5 | 78.1(300) | 81.9(60) | 81.6(0) | 79.3(180) | 79.7(120) | 78.3(240) |
| 6 | 74.2(240) | 74.2(300) | 76.1(60) | 77.6(0) | 73.9(180) | 74.8(120) |

*Table 4.3: Results of food product experiment*

Figure 4.1 shows the core temperature (°C) over time per food product specimen. The labels O1, O2 and O3 indicate operators.

### 4.6.3 Analysis

The data can be analysed with both ML and REML, but also with the ‘naive’ standard Gauge R&R study (which does not take into account the fixed differences between time instants). The estimates for the variance components using the three different methods are shown in Table 4.4. In this table $\boldsymbol{\sigma}^2_{repro}$ denotes the reproducibility which is the sum of $\boldsymbol{\sigma}_2^2$ and $\boldsymbol{\sigma}_3^2$, and $\boldsymbol{\sigma}^2_m$ denotes the total measurement variance.

| Approach | Variance components $\sigma_1^2$ | $\sigma_{repro}^2$ | $\sigma_2^2$ | $\sigma_3^2$ | $\sigma_0^2$ | $\sigma_m^2$ |
|---|---|---|---|---|---|---|
| ML | $10.68$ | $0.77$ | $0.77$ | $0.00$ | $1.12$ | $1.89$ |
| REML | $12.54$ | $0.81$ | $0.81$ | $0.01$ | $1.36$ | $2.17$ |
| Naive Gauge R&R | $11.71$ | $0.39$ | - | - | $6.35$ | $6.74$ |

*Table 4.4: Variance component estimates per approach*

From the results we conclude that taking into account the fixed differences between time instants largely improves (in this case reduces) the estimate for measurement spread. The estimate reduces from $6.74$ to $2.17$ and $1.89$ respectively. The differences in estimates for variance components between ML and REML are small. Table 4.5 shows the estimated core temperature per time instant. It confirms that the temperature decreases by about $1.0$°C per minute.

| Approach | Mean temperature per time instant $0$ | $60$ | $120$ | $180$ | $240$ | $300$ |
|---|---|---|---|---|---|---|
| ML | $79.26$ | $78.16$ | $77.06$ | $76.01$ | $74.25$ | $73.60$ |

*Table 4.5: Estimates of mean core temperature per time instant*

For both ML and REML sampling errors are computed with the bootstrap approach and using the asymptotic results. In Table 4.6 these are shown.

| Variance | ML | | REML | |
|---|---|---|---|---|
| component | Bootstrap | Asymptotic | Bootstrap | Asymptotic |
| $\sigma_1^2$ | $6.21$ | $6.40$ | $8.92$ | $8.08$ |
| $\sigma_2^2$ | $0.69$ | $0.86$ | $0.87$ | $0.92$ |
| $\sigma_3^2$ | $0.31$ | $0.31$ | $0.18$ | $0.46$ |
| $\sigma_0^2$ | $0.21$ | $0.37$ | $0.42$ | $0.51$ |
| $\sigma_m^2$ | $1.20$ | $1.46$ | $1.47$ | $1.89$ |

*Table 4.6: Estimated standard errors for variance components*

From the results in Table 4.6 two conclusions can be drawn:

- The standard error of the estimated measurement error is large; to improve this precision one needs to replicate the design.
- Bootstrapped and asymptotic standard error are quite similar.

## 4.7 Conclusions

In gauge R&R studies with destructive measurement systems, one is often forced to replace replications with measurements of different objects, thus confounding measurement variation with between-objects-within-sample variation. This chapter presents a methodology for reducing the resulting overestimation by exploiting patterns in the object-to-object variation. This is illustrated from the example in which the estimate of measurement variance reduces by a factor of three.

In the example, the standard errors of the estimates for variance components are quite large (see Table 4.6), especially the errors in the estimate for reproducibility. This is due to the fact that in the current design only three operators are used. Three is the typical number of operators in gauge R&R studies reported in literature and expected in practice. Also in the standard gauge R&R study, this small number results in very large standard errors for the estimated variance components, a problem that has been reported in scientific literature by for example Burdick and Larsen (1997), Vardeman and Van Valkenburg (1999). To get better estimates one has to replicate the design, but it only helps to use the replication scheme in which the number of operators is increased.

## Appendix

### Implementation details for the EM-algorithm

ML and REML estimation are implemented using the EM algorithm (Dempster, Laird and Rubin, 1977). Details of the algorithm applied to estimation of variance components are explained in McCulloch and Searle (2001, p.264). In the terminology of literature on the algorithm the incomplete data are the elements of the data vector $\mathbf{y}$, and the complete data are the elements of the data vector $\mathbf{y}$ together with the values of the unobserved random effects $(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3})$. The algorithm consists of an iteration of two steps, the E-step and the M-step. In the E-step (expectation step) the conditional expectation of the log-likelihood of the complete data is computed, given the incomplete data $\mathbf{y}$. The M-step (maximisation step) then maximises the conditional expectation of the log-likelihood function in order to determine the parameter estimates needed for the next iteration. The algorithm iterates between the E-step and M-step until convergence is reached.

Searle, Casella, and McCulloch (1992, pp.298–299) show that in cases such as the one at hand it suffices in the E-step to compute the conditional expected values of $\mathbf{u}_i'\mathbf{u}_i$ and $\mathbf{y} - \sum_{i=1}^{3} \mathbf{Z}_i\mathbf{u}_i$ given $\mathbf{y}$, i.e. $\mathrm{E}(\mathbf{u}_i'\mathbf{u}_i|\mathbf{y})$ (denoted by $\hat{t}_i$, where $i = 0, 1, 2, 3$), and $\mathrm{E}(\mathbf{y} - \sum_{i=1}^{3} \mathbf{Z}_i\mathbf{u}_i|\mathbf{y})$ (denoted by $\hat{\mathbf{s}}$). The length of the vector $\mathbf{u}_i$ is denoted by $q_i$.

Following Searle, Casella, and McCulloch (1992, p.300), we first have to choose starting values for $\boldsymbol{\gamma}^{(0)}$ and $\boldsymbol{\sigma}^{2(0)}$.

Step $1$ is to calculate ($m = 0, 1, 2, \ldots$):

$$\begin{aligned}\hat{t}_i^{(m)} = \sigma_i^{4(m)} & (\mathbf{y} - \mathbf{X}\boldsymbol{\gamma}^{(m)})'(\mathbf{V}^{(m)})^{-1}\mathbf{Z}_i\mathbf{Z}_i'(\mathbf{V}^{(m)})^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\gamma}^{(m)}) \\ + & \quad \mathrm{tr}(\sigma_i^{2(m)}\mathbf{I}_{q_i} - \sigma_i^{4(m)}\mathbf{Z}_i'(\mathbf{V}^{(m)})^{-1}\mathbf{Z}_i)\end{aligned}$$

where $\mathbf{V}^{(m)} = \mathbf{V}(\boldsymbol{\sigma}^{2(m)})$, and

$$\hat{\mathbf{s}}^{(m)} = \mathbf{X}\boldsymbol{\gamma}^{(m)} + \sigma_0^{2(m)}(\mathbf{V}^{(m)})^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\gamma}^{(m)}).$$

Step $2$, estimation of the parameters, takes on the following form:

$$\boldsymbol{\sigma}_i^{2(m+1)} = \hat{t}_i^{(m)}/q_i$$

and

$$\mathbf{X}\boldsymbol{\gamma}^{(m+1)} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\hat{\mathbf{s}}^{(m)},$$

from which $\boldsymbol{\gamma}$ can be obtained by solving the system of equations. The latter is possible because we already imposed the condition $\sum_{i=1}^{s} \gamma_i = 0$. In step $3$ it is determined whether convergence has been reached. If not, $m$ is increased by unity and one returns to step $1$. If so, we have ML estimates $\tilde{\sigma}_i^2 = \sigma_i^{2(m+1)}$ and $\tilde{\boldsymbol{\gamma}} = \boldsymbol{\gamma}^{(m+1)}$.
The REML equations are quite similar and can be found in Searle, Casella, and McCulloch (1992, pp.302–303) as well, albeit that they only involve estimates for the variance components $\boldsymbol{\sigma}^2$. In the EM algorithm for REML a matrix $\mathbf{P}$ is used, which has the following form:

$$\mathbf{P} = \mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}^{-1},$$

in which $\mathbf{K}$ is the matrix which defines the contrasts in REML. The EM algorithm for REML then takes the following form. First, starting values $\boldsymbol{\sigma}^{2(0)}$ are chosen. In step $1$ $\hat{t}_i^{(m)}$ is calculated:

$$\hat{t}_i^{(m)} = \sigma_i^{4(m)}\mathbf{y}'\mathbf{P}^{(m)}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}^{(m)}\mathbf{y} \quad + \quad \mathrm{tr}(\sigma_i^{2(m)}\mathbf{I}_{q_i} - \sigma_i^{4(m)}\mathbf{Z}_i'\mathbf{P}^{(m)}\mathbf{Z}_i).$$

In the second step the $\sigma_i^2$'s are estimated by

$$\sigma_i^{2(m+1)} = \hat{t}_i^{(m)}/q_i.$$

Finally, in step $3$ it is determined whether convergence has been reached. If not, $m$ is increased by one and one returns to step $1$. If so, one sets $\tilde{\sigma}_i^2 = \sigma_i^{2(m+1)}$.

### Convergence of the EM-algorithm

An important issue in any iterative estimation procedure is the question whether the iteration converges to a stationary point and whether — in the case of ML estimation — this stationary point is the global maximum of the likelihood function. Following Wu (1983), to ensure that the EM algorithm converges to a stationary point of the log-likelihood, it suffices to verify that:

1. $\Sigma_{\boldsymbol{\gamma}_0,\boldsymbol{\sigma}_0^2} = \{(\boldsymbol{\gamma},\boldsymbol{\sigma}^2) : l(\boldsymbol{\gamma},\boldsymbol{\sigma}^2) \geq l(\boldsymbol{\gamma}_0,\boldsymbol{\sigma}_0^2)\}$ is compact for any $l(\boldsymbol{\gamma}_0,\boldsymbol{\sigma}_0^2) > -\infty$,

2. $\mathrm{E}_{\boldsymbol{\theta}'}(l(\boldsymbol{\theta}|(\mathbf{y},\mathbf{u})|\mathbf{y}))$ is continuous in both $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$.

For condition (1) it is necessary to confine the parameter set to a compact subset of $R^{11}$. Hence we assume that there exists a sufficiently large $L > 0$ such that $\boldsymbol{\sigma}^2 \in [0, L]^4$ and $\boldsymbol{\gamma} \in [-L, L]^7$. Condition (2) is satisfied because the conditional density of $(\mathbf{y}, \mathbf{u})$ given $\mathbf{y}$ is multivariate normally distributed. The EM iteration proceeds until convergence is reached. As a criterion for reaching convergence the difference in the log-likelihood in the last two iterations is taken. If the increase in the log-likelihood between two iterations decreases below the threshold value $10^{-5}$, the iteration procedure stops.

Having thus established that the EM procedure converges to a stationary point, a second worry is whether this stationary point is the global maximum (and not a local maximum or a saddle point). A recommended practice in using the EM algorithm is to start with a set of overdispersed initial values, to increase the probability to detect all maxima. In our algorithm this is implemented by selecting 10 sets of starting values (one for every parameter). In the ML case for each of the six levels $\gamma_i$ $(i = 1, 2, \ldots, 6)$ and for the overall average $\mu$ ten starting values are selected from a uniform $(-L, L)$-distribution. $L$ is a positive integer large relative to the scale of the observations. For both the ML and the REML case for each of the four variance components $(\sigma_0^2, \sigma_1^2, \sigma_2^2, \sigma_3^2)$ ten starting values are selected from a uniform $(0, L)$-distribution. In both the ML and the REML cases all sets of starting values are iterated through 100 loops of the algorithm. The set resulting in the highest value for the log-likelihood function is selected and proceeded with.

### Replication of the experimental set-up

Two possible ways to replicate the design are described in the section on replication and extensions of the experimental design. Here we show how the formulation $\mathbf{y} = \mathbf{X}\gamma + \mathbf{Z}\mathbf{u}$. is adjusted accordingly. In the first case we get the following changes:

- The vector $\mathbf{y}$ becomes a $2pqr \times 1$ vector
- $\mathbf{X}$ becomes a $2pqr \times (qr + 1)$ matrix constructed in a similar fashion as explained in section 2, but now from the replicated experimental design.
- The $\mathbf{Z}$-matrices are similar to the ones in (4.4), but with $p$ replaced with $2p$.
- The vectors of random effects change to:

$$\begin{aligned}
\mathbf{u_0} &= (\epsilon_{111}, \epsilon_{112}, \ldots, \epsilon_{11r}, \epsilon_{121}, \ldots \epsilon_{1q1}, \ldots \epsilon_{1qr}, \epsilon_{211}, \ldots \epsilon_{2p,q,r}) \\
\mathbf{u_1} &= (a_1, a_2, \ldots, a_{2p}) \\
\mathbf{u_2} &= (b_1, b_2, \ldots, b_q) \\
\mathbf{u_3} &= (ab_{11}, ab_{12}, \ldots, ab_{1q}, ab_{21}, ab_{22}, \ldots, ab_{2p,q})
\end{aligned}$$

Finally the covariance structure does not change. Upon the changes prescribed the estimation procedure gives parameter estimators, which have smaller standard errors. Evidently, the design can be replicated an arbitrary number of times.
The second situation is somewhat more complicated and we get the following changes:

- $\mathbf{y}$ becomes:

$$\mathbf{y} = (y_{111}, y_{112}, \ldots, y_{11r}, y_{121}, \ldots, y_{12r}, \ldots, y_{1q1}, \ldots, y_{1qr}, y_{211}, \ldots, y_{p,q,r}, y_{p+1,q+1,1}, \cdots y_{2p,2q,r})'.$$

- $\mathbf{X}$ becomes a $2pqr \times (qr+1)$ matrix constructed in a similar fashion as explained in section 2, but now from the replicated experimental design.
- The $\mathbf{Z}$-matrices are similar to the ones in (4.4), but with $p$ replaced with $2p$, except for $\mathbf{Z_2}$, which becomes:

$$\mathbf{Z_2} = \mathbf{I}_2 \otimes \{_c\mathbf{I}_p\} \otimes \mathbf{I}_q \otimes \{_c\mathbf{I}_r\}.$$

- The vectors of random effects change to:

$$\begin{aligned}
\mathbf{u_0} &= (\epsilon_{111}, \epsilon_{112}, \ldots, \epsilon_{11r}, \epsilon_{121}, \ldots \epsilon_{1q1}, \ldots, \epsilon_{1qr}, \epsilon_{211}, \ldots, \epsilon_{p,q,r}, \epsilon_{p+1,q+1,1}, \ldots, \epsilon_{2p,2q,r}) \\
\mathbf{u_1} &= (a_1, a_2, \ldots, a_{2p}) \\
\mathbf{u_2} &= (b_1, b_2, \ldots, b_{2q}) \\
\mathbf{u_3} &= (ab_{11}, ab_{12}, \ldots, ab_{1q}, ab_{21}, \ldots, ab_{p,q}, ab_{p+1,q+1}, \ldots, ab_{2p,2q})
\end{aligned}$$

## References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV). American Psychiatric Association, Washington, DC.

Bennis, W.G. and O'Toole, J. (2005). "How business schools lost their way." *Harvard Business Review*, 83(5), 96–104.

Bisgaard, S. and Freiesleben, J. (2001). "Economics of Six Sigma programs – Quality Quandaries." *Quality Engineering*, 13(2), 325–331.

Breyfogle, F.W. (1999). *Implementing Six Sigma - Smarter Solutions Using Statistical Methods*. John Wiley and Sons, Inc., New York.

Bridgman, P.W. (1927). *The Logic of Modern Physics*. MacMillan, New York.

Burdick, R.K., Borror, C.M. and Montgomery, D.C. (2003). "A review of methods for measurement systems capability analysis." *Journal of Quality Technology*, 35(4), 342–354.

Burdick, R.K., Larsen,G.A. (1997). "Confidence intervals on measures of variability in R&R studies." *Journal of Quality Technology*, 39(3), 261–273.

Chrisman, J.J, Hofer, C.W. and Boulton, W.R. (1988). "Toward a system for classifying business strategies." *Academy of Management Review*, 13(3), 413–428.

Cochran, W.G. and Cox, G.M. (1950). *Experimental Designs*. John Wiley and Sons, Inc., New York.

Creveling, C.M., Slutsky, J.L. and Antis, D. (2003). *Design for Six Sigma – In Technology and Product Development*. Prentice Hall, Englewood Cliffs, New Jersey.

Davia, G. (1998). "Thoughts on a possible rational reconstruction of the method of 'rational reconstruction'." Paper presented at 20th *World Congress of Philosophy*, Boston.

De Koning, H. and De Mast, J. (2005). "Grounding of Six Sigma's Breakthrough Cookbook: how to research a methodology?" *International Journal of Six Sigma and Competitive Advantage*, 1(3), 263-275.

De Koning, H. and De Mast, J. (2006). "A rational reconstruction of Six Sigma's Breakthrough Cookbook." *International Journal of Quality and Reliability Management*, 23(7), 766–787.

De Koning, H. and De Mast, J. (2007). "The CTQ Flowdown as a conceptual model of project objectives." *Quality Management Journal*, 14(2), 19–28.

De Koning, H., De Mast, J., Does, R.J.M.M., Vermaat, M.B. and Simons, S. (2007). "Generic Lean Six Sigma projects in financial services." Submitted for publication in *Quality Management Journal*.

De Koning, H, Does, R.J.M.M. and Bisgaard, S. (2006). “Lean Six Sigma in financial services.” Submitted for publication in *International Journal of Six Sigma and Competitive Advantage*.

De Koning, H., Verver, J.P.S., Van den Heuvel, J., Bisgaard, S. and Does R.J.M.M. (2006). “Lean Six Sigma in healthcare.” *Journal for Healthcare Quality*, 28(2), 4–11.

De Mast, J. (2002). *Quality Improvement from the Viewpoint of Statistical Method*. PhD thesis, University of Amsterdam.

De Mast, J. (2003). “Quality improvement from the viewpoint of Statistical Method.” *Quality and Reliability Engineering International*, 19(4), 255-264.

De Mast, J. (2004). “A methodological comparison of three strategies for quality improvement.” *International Journal of Quality and Reliability Management*, 21(2), 198–213.

De Mast, J. (2007). “Integrating the many facets of Six Sigma.” Submitted for publication in *Quality Engineering*.

De Mast, J. and Bisgaard, S. (2007). “The science in Six Sigma.” *Quality Progress*, 40(1), 25–29.

De Mast, J. and Does, R.J.M.M. (2006). “Industrial statistics: a discipline with opportunities and challenges.” *Statistica Neerlandica*, 60(3), 1–13.

De Mast, J. and Trip, A. (2005). “Gauge R&R studies for destructive measurements.” *Journal of Quality Technology*, 37(1), 40–49.

De Mast, J., Does, R.J.M.M. and De Koning, H. (2006). *Lean Six Sigma for Service and Healthcare*. Beaumont Quality Publications, Alphen aan den Rijn, The Netherlands.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). “Maximum likelihood from incomplete data via the EM algorithm.” *Journal of the Royal Statistical Society B.*, 39(1), 1–38.

Does, R.J.M.M, Vermaat, M.B, De Koning, H., Bisgaard, S. and Van den Heuvel, J. (2006). “Generic Six Sigma Projects in Healthcare.” *Six Sigma Forum Magazine*, 6(1), 14–23.

Doggett, A.M. (2005). “Root cause analysis: a framework for tool selection.” *Quality Management Journal*, 12(4), 34–45.

Duncan, A.J. (1986). *Quality Control and Industrial Statistics*, 5th edition. IRWIN, Homewood.

Easton, G.S. and Jarrell, S.L (1998). “The effects of total quality management on corporate performance: an empirical investigation.” *Journal of Business*, 71(2), 253–307.

Eckes, G. (2001). *The Six Sigma Revolution – How General Electric and Others Turned Process into Profit*. John Wiley and Sons, Inc., New York.

Fine, T. (1988). “Foundations of probability.” In: *Encyclopedia of Statistical Sciences*, 8th edition, Kotz, S. and Johnson, N. (editors). John Wiley and Sons, Inc., New York.

Galbraith, J.K. (1991). *A History of Economics: The Past as the Present*. Penguin Books Ltd, UK.

Gano, D.L. (2001). "Effective problem solving: a new way of thinking." *ASQ Annual Quality Congress Proceedings*, 55, 110–122.

George, M.L. (2003). *Lean Six Sigma for Services*. McGraw-Hill, USA.

Goh, T.N. (2002). "A strategic assessment of Six Sigma." *Quality and Reliability Engineering International*, 18(5), 403–410.

Habermas, J. (1981). *Theorie des Kommunikativen Handelns*, Band 1. Suhrkamp, Frankfurt.

Hahn, G.J., Doganaksoy, N. and Hoerl, R.W. (2000). "The evolution of Six Sigma." *Quality Engineering*, 12(3), 317–326.

Hahn, G.J., Hill, W.J., Hoerl, R.W. and Zinkgraf, S.A. (1999). "The impact of Six Sigma improvement – a glimpse into the future of statistics." *The American Statistician*, 53(3), 208–215.

Harry, M. (1997). *The Vision of Six Sigma*, 5th edition. Tri Star, Phoenix.

Harry, M.J. (1998). "Six Sigma: a breakthrough strategy for profitability." *Quality Progress*, 31(5), 60–64.

Harry, M.J. and Schroeder, R. (2000). *Six Sigma – The Breakthrough Management Strategy Revolutionizing the Worlds Top Corporations*. Doubleday, New York.

Hoare, C.A.R. (1972). "Proof of correctness of data representations." *Acta Informatica*, 1(4), 271–281.

Hoerl, R.W. (2001). "Six Sigma black belts: what do they need to know?" *Journal of Quality Technology*, 33(4), 391–406.

Ishikawa, K. (1982). *Guide to Quality Control*. Asian Productivity Organization, Tokyo.

Kamlah, A. (1980). "Wie arbeitet die analytische Wissenschaftstheorie?" *Zeitschrift für allgemeine Wissenschaftstheorie*, 11, 23-44.

Kaplan, R.S. and Norton, D.P. (2001). *The Strategy-focused Organization. How Balanced Scorecard Companies Thrive in the New Business Environment*. Harvard Business School Publishing Corporation, Boston, Massachusetts.

Lind, M. and Goldkuhl, G. (2002). "Grounding of methods for business change: altering between empirical, theoretical and internal grounding." Paper presented at the *European Conference on Research Methods in Business and Management Studies (ECRM 2002)*.

Linderman, K., Schroeder, R.G., Zaheer, S. and Choo, A.S. (2003). "Six Sigma: a goal-theoretic perspective." *Journal of Operations Management*, 21(2), 193–203.

Lynch, D., Berolono, S. and Cloutier, E. (2003). "How to scope DMAIC projects?" *Quality Progress*, 36(1), 37–41.

Mayo, D.G. (1996). *Error and the Growth of Experimental Knowledge*. The University of Chicago Press, Chicago.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*, 2nd edition. Chapman and Hall, London.

McCulloch, C.E. and Searle, S.R. (2001). *Generalized, Linear, and Mixed Models*. John Wiley and Sons, Inc., New York.

Mintzberg, H. (1994). *The Rise and Fall of Strategic Planning*. The Free Press, New York.

Montgomery, D.C. (1997). *Introduction to Statistical Quality Control*, 3th edition. John Wiley and Sons, Inc., New York.

Montgomery, D.C. (2005). *Design and Analysis of Experiments*, 6th edition. John Wiley and Sons, Inc., New York.

Pande, P.S., Neuman, R.P. and Cavanagh, R.R. (2000). *The Six Sigma Way – How GE, Motorola, and Other Top Companies Are Honing Their Performance*. McGraw-Hill, New York.

Parasuraman, A., Zeithaml V.A. and Berry, L.L. (1985). "A conceptual model for service quality and its implications for future research." *Journal of Marketing*, 49(4), 41–50.

Park, S.H. (2003). *Six Sigma for Quality and Productivity Promotion*. Asian Productivity Organization, Tokyo.

Partington, D. (1996). "The project management of organizational change." *International Journal of Project Management*, 14(1), 13–21.

Polanyi, M. (1958). *Personal knowledge*. Routledge, London.

Poser, H. (1980). "Rekonstruktion, rationale." In: *Handbuch Wissenschaftstheoretischer Begriffe*, band 3, J. Speck (editor). Vandenhoeck und Ruprecht, Göttingen

Pyzdek, T. (2001). *The Six Sigma Handbook – a Complete Guide for Greenbelts, Blackbelts, and Managers at All Levels*. McGraw-Hill, New York.

Rasis, D. Gitlow, H. and Popovich, E. (2002). "Paper organizers international: a fictitious Six Sigma Green Belt case study. I." *Quality Engineering*, 15(2), 127–145.

Reed, R., Lemak, D.J. and Mero, N.P. (2000). "Total quality management and sustainable competitive advantage." *Journal of Quality Management*, 5(1), 5–26.

Sanders, D. and Hild, C. (2000). "Six Sigma on business processes: common organizational issues." *Quality Engineering*, 12(4), 603-610.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992). *Variance Components*. John Wiley and Sons, Inc., New York.

Shingo, S. (1989). *A Study of the Toyota Production System*. Productivity Press, New York.

Smith, J.M. and Smith D.C.P. (1977). "Database abstractions: aggregation and generalization." *ACM Transactions on Database Systems*, 2(2), 106–133.

Snee, R.D. (2004). "Six Sigma: the evolution of 100 years of business improvement methodology." *International Journal of Six Sigma and Competitive Advantage*, 1(1), 4–20.

Standard, C. and Davis, D. (1999). *Running Todays Factory. A Proven Strategy for Lean Manufacturing*. Hanser Gardner, Cincinnati.

Stephens, K.S. (editor) (2003). *Six Sigma and Related Studies in the Quality Disciplines – The Best on Quality*, 14. ASQ Quality Press, Milwaukee, Wisconsin.

Thagard, P. (1992). *Conceptual Revolutions*. Princeton University Press, Princeton, New Yersey.

Van den Heuvel, J., Does, R.J.M.M. and De Koning H. (2006). "Lean Six Sigma in a hospital." *International Journal of Six Sigma and Competitive Advantage*, 2(4), 377–388.

Van Wieringen, W.N. (2003). *Statistical Models for the Precision of Categorical Measurement Systems*. PhD thesis, University of Amsterdam.

Vardeman, S.B. and Van Valkenburg, E.S. (1999). "Two-way random-effects analyses and gauge R&R studies." *Technometrics*, 41(3), 202–211.

Vermaat, M.B. and Does R.J.M.M. (2006). "A semi-Bayesian method for Shewhart individual control charts." *Quality Technology and Quantitative Management*, 3(1), 111–125.

Wasserman, G.S. and Lindland, J.L. (1996). "A case study illustrating the existence of dynamics in traditional cost-of-quality models." *Quality Engineering*, 9(3), 119–128.

Womack, J.P. and Jones, D.T. (2003). *Lean Thinking*, 2nd edition. Free Press, New York.

Womack, J.P., Jones, D.T. and Roos, D. (1990). *The Machine that Changed the World: The Story of Lean Production*. Rawson Associates, New York.

Wren, D.A. (2005). *History of management thought*, 5th edition. John Wiley and Sons, Inc., New York

Wu, C.F. (1983). "On the convergence properties of the EM algorithm." *The Annals of Statistics*, 11(1), 95–103.

Zipkin, P.H. (1991). "Does manufacturing need a JIT revolution?" *Harvard Business Review*, 69(1), 40–50.

# Samenvatting: Wetenschappelijke fundering van de methode van Lean Six Sigma

## Het Lean Six Sigma programma

De twintigste eeuw was getuige van een snelle ontwikkeling van professionele organisaties. Innovaties op het gebied van managementstructuren en verbetermethoden waren hierin een belangrijke factor. Op het moment dat de race om concurrenten te verslaan echt op gang kwam, begonnen organisaties 'best practices' en verbetermethoden van elkaar te kopiëren. De praktijk heeft de verbetermethoden eruit gefilterd die echt iets bijdragen. Lean Six Sigma (LSS) is de meest recente incarnatie van een lange reeks verbetermethoden, die in de twintigste eeuw het licht zagen.

Het LSS programma bevat diverse elementen. In de eerste plaats heeft het een achterliggende bedrijfskundige rationale. LSS poogt een bijdrage te leveren om de concurrentiestrijd te overleven door kostenreducties te realiseren, klanttevredenheid te verhogen en het vermogen om te innoveren en continu te verbeteren te vergroten. Daarnaast geeft LSS een organisatiestructuur voor het organiseren van de verbetering van routinematige processen. LSS schrijft onder meer voor dat verbeteren op projectbasis plaats vindt. Om deze projecten effectief uit te kunnen voeren, voorziet LSS tenslotte in een gestructureerde aanpak voor verbeterprojecten.

## Onderzoeksmethode

Het LSS programma helpt projectleiders bij de uitvoering van verbeterprojecten. Het onderwerp van dit proefschrift – de methodologische elementen van het LSS programma – kan daarom worden opgevat als een systeem van *voorschriften*; richtlijnen, die een projectleider aangeven wat hij moet doen om een bepaald doel te bereiken. Het moge duidelijk zijn dat onderzoek naar LSS niet kan worden uitgevoerd volgens de standaardaanpak die wordt gebruikt in mathematisch onderzoek. De vraag is op welke manier LSS dan wel onderzocht dient te worden. Het eerste hoofdstuk van het proefschrift zoekt naar een kader om de studie vorm te geven.

Hoewel empirisch onderzoek tot de mogelijkheden behoort, geeft dat geen goed inzicht in de vraag waarom LSS effectief is. In dit onderzoek is ervoor gekozen om LSS vanuit vijf invalshoeken te onderzoeken:

- Rationele reconstructie: De LSS methode is in de huidige literatuur op een niet-wetenschappelijke wijze geformuleerd. De eerste stap van het onderhavige onderzoek is erop gericht om uit de los geformuleerde en vaak niet volledig coherente beschrijvingen van LSS, een precieze, consistente en helder geformuleerde beschrijving te destilleren.

- Waardefundering: Voorschriften worden gerechtvaardigd door hun achterliggende doel. In beschrijvingen van LSS lopen projectdoelen uiteen van kwaliteitsverbetering, variatiereductie, het realiseren van doorbraken, en foutreductie. Deze worden vervolgens weer gerechtvaardigd met concepten zoals kosten-van-lage-kwaliteit en modellen zoals de verborgen fabriek. Er is echter op dit punt meer onderzoek nodig, met name gericht op het integreren van bestaande verklaringen met bedrijfseconomische theorieën.
- Theoretische fundering: De effectiviteit van voorschriften kan ook worden gevalideerd door vanuit een externe theorie te verklaren waarom ze werken.
- Empirische fundering: Empirische fundering betekent dat de effectiviteit van voorschriften wordt aangetoond op basis van empirische data.
- Tenslotte poogt dit onderzoek een specificatie te geven van de condities waaronder LSS toegepast kan worden.

Bestudering van de literatuur die voor handen is laat zien dat deze op geen enkele van de genoemde punten een adequate beschrijving geeft.

### Onderwerp en doelstelling van het proefschrift

De doelstelling van het proefschrift is om methodologische aspecten van LSS wetenschappelijk te funderen. Methoden zoals LSS bestaan uit vier klassen van elementen:

- Bedrijfskundige rationale: De bedrijfskundige motivatie die ten grondslag ligt aan de invoering van LSS.
- Stapsgewijze aanpak: Een stappenplan, dat de projectleider kan volgen bij het uitvoeren van een verbeterproject.
- Hulpmiddelen en technieken: LSS voorziet in een breed spectrum aan procedures, dat een projectleider helpt tussenresultaten te bereiken.
- Concepten en classificaties: Om de bovenstaande elementen te kunnen communiceren, bedient de LSS methode zich van concepten (zoals CTQ) en classificaties (bijvoorbeeld de DMAIC fasering).

### Motivatie voor dit onderzoek

De motivatie voor het uitvoeren van dit onderzoek komt voort uit twee doelstellingen, namelijk het verschaffen van begrip met betrekking tot LSS en het faciliteren van het effectief gebruik van LSS. Voor beide doelen is het essentieel om eerst een rationele reconstructie uit te voeren en vervolgens LSS op diverse manieren te funderen.

Dit onderzoek is niet zuiver wiskundig. Vandaar dat de auteur het belangrijk vindt in hoofdstuk 1 zijn visie te verduidelijken op de relatie tussen industriële statistiek en

wiskunde. De conclusie is dat industriële statistiek zich wel bedient van wiskundig redeneren, maar zich niet daartoe beperkt. Andere methoden – zoals rationele reconstructie en funderingsonderzoek – spelen ook een belangrijke rol, hetgeen door de geschiedenis van het vakgebied nog eens onderstreept wordt.
Tenslotte wordt de vraag beantwoord of dit onderzoek zich kwalificeert als valide promotieonderzoek. De stelling wordt ingenomen dat er verschillende typen wetenschappelijk onderzoek bestaan. Aan de ene kant onderzoek dat goed gedefinieerde, welomschreven problemen aanpakt met bekende methoden. Aan de andere kant bestaat er onderzoek, zoals dit onderzoek, waarin een echt, maar typisch slecht gedefinieerd probleem wordt geanalyseerd. In dit laatste geval probeert men dit vaag geformuleerde en slecht gedefinieerde probleem adequaat te definiëren en met wetenschappelijke precisie en objectiviteit te analyseren.

### Een rationele reconstructie van Six Sigma en Lean Six Sigma

In het tweede hoofdstuk wordt een rationele reconstructie gemaakt van de eerder genoemde methodologische elementen van LSS. Het volgende materiaal wordt hierbij gepresenteerd:

- De rationele reconstructie van de bedrijfskundige motivatie van LSS. In de bestaande literatuur wordt de toepassing van LSS onderbouwd door te verwijzen naar ‘showcases’, het model van de verborgen fabriek (verbetering kostenstructuur) en het verkrijgen van strategische voordelen.
- De rationele reconstructie van de strategie en het stappenplan. Het blijkt dat de fasering en onderliggende stappen, die genoemd worden in de literatuur, in hoge mate convergeren. Daarom is het mogelijk een generieke fasering en een generiek stappenplan te construeren. Een aantal afwijkingen van de generieke aanpak wordt besproken.
- De rationele reconstructie van de gereedschapskist van LSS.

In de laatste sectie van hoofdstuk twee wordt de integratie tussen Lean en Six Sigma besproken. Lean bestaat uit een verzameling hulpmiddelen en technieken en is erop gericht problemen in productie- en dienstverleningsprocessen op te lossen. Lean verbreedt het vocabulair van Six Sigma en voegt een aantal technieken en hulpmiddelen toe aan de gereedschapskist van Six Sigma.

### De CTQ flowdown als conceptueel model van projectdoelen

In het derde hoofdstuk wordt een hulpmiddel beschreven en verduidelijkt, dat de kern vormt van de definitie van de meeste LSS projecten. Dit hulpmiddel heet de CTQ flowdown. Het laat de relatie zien tussen strategische focuspunten en projectdoelen.

Deze projectdoelen worden op hun beurt vertaald naar en opgesplitst in CTQ's, die vervolgens worden geoperationaliseerd in de vorm van metingen. In dit hoofdstuk wordt de precieze aard van de relaties tussen strategische focuspunten, projectdoelen, CTQ's en metingen uitgediept. De CTQ flowdown dient diverse doelen. Hij draagt zorg voor een heldere projectdefinitie, verduidelijkt de bedrijfskundige motivatie van een verbeterproject, maakt de aannames achter de projectdefinitie expliciet, helpt te focusen op de meest urgente organisatieproblemen en faciliteert het optimaal oplossen van trade-off problemen.

In de tweede helft van hoofdstuk drie worden de ontwikkelde theorie en de bijbehorende technieken gebruikt om te komen tot een aantal generieke modellen (sjablonen) voor LSS projectdefinities. De hier gepresenteerde sjablonen zijn van toepassing op projecten in de financiële dienstverlening. Projectleiders kunnen deze sjablonen gebruiken als voorbeeld en leidraad in de definitiefase. Dit helpt hen te komen tot kristalheldere projectdefinities met expliciete doelen en een solide bedrijfskundige rationale.

### Een experimentele opzet voor destructieve gauge R&R Studies

In LSS projecten wordt in hoge mate gebruik gemaakt van data. De betrouwbaarheid van die data is derhalve een belangrijk issue, dat vaak weinig aandacht krijgt. Binnen LSS krijgt zij wel voldoende aandacht. Onderdeel van de studie naar de betrouwbaarheid van data is een onderzoek naar de precisie van het meetsysteem.

De standaardmethode om de precisie van een meetsysteem vast te stellen is een zogenoemde 'gauge repeatability and reproducibility' (gauge R&R) studie. Daarin worden herhaalde metingen benut om variantiecomponenten te schatten, die vervolgens als meetspreiding worden geïnterpreteerd. Voor destructieve metingen is het niet mogelijk om herhaalde metingen uit te voeren. Een mogelijke oplossing is om de herhaalde metingen te vervangen door metingen aan meerdere objecten. Vervolgens worden deze metingen gemodeleerd met behulp van een vast patroon. We beschouwen een gauge R&R studie met een onderzoeksopzet gebaseerd op een Latijns vierkant. Het model dat hieruit voortvloeit is een ongebalanceerd, gemengd, lineair model. Voor dit model worden de variantiecomponenten geschat met behulp van twee schattingmethoden (maximum likelihood en restricted maximum likelihood). De numerieke optimalisatie is gedaan met behulp van een implementatie van het EM-algoritme. De benadering blijkt goed van toepassing op een voorbeeld uit de praktijk, waarin de temperatuur van een voedselproduct wordt gemeten.

# Curriculum vitae

Henk de Koning is geboren op 25 februari 1977 te Langbroek. Na zijn middelbare school te Doorn is hij naar Utrecht gegaan om natuurkunde en psychologie te studeren. Na deze studies cum laude te hebben afgerond, besloot hij eerst te gaan werken in de strategische advisering en trad hij in dienst bij OC&C Strategy Consultants. Na ruim een jaar besloot hij de strategische advisering te verruilen voor IBIS UvA, om training en advisering te combineren met wetenschappelijk onderzoek.

Bij het IBIS UvA houdt hij zich bezig met onderzoek en advisering in de industriële statistiek en met Lean Six Sigma in het bijzonder. Naast deze twee taken heeft de auteur in het academisch jaar 2005/2006 het mastervak "Industrial Statistics" aan de Universiteit van Amsterdam verzorgd. Het onderzoek bij het IBIS UvA mondde uit in dit proefschrift, en daarnaast in meer dan tien wetenschappelijke publicaties over Lean Six Sigma. Onder meer verscheen het boek "Lean Six Sigma for Service and Healthcare", waarvan hij co-auteur is. Als adviseur bij het IBIS UvA concentreerde Henk zich met name op het begeleiden van de invoering van het verbeterprogramma Lean Six Sigma. In deze rol heeft hij meer dan 250 projectleiders en medewerkers opgeleid en ondersteund. Hij was onder meer actief bij Burgers Ergon, Getronics PinkRoccade, LG Philips-Displays, Nationale Nederlanden, TNT Post, en Wolters Kluwer.