UNIVERSITY OF TURKU

Identification and Root Cause Analysis of the Key Bottlenecks Affecting the Lead Time of Factory Acceptance Testing of Large Bore Engines

Mechanical Engineering/Department of Mechanical and Materials Engineering

Master's thesis

Author:

Atte Rantanen

7.5.2026

Turku

Master's thesis

Subject: Mechanical Engineering
Author: Atte Rantanen
Title: Identification and Root Cause Analysis of the Key Bottlenecks Affecting the Lead Time of Factory Acceptance Testing of Large Bore Engines
Supervisors: University Lecturer Jari Böling, Mikko Kekkonen
Number of pages: 75 pages
Date: 7.5.2026

This thesis investigated the root causes of delays in the factory acceptance testing (FAT) process at a company manufacturing large engines. The objective was to identify why engines experience extended lead times during these tests and to propose improvements that could increase the throughput capacity of existing test cells infrastructure.

The study followed a mixed methods approach within a case study design, combining quantitative data analysis with qualitative interviews, hands-on training sessions at the factory, and a questionnaire. Root cause analysis was conducted using Fishbone and Five Why's methods. The quantitative analysis established that the LB3 engine represented the most significant bottleneck in testing and revealed through Pareto analysis that five issue categories – tightening errors, installation errors, pipe tightening errors, missing parts, and supplier quality – accounted for approximately 80% of all categorized problems.

The root cause analysis traced these issues back to an incorrect assembly order in the company's manufacturing control system, insufficient tightening procedures for pipe fittings and tight installation clearances in the engine design. The questionnaire validated these findings, with the majority of respondents confirming that unclear tightening procedure and part reinstallations caused by incorrect assembly order are the most significant factors contributing to leaking connections.

The study concluded that assembly work is the largest source of issues encountered during FAT. However, the errors are not caused by a lack of care from the workers but are enabled by systemic factors such as incorrect assembly instructions leading to unnecessary part reinstallations and tight installation clearances. Two main proposals were made: correcting the assembly order in the manufacturing control system and ensuring correct installation methods for the pipe fittings through training. These changes would target the identified root causes and deliver improvements in FAT lead time and lead time variation.

Key words: root cause analysis, factory acceptance test, Pareto analysis, data analysis, questionnaire, lead time, Five Why's, large bore engine.

During the preparation of this work, the author used GPT-5 by OpenAI and Claude Opus 4.6 by Anthropic large language models to check and improve English language expressions. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication. [1], [2]

Table of contents

1 Introduction 6

1.1 Background of the work 6

1.2 Goals and scope 7

1.3 Research plan 8

2 Literature review 9

2.1 Engine manufacturing and testing 9

2.1.1 Production and operations management 9

2.1.2 Factory acceptance testing 11

2.2 Root cause analysis methods 13

2.2.1 Overview 13

2.2.2 Fishbone diagram 14

2.2.3 Five Why's technique 16

2.3 Research methodology 19

2.3.1 Research design 19

2.3.2 Quantitative methods: data analysis 21

2.3.3 Qualitative methods: interviews and participant observation 23

2.3.4 Pareto analysis 25

2.3.5 Thematic analysis 26

3 Methodology 28

3.1 The topic of the thesis 28

3.1.1 Motivation 28

3.1.2 Thesis meetings with company supervisor 29

3.1.3 Thesis meeting with key stakeholders 30

3.1.4 Thesis meeting with company and university supervisors 31

3.1.5 Understanding the topic 31

3.2 Data analysis 32

3.2.1 Data sources 33

3.2.2 Data cleanup 35

3.2.3 Problems during the data preparation 38

3.3 Hands-on training and interviews 38

3.3.1 Training at large bore engine assembly 39

3.3.2 Training at Testing and Validation 40

3.4 Interviews 41

3.5 **Thematic analysis** 43

3.6 **Root cause analysis: Fishbone and Five Why's method** 44

3.7 **Questionnaire** 44

4 **Results** 46

4.1 **Results from the data** 46

4.1.1 Why did the work focus on LB3? 46

4.1.2 What are the most frequent issue categories? 51

4.2 **Discoveries from interviews and participant observation** 53

4.2.1 Operations control 54

4.2.2 Existing processes effect on factory operations 55

4.2.3 Facility's effect on factory operations 57

4.2.4 Personnel's effect on factory operations 58

4.2.5 Observations made about engines' design 62

4.3 **Root cause analysis of identified issues** 63

4.3.1 General overview of causes using a Fishbone diagram 63

4.3.2 Root cause analysis utilizing Five Why's method 65

4.3.3 Questionnaire results 70

5 **Conclusion and discussion** 73

5.1 **Proposals for the case company** 74

5.1.1 Proposal 1: Correcting the PMCS assembly order 75

5.1.2 Proposal 2: Ensuring correct installation of Ermeto fittings 76

5.2 **Discussion** 76

5.2.1 Contextualization within existing literature 77

5.2.2 Limitations of the study 78

5.2.3 Reliability and validity 79

5.2.4 Future research 79

References 81

Appendices 85

Appendix 1: Excel formulas 85

Appendix 2: Thematic analysis themes and codes 87

Appendix 3: Questionnaire 88

List of symbols and abbreviations

CRISP-DM	Cross Industry Standard Process for Data Mining
CTE	Chief Technical Expert
Data 1	Issue data gathered from cell specific Teams channels
Data 2	Test slot data gathered from production control software
FAT	Factory acceptance test
FIFO	First in, first out
GM	General Manager
LB	Large bore
MAR	Missing at random
MB	Medium bore
MCAR	Missing completely at random
NMAR	Not missing at random
PCT	Product conformity test
PMCS	Proprietary manufacturing control software
POM	Production and Operations Management
R&D	Research and development
RCA	Root cause analysis
Re-run	Removing the engine from the test cell for repairs and then repeating the test.
SB	Small bore
SSI	Semi-structured interview
T&V	Testing and Validation
TPS	Toyota Production System
UAT	User acceptance test
VSM	Value stream mapping

1 Introduction

The global demand for energy continues to grow, and with it, the need for reliable power generation and marine propulsion solutions [3]. The global marine internal combustion engine market was valued at USD 9.28 billion in 2025 and is projected to grow to USD 16.54 billion by 2031 [4]. In this growing market, the ability to deliver products on time is an important competitive factor. The case company has recently experienced an increase in demand, creating a need to raise production throughput. This can be achieved either by increasing infrastructure, or by improving the efficiency of existing operations. Since infrastructure changes are slow and expensive, the focus shifts to finding ways to make the current processes more efficient.

1.1 Background of the work

Factory acceptance testing (FAT) is a critical step in the production of large engines. It is the final verification that the engine meets its performance and quality specifications before it is delivered to the customer [5]. Because every engine must pass this test, the test cells where the FATs are conducted can easily become the bottleneck that set the pace for the whole operation. If engines experience faults during testing, they occupy the test cell longer than planned, delaying subsequent engines and so affecting the whole production chain.

This thesis originates from a practical need within the case company to understand why engines spend more time in the test cells than necessary and what can be done about it. The company had observed that a significant portion of testing time is consumed by troubleshooting and repairing faults that appear during the tests, rather than by the testing itself. In some cases, engines must be removed from the test cell entirely, repaired, and then re-tested, which effectively doubles the use of the limited test cell capacity for that engine.

Addressing these inefficiencies is important for two reasons. First, improving the utilization rate of existing test cells is far more cost-effective than building new ones, which requires large capital investments and long lead times. Second, even if market conditions would change in the future, a more efficient testing process would still reduce operational costs and improve delivery reliability and make the process planning easier. The topic is therefore relevant both in the current high-demand environment and in the long term.

1.2 Goals and scope

The primary goal of this thesis is to identify and categorise the issues that delay factory acceptance testing of large bore engines at the case company's factory, and to find the root causes behind those issues. Ultimately, proposals should be made based on the findings that would reduce the testing lead times.

The following objectives were set to achieve these goals:

1. Identify and categorize the issues that cause delays during engine testing.
2. Determine which issue categories have the largest impact on testing lead times.
3. Investigate the root causes behind the most critical issues.
4. Propose improvements to reduce lead times, reduce lead time variation, and reduce the need for re-runs.

Since faults appearing during testing can originate from any earlier stage of production, understanding the full manufacturing process is necessary. For this reason, the scope of the work covers the entire engine assembly process at the factory, from the beginning of assembly through to the completion of the final FAT. The case company has divided its engine portfolio into three categories based on the relative bore sizes (cylinder diameter) of the engines: small bore (SB), medium bore (MB), and large bore (LB). However, to keep the work feasible within the given timeframe, the scope is limited to the large bore engine lineup, with a specific focus on the LB3 engine. This restriction was agreed upon with the case company's management and is justified by the LB3's significance in the current product portfolio.

The work does not try to find causes for issues related to test cell infrastructure, such as the physical condition or availability of the facilities themselves, but observations regarding these topics will be noted. Similarly, supply chain issues occurring before parts arrive at the factory are excluded, as addressing those would require a separate study with access to supplier data. The focus is on problems that originate within the company's own processes and that can realistically be improved through internal changes.

1.3 Research plan

The first step in the thesis process is to create a research plan. The plan is used to guide the completion of the project and to assure a successful and on-time delivery of the work. The most important part of the process is to develop a solid plan at the beginning of the work and then use it as a guide as the study progresses. The work starts with arranging multiple meetings with different stakeholders within the case company to come up with the scope, goals, and research questions for the work. During these meetings, possible problems that might arise during the process are also discussed.

Research questions:

1. What are the reasons for the delays in FAT?
2. Does assembly work contribute to these delays?
3. How can we mitigate these issues?

The plan to reach the goals is as follows. The work begins with learning the manufacturing and testing processes of large bore engines by using the material available in the case company's internal systems. Once a general understanding has been developed, the data found in these systems is analysed to get an idea of what the issues are and how significant their impact is on testing lead times. After the data analysis, hands-on training sessions are conducted at the factory, where the author's understanding of the processes is deepened and simultaneously the root causes behind the issues are investigated by interviewing the employees. Lastly, a questionnaire is conducted to validate the findings made during the hands-on training sessions. Once it can be shown that certain problems are real and their causes are understood, improvement proposals are created to address them.

The overall approach of this work is exploratory and combines quantitative data analysis with qualitative methods. This mixed approach is chosen because the data alone can show what the problems are and how often they occur, but understanding why they happen requires input from the people involved in the processes. By combining these perspectives, the work aims to provide both evidence-based prioritization and actionable root cause insights.

In this thesis, a theoretical foundation is first built, which is then utilized later as the actual research work about the issues in factory acceptance testing is presented.

2 Literature review

This chapter reviews the theoretical foundations and prior research relevant to the investigation of delays in the factory acceptance testing process. The review is organized into three main sections. The first section covers engine manufacturing and testing, introducing the key production and operations management metrics used to evaluate manufacturing system performance, as well as the factory acceptance testing process. The second section examines root cause analysis methods, focusing on the fishbone diagram and the Five Why's technique, which together form the analytical framework used in this thesis to trace delay-causing issues back to their fundamental origins. The third section presents the research methodology, including the case study design, the mixed methods approach combining quantitative data analysis with qualitative interviews and participant observation, and the Pareto and thematic analysis methods used to process and interpret the collected data. Together, these theoretical and methodological foundations provide the basis for the empirical investigation carried out in this thesis.

2.1 Engine manufacturing and testing

This section provides an overview of the key concepts related to engine manufacturing and testing that are relevant to this thesis. It begins by introducing fundamental production and operations management metrics – throughput, capacity, cycle time, and lead time – which are used to describe and evaluate the performance of manufacturing systems. The section then moves on to factory acceptance testing, which represents the final stage of the engine manufacturing process and serves as the quality gate before delivery to the customer. Understanding both the production metrics and the testing process is essential for analysing the delays and inefficiencies investigated in this thesis.

2.1.1 Production and operations management

Production and operations management relies on several fundamental metrics that help to describe and evaluate how a production system performs. These metrics – throughput, capacity, cycle time, and lead time – are closely related to each other and understanding them is important for anyone who wants to improve the efficiency of a production process.

Throughput describes how fast a production system is able to produce output during a certain period of time. In manufacturing, it is one of the most important indicators of how well a production line is performing. According to Yamashita et al. [6], the throughput of a production line depends heavily on the capacity of the buffers between the workstations. Their research shows that when buffer capacities are allocated properly, it is possible to reach a desired throughput level for the whole line. This means that throughput is not only determined by the speed of individual machines, but also by how the system is designed as a whole, including the storage spaces between the processing stages [6].

Capacity is closely connected to throughput, as it defines the upper limit of what a system or a workstation can produce. Yamashita et al. [6] explain that in production lines with unreliable machines, the actual throughput is often lower than the theoretical capacity, because machines can break down or have other disruptions. In their study, they present a method for finding the right buffer capacity allocation to achieve a target throughput level. This highlights the fact that managing capacity is not just about individual machines, but also about making sure the whole system is balanced and that buffers are large enough to handle variability in the process [6].

Cycle time refers to the actual time that is needed to complete one unit or one step in a process. Tyagi et al. [7] make a clear distinction between cycle time and lead time in the context of product development. In their case study, they found that the total lead time for the studied process was 620 days, while the cycle time, meaning the actual time spent on value-adding work, was only about 122.55 days. This large difference shows that a significant portion of the time in the process is not spent on productive work, but rather on waiting, rework, or handling duplicate information [7]. Value stream mapping (VSM) was used in their study as a method to visualize the process and find where the non-value-added time exists, with the goal of reducing it [7].

Lead time covers the total time from the start of a process until it is finished, including all the waiting and delays that happen along the way. Tyagi et al. [7] describe lead time as the "total time in system" and their research demonstrates how lean tools such as VSM can be used to identify and eliminate waste in order to reduce lead time. In their case study, the application of the proposed VSM framework was expected to reduce the product development lead time by approximately 50% [7].

These metrics are not independent from each other, and their relationships can be described mathematically. One of the most well-known formulations is Little's Law, which states that the average number of items in a system, $E[N]$, is equal to the arrival rate λ multiplied by the average time each item spends in the system, $E[T]$: $E[N] = \lambda E[T]$ [8]. Keilson and Servi [5] take this idea further by showing that in many queueing systems, the connection between the number of items and the time spent in the system goes deeper than just averages. They show that in systems where items arrive randomly (Poisson arrivals) and are served in the order they arrive (FIFO), the number of items in the system doesn't just match the time in average – it also follows a related probability distribution [8].

Why does this matter in practice? As Keilson and Servi [8] explain, if you know how the time in the system is distributed, you can directly figure out how the number of items in the system is distributed without needing extra calculations.

So, throughput, capacity, cycle time, and lead time are fundamental concepts that together describe the performance of a production system. Understanding how these metrics relate to each other is essential for the identification of bottlenecks, the elimination of waste, and the improvement of overall system efficiency.

2.1.2 Factory acceptance testing

Before a manufactured product is delivered to the customer, it must be verified that the product functions as intended and meets the agreed-upon requirements. A factory acceptance test is a quality assurance procedure designed to accomplish exactly this: it is a structured test conducted at the production site to confirm the functional integrity of the product before it leaves the factory. Krause et al. [5] describe the FAT as an end-of-line test that extends the standard quality control measures performed during manufacturing, enabling comprehensive functional verification directly at the production site. They emphasize that standard quality control procedures – such as visual inspections – can verify mechanical integrity but cannot always detect functional defects, which is why a dedicated FAT is needed [5].

The FAT occupies a specific position within the broader quality assurance chain. In their description of a multi-stage quality assurance process, Krause et al. [5] explain that the FAT takes place after the manufacturer's own inspections and before the product is shipped to the end user. Once the product arrives at its destination, a separate user acceptance test (UAT) may be performed to verify that no damage has occurred during transport [5]. The FAT therefore serves as the final functional checkpoint at the factory, and its role is to enable early fault detection and short repair cycles so that only verified products are delivered [5]. Experience from using FATs in production has demonstrated their capability to identify assembly and component issues at an early stage, allowing defective units to be repaired and retested before shipment [5].

At the case company, the factory acceptance test is carried out as an overall check of the manufacturing quality and to establish that the contractual commitments have been fulfilled [9]. The test is conducted in accordance with ISO 15550, and the results are recorded in a test protocol signed by the manufacturer, the customer (if present), and a representative of a classification society [9], [10].

The FAT at the case company comprises two main parts: a running-in phase and an official FAT phase [9]. The purpose of the running-in phase is to ensure that possible manufacturing defects and leakages are eliminated and that all bearing surfaces and piston rings are run in smoothly before the engine is delivered. This phase also serves to verify the functionality of the engine control system [9]. The running-in begins with a bearing run, after which the engine is loaded in steps up to full load [9].

Once the running-in is completed, the test proceeds to the official FAT phase, which follows the programme stated in the contract and the requirements of the classification society concerned. During this phase, the engine is operated at multiple load points in different fuel modes (based on the model), and a range of performance parameters are measured in accordance with ISO 15550 [9], [10]. Safety systems are also tested, including shutdown functions for critical parameters such as lubricating oil pressure and overspeed protection [9]. Additionally, fuel changeover capability (if available) is verified by switching between diesel and gas operation at the highest possible load [9].

After the test run, the customer and the classification society representative are invited to supervise a physical inspection of the engine. Standard inspection activities include opening selected bearings and visually inspecting the crankcase and camshaft compartment [9]. Additional tests – such as noise, vibration, or emission measurements – can also be performed if specifically stated in the contract [9].

The FAT is therefore a comprehensive process that covers not only the functional performance of the engine but also the verification of safety systems, control systems, and the mechanical condition of internal components. Any issues discovered during the test must be addressed before the engine can be approved for delivery. It is this process, and the delays that can occur within it, that forms the central focus of the empirical investigation in this thesis.

2.2 Root cause analysis methods

Root cause analysis (RCA) is a systematic problem-solving approach that seeks to go beyond the surface-level symptoms of a problem to identify and address its fundamental, underlying causes. The basic premise of RCA is straightforward: to solve a problem, one must first recognise and understand what is truly causing it [11], [12]. Rather than treating visible symptoms, which may temporarily alleviate the situation but allow the problem to persist, RCA aims to trace the chain of causation back to its origin, so that targeted and lasting corrective actions can be implemented.

2.2.1 Overview

The central purpose of RCA is to identify and eliminate the fundamental causes of problems to prevent their recurrence. As Dew and Sproull emphasise, if the real cause of a problem is not identified, then one is merely addressing the symptoms, and the problem will continue to exist. For this reason, identifying and eliminating root causes of problems is of utmost importance [13], [14]. The Canadian Root Cause Analysis Framework provides a comprehensive description of the purpose of RCA, defining it as “an analytic tool that can be used to perform a comprehensive, system-based review of critical incidents. It includes the identification of the root and contributory factors, identification of risk reduction strategies, and development of action plans along with measurement strategies to evaluate the effectiveness of the plans” [15]. Thus, RCA is not just a diagnostic exercise, but it is a structured process that extends from problem identification through to the development and evaluation of corrective measures.

The philosophical idea of RCA is the belief that problems are best solved when the underlying causes are identified and addressed, rather than just trying to fix the observable symptoms. Mahto and Kumar describe this as “the essence of root cause analysis”: to solve a problem, one must first recognise and understand what is causing it [12]. They also say that RCA “is an important component of a thorough understanding of ‘what happened’”. The process begins with an initial understanding of the event, identifying unanswered questions and information gaps, gathering information through interviews and observation, and synthesising this into a final understanding, which is then used to begin the “why” portion of the analysis. This iterative questioning continuously asking “why” reflects the core philosophy of RCA: tracing each cause that led to the previous cause, continuing until the trail can be traced back no further, with each result of a cause clearly flowing from its predecessor [12]. RCA is thus inherently a systems-based approach, seeking to understand causal chains in their entirety rather than isolating individual failure points.

Defining the concept of a “root cause” is essential to understanding RCA. According to Mahto and Kumar, “a root cause is the most basic reason for an undesirable condition or problem” [12]. This definition implies that in addition to the immediately observable causes to a problem, there exists a deeper, more fundamental factor (the root cause) whose elimination would break the causal chain and prevent the problem from recurring. The root cause draws a line between the obvious or partial causes and the actual starting point of a problem. The Canadian Root Cause Analysis Framework supports this definition by differentiating between “root” and “contributory” factors, recognising that while multiple factors may contribute to an incident, the root cause represents the primary, most actionable point of intervention [15].

2.2.2 Fishbone diagram

In this thesis, the fishbone diagram is used as the first step during the root cause analysis. The fishbone diagram, also known as the Ishikawa diagram or the cause-and-effect diagram, provides a structured and visual approach to systematically identify, organize, and analyse the underlying causes contributing to a specific problem.

The fishbone diagram was developed and popularized in the 1960s by Dr. Kaoru Ishikawa (1915–1989), a Japanese professor and quality management pioneer who made significant advancements in quality improvement while working with the Kawasaki shipyards [16]. The diagram derives its common name from its visual resemblance to the skeleton of a fish [16], [17]. It is recognized as one of the seven basic quality tools used in organizations, alongside histograms, Pareto diagrams, correlation diagrams, control charts, data stratification, and brainstorming [16]. Today, the Ishikawa diagram remains widely used across industries and domains for diagnosing problems and identifying root causes [16], [17].

The fishbone diagram is a visual aid that displays the relationships between various factors contributing to a particular effect or problem. The structure consists of three main elements: the “head” of the fish, which represents the focused problem statement; the long bones, which represent the main cause categories and their relationship to the problem; and the short bones, which represent contributing factors and specific sub-causes related to each main category [17].

The fishbone diagram is commonly used in conjunction with the “Five Why’s” technique, this work included, which involves repeatedly asking “why” an issues occurs in order find the fundamental root cause of a problem [17].

The construction and use of a fishbone diagram follow a systematic process. Liliana [16] divides it into four principal steps:

1. Identify the problem.
2. Work out the major factors involved.
3. Identify possible causes.
4. Analyse the diagram.

However, there can be variation to these steps based on the problem at hand. For example, Kumah et al. [17] have provided a more detailed procedure which includes six steps instead of four.

A critical step in constructing a fishbone diagram is the selection of the main cause categories, as these define the analytical framework for the investigation. In the classic cause-and-effect diagram, the usual categories are materials, methods, equipment, environment, and people [17]. However, the literature shows that the categories are highly adaptable and can be tailored to the specific domain and problem under investigation. Liliana [16] provides a comprehensive overview of various category configurations that have been used in different Ishikawa diagram applications, including but not limited to:

- People, Methods, Machines, Materials, Measurements, and Environment.
- Material, Personnel, Process, External factors, and Management.
- Software, Users, Hardware, and Environment.
- Equipment, Process, People, Materials, Environment, and Management.
- Working conditions, Raw materials, Management, Technology, Machine, and Workers.

This flexibility in category selection makes the fishbone diagram a highly versatile tool. The users are encouraged to select categories that are most relevant to the specific problem context, ensuring that all significant contributing factors are captured.

The fishbone diagram offers several advantages as an RCA tool. It enables a structured team process for identifying the underlying factors and causes of problems, and it allows the team to focus on root causes rather than symptoms [17]. The tool is excellent for capturing team brainstorming output and provides a sequential, organized view of contributing factors [16]. Creating the Ishikawa diagram in a more detailed form has the advantage of offering the possibility to identify and analyse all factors related to the studied problem [16].

However, certain limitations should be acknowledged. The diagram deals with the direction and categorization of causes but does not address quantity or the relative significance of individual causes [16]. Also, it can get very complicated for complex problems, and it can be hard to show how causes in different categories are connected to each other [16]. Additionally, if there is more than one root cause, it may be necessary to address each one with a separate fishbone diagram process [17].

Despite these limitations, the fishbone diagram remains a foundational and widely used tool in quality management and root cause analysis, and its structured, visual approach makes it well suited for the RCA conducted in this thesis.

2.2.3 Five Why's technique

In this thesis, the Five Why's technique is used in the final part of the root cause analysis. The Five Why's is a widely recognized method for determining the root cause of a problem by repeatedly asking the question "why" [18], [19].

The Five Why's technique originated in the Toyota Production System (TPS) and was developed by Sakichi Toyoda during the evolution of Toyota's manufacturing methodologies [19]. Taiichi Ohno, the architect of TPS, described the method as central to Toyota's scientific approach: "by repeating why five times, the nature of the problem as well as its solution becomes clear" [19], [20]. Ohno further designated the solution, or the "how-to," as "1H," coining the expression "Five Why's equal one how (5W = 1H)" [20]. Since its inception at Toyota, the tool has seen widespread use beyond its original context and is now employed within Kaizen, Lean manufacturing, and Six Sigma methodologies [19], [20]. Five Why's is considered a basic tool that should always be used when performing a root cause analysis, and it serves as a method for diagnosis that is essential for solving problems [18].

Barsalou et al. [18] define the Five Why's technique as a type of proximate cause strategy for RCA. A proximate cause strategy begins with an evaluation of the symptoms of the problem, after which the investigator seeks to identify each underlying cause [21]. The method functions as a cause-and-effect chain, where each successive "why" delves deeper into the causal structure of a problem [19]. Five Why's is described as "a repetitive questioning technique for probing deeper to surface the root cause of a problem", and it is well suited to identifying the cause of a failure through cycles of investigation and analysis [18], [22].

Although the technique carries the number five in its name, this figure is more of a rule of thumb; more or fewer iterations may be needed depending on the complexity of the problem and users should not feel constrained by the arbitrary number in the tool's title [18], [20]. As each successive "why" digs deeper, the questions become less superficial and increasingly difficult to answer [18]. A key structural principle is that the causal chain should make clear sense when read in reverse using "therefore" meaning that, the logic should hold when moving from the root cause back up to the problem statement [19].

An important part of the analysis process is determining when to stop asking "why." Barsalou and Starzyńska [18] have proposed several criteria that, when fulfilled, indicate the stopping point has been reached:

- A small investment can yield a significant improvement.
- A few causes are found to have a major impact.
- The root cause is within the organization's ability to correct.
- A person, policy, or process is identified as the root cause.

Importantly, once identified, root causes must be validated to confirm their accuracy [18].

A central advantage of the Five Why's technique lies in how straightforward it is to apply. By iteratively questioning the cause behind each answer, practitioners can uncover underlying root causes that would otherwise remain hidden, and this process can also point toward an effective corrective action [20]. The method encourages investigators to persist in their analysis rather than settling on an early or surface-level explanation for a failure, driving the inquiry past obvious symptoms toward a genuine understanding of the fundamental causes [18], [23]. In the absence of such a structured questioning approach, there is a risk that investigations conclude prematurely, for example, attributing a failure to human error and responding only with retraining, while the deeper systemic factors remain unaddressed [18]. Another notable strength is the versatility of Five Why's: it can function independently as a diagnostic tool or be integrated into larger structured problem-solving frameworks, including 8D reports, A3 reports, and Lean Six Sigma projects [18], [24], [25]. The practical benefits of the technique have also been demonstrated in practice. One documented case shows that applying Five Why's as part of an RCA effort resulted in annual cost savings of over $32,000 at a manufacturing facility [26].

Despite its widespread use, the Five Why's technique has been the subject of a number of caveats and critiques. Card [20] argues that the fundamental problem with Five Why's is that it focuses on only one root cause at the end of one causal pathway, providing a "blinkered view of the complex causal pathway" that led to an incident. Card [20] illustrates this by comparing a Five Why's analysis to a causal tree diagram: in one example, the tree diagram uncovered more than 75 causes and contributing factors, while the Five Why's approach would identify only one or possibly two, representing less than 3% of the opportunities for improvement. This inherently narrow, linear structure means that Five Why's may miss the breadth required for complex system failures [20]. Gangidi [19] identifies several additional shortcomings of the traditional technique:

1. A lack of tendency or interest of analysts to follow through with deeper questioning, thereby missing sub-system root causes.
2. Confusing a symptom with the actual root cause.
3. Inability or unwillingness of experts to think beyond their area of expertise.
4. Lack of repeatability of the results – different people could arrive at different root causes for the same problem by asking different questions.
5. Difficulty in applying known principles to new problems.

Card [20] echoes the repeatability concern, noting that it is unlikely any two teams would independently arrive at exactly the same results when using Five Why's on the same incident.

To address some of these limitations, Gangidi [19] proposes the 3 × 5 Why's technique, which extends the traditional method by adding three legs to the analysis (occurrence, human, and systemic issues) enabling the identification of deeper root causes that may span across multiple groups within an organization. This expanded approach aims to ensure that both the technical failure mode and the underlying human and systemic factors are systematically investigated [19].

2.3 Research methodology

This chapter presents the research methodology chosen for this thesis and explains the reasoning behind the methodological decisions. The overall research design is a case study with a mixed methods approach, which combines quantitative data analysis and qualitative interviews together with participant observation. The chapter also describes the thematic analysis and Pareto analysis methods that were used for analysing the qualitative and quantitative data. Throughout the chapter, the methodological choices are supported by relevant literature, and also the limitations of the chosen approaches are discussed.

2.3.1 Research design

This study follows a case study research design, which is well suited for investigating complex, real-life phenomena within their actual context. According to Flyvbjerg [27], context-dependent knowledge, such as that produced by case studies, is central to expert-level understanding in any field. Flyvbjerg [27] argues that beginners rely on general rules and theories (e.g., from textbooks), but to become a true expert, a person must accumulate hands-on experience with many real, concrete cases. A case study allows the researcher to get close to the studied situation and observe it as it develops in practice, which is especially useful when the aim is to understand in detail how a specific process works [27].

In this thesis, the case study approach was chosen because the research problem (delays in factory acceptance testing) is closely tied to the operational context of the case company. The issues cannot be studied in a meaningful way without understanding the specific processes, people, and systems that are involved. This is in line with what Flyvbjerg [27] describes as the strength of case studies: their ability to capture the complexity and nuances of real-world situations that more abstract or generalizable methods might overlook.

However, case study research has also faced criticism. One common concern is that findings from a single case cannot be generalized to a broader population. Flyvbjerg [27] responds to this by pointing out that through careful and strategic selection of the case (choosing critical or paradigmatic cases), a researcher can make conclusions from a single case that have relevance also beyond that specific case. Another concern that is often raised is that case studies are prone to subjective bias, where the researcher has a tendency to confirm their initial assumptions. Flyvbjerg [27] opposes this as well, noting that researchers who conduct intensive case studies often report having to revise their original hypotheses because the evidence contradicts them. He argues that the closeness to the subject in case study research actually promotes falsification rather than verification, suggesting that the method is self-correcting by design [27]. A third criticism is that case studies are difficult to summarize into general theories. Flyvbjerg [27] acknowledges this but argues that this difficulty is not a weakness of the method itself but rather reflects the richness and complexity of the reality being studied.

Within the case study, this thesis employs a mixed methods approach, combining quantitative and qualitative data collection and analysis. Mixed methods research draws on the strengths of both approaches: quantitative methods are useful for measuring the magnitude and frequency of phenomena, while qualitative methods are better suited for understanding the reasons and mechanisms behind them [28]. The combination of the two can provide a more complete picture than either approach could offer alone [28], [29].

Bryman [29] identifies several rationales for combining quantitative and qualitative methods. These include triangulation, where different methods are used to confirm the findings of each other; complementarity, where one method is used to elaborate or clarify the results from another; and development, where the results from one method are used to inform the design of the other [29]. In his analysis of 232 mixed methods articles, complementarity was found to be the most common way in which the two types of data were actually used in practice, followed by expansion, which involves using different methods to broaden the scope of the study [29].

In the context of this thesis, the quantitative data analysis was used first to identify and prioritize the most common issue categories causing delays, after which qualitative interviews were conducted to explore the root causes behind those issues. This sequential structure corresponds to what Fetters, Curry, and Creswell [28] describe as integration through building, where the results from one dataset inform the data collection of the other. Specifically, the quantitative findings served as the basis for formulating the interview questions used during the qualitative phase. Fetters et al. [28] also describe three other approaches to integration at the methods level: connecting, where one dataset links to the other through sampling; merging, where both datasets are brought together for joint analysis; and embedding, where data collection and analysis are linked at multiple points throughout the study.

Bryman [29] does note that in practice, there is not always a clear alignment between the stated reasons for combining methods and how the integration is actually carried out. In some cases, researchers do not explicitly state any rationale at all [29]. This shows the importance of being transparent about why and how the methods are combined, which this thesis aims to do by describing clearly the role that each method plays in answering the research questions.

2.3.2 Quantitative methods: data analysis

The quantitative data analysis conducted in this thesis follows loosely the structure of the CRISP-DM (Cross Industry Standard Process for Data Mining) model. CRISP-DM is an industry-independent process model that was originally developed by practitioners to provide a standardized approach for conducting data mining projects [30]. It consists of six iterative phases: business understanding, data understanding, data preparation, modelling, evaluation, and deployment [30]. Although CRISP-DM was designed primarily for data mining, its structured and iterative nature makes it applicable also as a general framework for organizing the work of data analysis.

In the business understanding phase, the goal is to assess the business situation and determine the objectives of the analysis [30]. In the context of this thesis, this phase corresponds to the meetings with the case company’s stakeholders, during which the key questions to be answered through the data were defined. The data understanding phase involves identifying data sources, describing their structure, and building familiarity with the data through methods such as descriptive statistics, data visualization, or expert interviews [30]. In this study, data understanding was achieved by exploring the company’s internal systems, scrolling through communication channels, and discussing data availability with employees. The data preparation phase covers tasks such as data selection, transformation, and cleaning to make the data suitable for analysis [30]. Schröer et al. [30] note that not all studies describe their data preparation tasks explicitly, but emphasize that it is important to do so, as documenting all preparatory steps makes the results reliable and traceable. The data preparation carried out in this thesis is described in detail in the following chapters.

It should be noted that this thesis does not apply all six phases of CRISP-DM, as the work does not involve predictive modelling or system deployment. However, the first three phases provide a useful structure for describing how the quantitative analysis was planned and executed.

Before any analysis can take place, the raw data must be prepared. Data preparation is a critical step because the quality of the data has a direct effect on the reliability of the results. As Kwak and Kim [31] explain, problems such as missing values and outliers can compromise the statistical power of a study and introduce significant bias into the findings. Similarly, Van den Broeck et al. [32] state that errors occur in research data regardless of how carefully a study is designed, and that data cleaning is needed to identify and correct these errors or at least minimize their impact on the results.

Van den Broeck et al. [32] define data cleaning as a process consisting of three stages: screening, diagnosing, and editing of suspected data abnormalities.

1. Screening: actively searching for potential errors in the dataset.
2. Diagnosing: determining whether a flagged value is a true error or a legitimate observation.
3. Editing: correcting or otherwise handling the confirmed errors.

They also highlight that errors can occur at any stage of the data flow – not only during the original measurement, but also during data entry, extraction, transfer, and transformation – including during the cleaning process itself [32]. For this reason, transparency and proper documentation of all cleaning procedures is important for maintaining the validity of the study [32].

In the context of missing values, Kwak and Kim [31] describe three types: missing completely at random (MCAR), missing at random (MAR), and not missing at random (NMAR). The type of missingness matters because it determines how the missing data should be handled. They recommend using more than one method for dealing with missing values so that results can be compared [31]. Regarding outliers, Kwak and Kim [31] describe them as extreme values that fall outside the overall pattern of a distribution. Outliers can result from various causes, including participant response errors and data entry errors. Three general approaches for treating outliers are described [31]:

- Trimming: removing them from the dataset.
- Winsorization: replacing their values or adjusting their weights.
- Robust estimation method: estimating the outliers with statistical approach.

However, since outliers are also observed values, simply excluding them may not always be appropriate [31].

In this thesis, Microsoft Excel was used to prepare and clean the data [33]. The cleaning process included, for example, verifying that all rows had a consistent formatting and that there were no typing errors in the data. The specific processes are described in more detail in Chapter 3.2.2.

2.3.3 Qualitative methods: interviews and participant observation

Interviews are the most widely used method for gathering qualitative data across academic disciplines. Thelwall and Nevill [34] found that interviewing is by far the most prevalent qualitative approach, with its use increasing substantially between 1996 and 2019. This prevalence is not surprising, as interviews are a natural method for obtaining data in people-focused research [34]. The interview method used in this thesis sits somewhere between unstructured and semi-structured interviews. In general, the interviews were quite open where the participants were able to lead the conversation suggesting unstructured definition. However, there were also some pre-generated guiding questions presented based on the data, which would suggest semi-structured interviews.

McIntosh and Morse [35] describe the semi-structured interview (SSI) as a method that employs a relatively detailed interview guide and is particularly suitable when objective knowledge about a phenomenon exists, but subjective knowledge is lacking. The interview questions are open-ended, and while all participants are asked the same questions in the same order, they are free to respond as they wish. The researcher may also probe the responses to seek further clarification or elaboration [35]. This combination of a predefined framework with flexibility in the responses is what gives the SSI its semi-structured character, making it unique among interview methods for the balance it provides between topical relevance and responsiveness to the participant [35].

McIntosh and Morse [35] also explain that because all participants answer the same questions, the data collected through SSIs are comparable and can be analysed systematically item by item. Furthermore, the SSI is not tied to any particular philosophical foundation, which makes it compatible with qualitative, quantitative, and mixed methods research designs [35]. This versatility is one of the reasons the SSI has proliferated and diversified since the 1990s, evolving from simple research strategy into an independent research method [35].

In this thesis, a set of guiding questions was prepared in advance based on the findings from the quantitative data analysis, providing the objective framework that McIntosh and Morse [35] describe as the basis for SSI question development. However, the participants were free to elaborate on their answers and raise additional topics, and the author asked follow-up questions as needed. This structure ensured that the conversations remained relevant to the research problem while still allowing for unexpected insights.

In addition to the interviews, this thesis employed elements of participant observation during the hands-on training sessions at the factory. Participant observation is a method closely associated with ethnographic research. Thelwall and Nevill [34] note that ethnography is used the most in the arts & humanities and social sciences, where it is used to gain insights into localized social or cultural contexts. Compared to interviews, ethnographic approaches are considerably rarer across most other fields [34]. Thelwall and Nevill [34] suggest that this may be partly because professionals in fields such as healthcare already have regular contact with their subjects, which can reduce the perceived need for dedicated participant observation.

In the context of this thesis, the participant observation was not a full ethnographic study, but it utilized some features from it: by being physically present and taking part in the daily work at the assembly and testing facilities, the author gained access to tacit knowledge and contextual understanding that could not have been obtained through interviews or from internal systems alone. This included observing the work routines, hearing informal conversations between employees, and seeing the problems as they occurred in practice.

A key feature of the qualitative data collection in this thesis was the decision to integrate most of the interviews into the hands-on training sessions rather than conducting them as separate, formal meetings. This decision was made because the employees working directly with the engines were considered to have the most practical understanding of the issues causing the delays, and the hands-on trainings were a natural environment for conducting the inquiry.

This approach of combining observation with interviewing meant that the qualitative data came from two types of sources: informal, context-driven conversations during the training sessions, and individual meetings with employees selected based on their roles within the organisation. McIntosh and Morse [35] note that SSIs can be administered in various modes and settings, including face-to-face, and that their flexible structure makes them adaptable to different research contexts. The approach used in this thesis took advantage of this flexibility by adapting the interview method to the practical realities of the factory environment.

After each conversation, the author took notes and later refined them into clear bullet points. These notes, together with the observations made during the training periods, formed the qualitative dataset that was subsequently analysed using thematic analysis (Chapter 3.5). Elo and Kyngäs [36] note that qualitative data can come from a variety of sources, and that the preparation phase of content analysis involves becoming immersed in the data to build a thorough understanding of it. This principle was followed in this thesis, as the author revisited the notes multiple times before beginning the coding process.

2.3.4 Pareto analysis

In the early 20th century, the economist Vilfredo Pareto stated that 20% of the population possesses 80% of the wealth [37]. This basic 80/20 principle was later adopted to describe other realities and refined into the “A”, “B”, “C” classification, where attributes are grouped by their relative importance [37]. In its general form, the Pareto principle states that for natural events, roughly 80% of the effects come from 20% of the causes, which are referred to as the “vital few” [38]. For example, 20% of part numbers in a bill of materials may account for 80% of the total value, or 20% of customers may generate 80% of a company’s revenue [37].

The central tool of this methodology is the Pareto chart, which is a frequency block diagram that displays the relative frequency of different attributes in descending order [37]. This means the most frequent events are placed on the left side and the least frequent on the right. A cumulative line graph is typically overlaid on top of the bars to show the running total as a percentage. A useful or ideal Pareto chart is one where approximately 20% of the attributes carry about 80% of the total weight, thereby clearly revealing the critical attributes that deserve attention [37]. However, Grosfeld-Nir et al. [37] point out that this is not always the case – sometimes the relative frequencies are distributed almost uniformly across the attributes, making it difficult to identify a small group of dominant causes. In such situations, the chart becomes less informative as a focusing tool [37].

In manufacturing, Pareto analysis has been applied to identify and prioritize defects. Raghuram et al. [38] used a Pareto chart to analyse end-of-line rejections in an engine assembly line. The analysis revealed the vital few defects that accounted for the majority of rejections, after which the root causes of those defects were investigated using cause-and-effect diagrams [38]. This illustrates a typical application of the Pareto methodology in practice: first the most significant problem categories are identified through classification, and then the improvement efforts are directed toward those specific categories. As Raghuram et al. [38] note, "end of line rejections and rework contribute towards increasing the component costs, which in turn affects the customer pricing or company profits." By focusing on the vital few defects, manufacturers can address the issues that have the largest impact first.

In this thesis, Pareto analysis was used in a similar manner. The defect and delay data collected from the case company's factory acceptance testing process were classified and visualized using a Pareto chart. This allowed the most frequently occurring issue categories to be identified, which then guided the focus of the subsequent qualitative investigation. The specific application of the Pareto analysis is described in Chapter 4.1.2.

2.3.5 Thematic analysis

Thematic analysis is a widely used qualitative analytic method for identifying, organizing, and interpreting patterns from interview data. Despite its widespread use, Braun and Clarke [39] note that historically the method has been poorly defined and has rarely been explicitly acknowledged as a distinct analytic approach. They argue that thematic analysis offers an accessible and theoretically flexible method for analysing qualitative data and recommend it as a foundational method that researchers should learn first, as it provides core skills applicable to many other forms of qualitative analysis [39].

A key characteristic of thematic analysis is its theoretical freedom. Unlike methods such as grounded theory or interpretative phenomenological analysis, thematic analysis can be utilized with various epistemologies and research questions [39]. This makes it compatible with a wide range of research designs, including mixed methods studies [40]. However, Braun and Clarke [39] caution that this flexibility also means that the researcher must be transparent about the choices made during the analysis – what is being done, and why – to avoid the “anything goes” critique [41].

Braun and Clarke [39] propose a six-phase framework for conducting thematic analysis, which has since become a widely adopted standard [40]. The phases are [39], [40]:

1. Familiarization with the data, where the researcher reads and rereads the material to build a thorough understanding of it.
2. Generating initial codes, where meaningful features of the data are systematically labelled.
3. Searching for themes, where codes are grouped into broader patterns.
4. Reviewing themes, where the candidate themes are checked against the coded data and the full dataset for coherence and consistency.
5. Defining and naming themes, where each theme is refined and given a clear description.
6. Producing the report, where the final themes are presented with supporting data extracts and linked to the research question.

Braun and Clarke [39] also distinguish between two levels at which themes can be identified. A semantic approach focuses on the explicit, surface-level meanings in the data, while a latent approach goes deeper to examine the underlying ideas, assumptions, or structures that shape what participants say [42].

In this thesis, the thematic analysis followed the six-phase framework described by Braun and Clarke [39] and was conducted at the semantic level, as the goal was to identify and describe the practical issues causing delays rather than to theorize about underlying ideologies or assumptions. The analysis was carried out manually, with the author coding the interview notes and observation records. Ahmed et al. [40] note that well-defined codes form the foundation for constructing themes, and that transparent and consistent coding practices are essential for ensuring the credibility of the analysis. These principles were followed throughout the process, and the resulting themes are presented in Chapter 4.2.

3 Methodology

This chapter presents the chronological progression of the work conducted with the case company. It explains how the thesis topic was selected and how the scope was defined through a series of meetings with different stakeholders in the company. It then describes how the necessary data was gathered from the company's internal systems and through interviews with employees. The chapter also covers the practical, hands-on training sessions used to build understanding of the topic and to investigate the issues causing delays in FAT.

3.1 The topic of the thesis

This chapter explains why the case company chose to research the topic and outlines the onboarding process that enabled the author to complete the work. The onboarding was especially important because, before starting, the author had no prior knowledge of the topic beyond a general understanding of the industry.

3.1.1 Motivation

Commissioning and testing of engines are a crucial part of the case company's operations. Recent changes in market conditions have led to a need to increase the factory's throughput capacity to meet the demand. This can be achieved in two ways: by increasing the number of available test cells, or by improving the efficiency of the existing ones through a higher capacity utilization rate. The problem with the first approach is that infrastructure changes are slow and expensive operations making it unsuitable for addressing urgent capacity needs, which is why the focus should be on the second approach: maximizing the utilization rate of existing test cells.

One potential way to boost the performance of the existing facilities, is to move from a three-shift system to a five-shift system, with testing staffed around the clock every day of the year. However, this would bring significant challenges, including higher labour costs and the need to recruit and train large numbers of employees all at once. To mitigate the need for this change, or at least to give the company more time to prepare, this thesis was initiated with the added benefit that the proposed improvements in FAT efficiency would remain valuable even during periods of lower throughput demand.

Prior to starting this work, the case company had already identified, that a lot of time is wasted in the test cells if the engines under testing have faults that need fixing. These faults create a further challenge: when some engines have issues and others do not, planning the testing schedule becomes difficult and inefficient. The optimal outcome of this study for the company would be, that the reasons for unnecessary time loss could be identified and then eliminated. If this could be achieved solely through changes to the internal processes, it would improve operational efficiency with little to no financial investment.

3.1.2 Thesis meetings with company supervisor

The thesis process started in August 2025 with a meeting between the author and the case company supervisor. During this first meeting the supervisor presented the author with multiple possible topics for the thesis which were discussed and ranked based on the preferences of the author. After the meeting the supervisor started to investigate if the topics could be approved by the management, but unfortunately in the next meeting the author was told that none of those topics could be used for the thesis. However, during this meeting the brainstorming continued which eventually led to the final topic about the factory acceptance tests and so the work started.

During the next meetings, the supervisor explained more about the delays in FAT and how they affected the case company. While explaining the problems, the supervisor also provided the author with valuable new information, as the author had no prior knowledge of FAT or of building large engines. These meetings were also used to iron out the practicalities involved in making this work for the case company. It was decided that the author would be employed by the case company for the duration of the work.

Last meetings before the contract started were used to define the research plan and schedule for the work as well as the scope for it. However, the scope couldn’t be finalized using the knowledge only from the author and the supervisor, for which reason, the last thing that was prepared was a presentation that would be presented to the management team of the function responsible for executing the test runs of the engines at the factory, to get feedback about the project and its scope.

3.1.3 Thesis meeting with key stakeholders

A first trip to the factory was arranged on the 13th of November 2025 to meet the stakeholders and get the final approval to start the work. In this meeting, the author presented the topic and proposed a research plan that would be used to investigate the issue. This plan included the objective of the work, business case, scope of the work, research questions and schedule. While presenting these, the stakeholders gave valuable input for the author, which helped to modify the work to be as helpful as possible to the case company while also being feasible to complete in the given eight-month time frame.

In this meeting the stakeholders expressed that the author should come up with a general picture of the production process across the whole engine lineup. Based on this, it should be explained that why the work focuses on LB engines and the LB3 specifically. In this meeting it was also mentioned that both PCT (product conformity test) and FAT are included in the scope. PCT is a lighter version of FAT, lasting a shorter amount of time and with less regulation, used more in the energy sector. The type of test does not affect the actual issues, which is why this work refers only to FAT, even though in reality there exist two types of tests.

The management talked about their desire to reduce the lead time in testing as well as the lead time variation of the tests. Also, re-runs, which involve removing the engine from the test cell for repairs and then repeating the test, were identified as a significant problem that would be beneficial to eliminate.

The management gave the author more details about where the information about the problems could be found and cell specific Microsoft Teams channels were mentioned as a potential gold mine for data [43]. In the meeting it was also agreed that the author would attend two hands-on training sessions at the factory. One at the engine assembly and one at the engine testing. The goal was to arrange this so, that the author could be a part of assembling one of the LB3 engines and after the engine was finished, the author could attend the testing of the same engine. In this way, the author could develop a practical and comprehensive understanding of the entire building and testing process. Lastly the deliverables for the case company were defined.

Deliverables for the case company:

1. Root causes of bottlenecks.
2. Proposals for reducing lead times.
3. Proposals for reducing lead time variation.
4. Proposals for avoiding the need to repeat FAT tests.

3.1.4 Thesis meeting with company and university supervisors

The last preliminary meeting before starting the work, was between the author, the case company supervisor and the supervisor from the University of Turku. The idea of this meeting was to familiarize each other and discuss about the expectations that the different parties have for the thesis. The practicalities were discussed regarding the writing process, and the responsibilities were shared among the supervisors. It became apparent, that most of the expectations for the work were from the case company's side. For this reason, and for university supervisor's time constraints, it was agreed that the main supervisor for the work would be the one from the case company, and the university supervisor would be responsible for making sure that the work meets the academic standards.

3.1.5 Understanding the topic

Once the practical arrangements were finalized, the research began. The initial task was for the author to deepen their understanding of the topic. To support this, meetings were arranged with personnel from across the organization. The goal of these meetings was to introduce the author to key employees that could have useful information regarding the study and ask questions, which were used to learn more about the topic of the work. During these meetings, the author already heard from the employees' point of view about what the most pressing issues are, and what are the causes for them. Already at this point, it was interesting to observe that employees from different sides of the organization described the same issues, but in a different way and the causes they suggested were different.

To enhance the learning process, a trip to the factory was arranged on 16–17.12.2025 to introduce the manufacturing facilities. During this visit the author met with many workers from different teams face to face. The goal was to find additional connections that could be useful when trying to find information about the problems causing the delays in FAT. Another reason for the visit was to give a general overview of the operation to make it easier to understand the findings in the future. During the visit, the author saw the test cells where the engines were tested and the factory floor where the engines were assembled. While showing the author around, the host from the case company informed the author already about various problems causing delays in the operations and what, in their opinion, should be done about them.

3.2 Data analysis

In a business setting, decisions can’t be made based on feeling. Hard data is needed to show that problems exist, and that they need to be fixed. The data also allows the company to identify the most critical problems and prioritize resources to address them first. This need for hard data was emphasized in the meeting with the stakeholders while they presented the two questions that would need to be answered based on the data. First, why should the author focus the study to the LB3 engine? Second, what are the most important issue categories affecting the engine testing lead time? These questions were the groundwork upon which the data analysis was built.

At this stage of the study, the author did not yet have a clear understanding of the steps involved in Testing and Validation (T&V), where the FATs are conducted. Some information was gotten from the preliminary meetings but those didn’t give any specific information. An easy way to learn more about the testing, was to scroll through the cell specific Teams channels that the management had mentioned. From here the author could see the day-to-day problems that the employees at the test cells were facing and get a general understanding of what are the steps involved in the engine testing.

FAT operation in a nutshell:

1. Prepare the test cell (e.g., install engine mounts).
2. Lift the engine to the test cell.
3. Connect all necessary pipe and electrical connections.
4. If not already installed, connect a generator to the flywheel.
5. Commissioning of the engine.
6. Running the FAT.
7. Disconnecting and lifting the engine from the cell.

Once a general understanding of the problems and processes was gained, the next step was to find the data that could be used to find the issues and categorize them. The author was offered many different systems to look for the data from, but there weren't any clear instructions on where to exactly find the data or what to even look for. The strategy from here was to see the questions set by the management and then try to find some data that could be linked to those requirements. This involved quite a bit of manual work going through different systems and different graphs within the systems. Most of these were hard to understand without knowing the context behind them. The amount of data available in the systems was vast, which had its pros and cons. The pros being that there was data to be analysed. On the other hand, the drawbacks were that the sheer volume of data made it difficult to find exactly what the author needed, and it raised the concern that analysing a particular dataset might duplicate work already done by others but not yet found in the system. Eventually, a decision had to be made to stop looking for more data and to start the analysis with the data that was found until that point.

3.2.1 Data sources

After all the searching, one of the data sources ended up being the cell specific Teams channels. These channels are the working environment that the employees in the test cells and the control room use to log different events during the T&V of the engine. The advantage of this data is that there is very little bureaucracy to this kind of logging resulting in, the channels including most, if not all, the problems the engines are facing during testing. This is exactly the data that is needed to answer the management's questions.

However, there was one major problem. The data was in Teams and no way in a form that could be analysed. One way to get the data would have been to copy all messages one by one to some table, but because there were thousands of them, that wouldn't have been possible. Luckily, during the preliminary meetings, the author found an employee who had transcribed the problems from the channels into an Excel spreadsheet. The data began from the last quarter of 2024 and continued all the way until the present day. The author got a permission to use this data. The data being available already in a spreadsheet made it viable to be used in this work. A snippet of this dataset can be seen in Figure 1, and it will be referred to as "Data 1" going forwards.

The Data 1 being available in Excel format as well as the nature of the data meant, that the most suitable way to process and analyse it, was to use Microsoft Excel. Another reason for this decision was the author's considerable prior experience with Excel, which allowed the author to just focus on the data rather than the software.

Note date	Project	Model	Engine type	Engine number	Problem description	Category	Message date
18.12.2024	********	SB2	SB2V1	********	********	Missing part	17.12.2024 21:34
18.12.2024	********	LB3	LB3V2	********	********		18.12.2024 12:04
19.12.2024	********	LB3	LB3V2	********	********	Installaiton error	19.12.2024 11:46
19.12.2024	********	LB3	LB3V2	********	********	Supplier quality	19.12.2024 14:28
30.12.2024	********	SB2	SB2V1	********	********	Supplier quality	21.12.2024 8:27
30.12.2024	********	LB4	LB4V2	********	********	Tightening	21.12.2024 10:16
30.12.2024	********	LB4	LB4V2	********	********	Tightening, pipes	21.12.2024 11:40
30.12.2024	********	MB2	MB2V5	********	********	Installaiton error	21.12.2024 13:20
30.12.2024	********	SB2	SB2V1	********	********	Supplier quality	22.12.2024 7:14
30.12.2024	********	LB4	LB4V2	********	********	Tightening, pipes	22.12.2024 7:28
30.12.2024	********	SB2	SB2V1	********	********	Test run error	22.12.2024 9:39
30.12.2024	********	MB3	MB3V2	********	********	Tightening	22.12.2024 10:07
30.12.2024	********	MB3	MB3V2	********	********	Tightening, pipes	22.12.2024 13:34
30.12.2024	********	MB3	MB3V2	********	********		22.12.2024 14:22
30.12.2024	********	MB3	MB3V2	********	********	Tightening, pipes	22.12.2024 14:57
30.12.2024	********	MB3	MB3V2	********	********	Tightening	22.12.2024 15:24
30.12.2024	********	MB3	MB3V2	********	********		22.12.2024 15:43
30.12.2024	********	LB4	LB4V2	********	********	Electrical connection	27.12.2024 9:04

Figure 1: Data 1: Issue description data from cell specific Teams channels (after cleaning).

Each line in this table represents a message that had been extracted from the Teams chats. Some of the columns had to be classified to protect the case company. This was done in practice by formatting the cells to show "********". The data included some basic logging information like the date the message was sent to the chat and the engine type as well as the engine number. The "Problem description" contained the actual message sent to the chat and the most interesting column regarding this research, the "Category", includes the issue type categorization that was made based on the said message. The categorization was made by the same employee that had created this table.

However, to be able to draw conclusions, for example how many issues the engines have on average, it was necessary to know how many engines there were in total during the given timeframe. It would have been possible to count the different engine types that appear in this data and then divide the number of issues with the number of engines. However, for this to be accurate, it would have been necessary to assume, that every engine can be found from this data which wouldn't be likely, considering the origin of the data. For this reason, the author didn't make this assumption to avoid the results being misleading. Next step was to find this other source of data where the true number of engines could be found. This data was eventually found from a production control software that the case company uses to control and plan its operations.

A snippet of this dataset can be seen in Figure 2, and it will be referred to as “Data 2” going forwards.

Test cell	Model	Engine type	Engine class	Engine number	Test slot type	Test start	Test end	Duration
Cell 13	MB2	MB2V1	MB	********	Customer ET (functional test)	23.12.2024	30.12.2024	8
Cell 17	LB4	LB4V2	LB	********	Customer ET (standard)	19.12.2024	31.12.2024	13
Cell 14	MB3	MB3V2	MB	********	Customer ET (standard)	21.12.2024	31.12.2024	11
Cell 15	SB2	SB2V1	SB	********	Customer ET (classification test)	17.12.2024	2.1.2025	17
Cell 10	MB2	MB2V5	MB	********	Customer ET (extended)	22.12.2024	2.1.2025	12
Cell 14	MB3	MB3V2	MB	********	Customer ET (standard)	2.1.2025	7.1.2025	6
Cell 10	MB3	MB3V1	MB	********	Customer ET (extended)	3.1.2025	8.1.2025	6
Cell 11	MB2	MB2V3	MB	********	Customer ET (standard)	4.1.2025	10.1.2025	7
Cell 14	MB3	MB3V1	MB	********	Customer ET (standard)	8.1.2025	10.1.2025	3
Cell 10	MB2	MB2V5	MB	********	Customer ET (standard)	8.1.2025	11.1.2025	4
Cell 13	MB2	MB2V1	MB	********	Customer ET (functional test)	31.12.2024	12.1.2025	13
Cell 13	MB2	MB2V2	MB	********	Customer ET (functional test)	12.1.2025	14.1.2025	3
Cell 14	MB3	MB3V1	MB	********	Customer ET (standard)	13.1.2025	15.1.2025	3
Cell 11	MB2	MB2V3	MB	********	Customer ET (standard)	11.1.2025	17.1.2025	7

Figure 2: Data 2: Engine test slot reservation data (after cleaning).

The Data 2 includes all past test slot reservations for different engines. These were marked in the data by stating the engine model and engine number, as well as the test slot start and end dates (Figure 2). Additionally, the data included the test cell where the test was carried out. With Data 2, the author could calculate the quantity of the tests. Since it was known that every engine needs to be tested, it was possible to extract the number of engines based on this information. It’s important to note that the total number of engines might be different than the total number of tests, because some engines might need to be tested more than once. More on this in Chapter 4.1.1.

3.2.2 Data cleanup

The table that is presented in Figure 2 isn’t what was found from the control system, but the final form that was reached after cleaning the data. The first step was to create a new Excel environment where all the data would be consolidated and later analysed. Next, the Data 1 was copied from its original file to this new environment. This was done because of a couple of reasons. First the information in the original file was constantly updated, which wasn’t desirable. Second, the data would need some cleaning which involved modifying and deleting selected lines. For these reasons, the data was copied to ensure the integrity of the original data. Next step was to extract the Data 2 from the production control software’s web user interface into the same Excel. Unfortunately, the software didn’t have a simple extract feature that could have been used to extract the data straight into a numerical format like .xlsx or .csv. However, export to PDF was available. For this reason, a workaround had to be used, and in this case, it involved exporting the data first into PDF and then using the Excel's built-in feature for importing data from PDF files. The data was imported one page at a time, creating multiple tables, after which, these tables were combined within Excel to create a one big table, including all the lines from the original data.

Now that all the data was in the same spreadsheet, the data could be cleaned. First thing that was checked, was that all the rows were in a same format. For example, in the Data 1 there were the date columns. It was necessary to check that Excel understood the dates correctly and that there weren't any typing errors. Checking this was a good choice since there were a hand full of lines that had indeed wrong formatting that Excel couldn't understand. These were then fixed by hand by utilizing the Excel's table filtering function to easily identify the "broken lines". If these lines would not have been fixed, the filtering wouldn't have worked correctly resulting in misleading findings in the future. Once the dates were correct, the whole table was sorted based on the "Message date" column.

Another major modification that was done to the data, aside from modifying the values, was to remove entire rows from the Data 1. There was a need to set a clear timeframe from where the data would be examined in the analysis. This decision was made based on two main requirements. First, like mentioned in the Chapter 3.2.1, the Data 1 had been gathered from 2024 onwards, so the amount of data was limited. Second, when trying to identify issues that needs to be addressed from historical data, it is important to note that some changes to the processes are always ongoing during the historical period. This means that some problems which have been relevant in the past, might have been already addressed, so it wouldn't make sense to include those in this analysis. To mitigate this issue, and because of the limited data available, a choice was made to include only data from the year 2025.

To restrict the Data 2 to 2025, the control software's built-in filter was used to select all slot reservations whose end date fell within a week belonging to 2025. A side effect of this approach was that some reservations begun in 2024, and a few even ended in 2024, because calendar week numbering makes those entries valid for the filter. This rule was adopted to consistently handle the year transition and determine which reservations to include. As a result, the final dataset contained some extra lines from 2024 (Figure 2), but it similarly omitted some reservations from late 2025, keeping the total analysis period effectively at one year.

After filtering the Data 2, it could be used to select the necessary lines from the Data 1. Since both sources include the engine numbers, those were utilized to link the two datasets. In practice this was achieved in the following way. An Excel function called “UNIQUE” was used in conjunction with “SORT” to list all unique engine numbers in alphabetical order from the Data 2 into column “A” (all Excel formulas can be found from Appendix 1). Next, the same was done with the Data 1 to column “B” (shown in Figure 3). At this point, the author saw that the lists had different lengths, which was expected and indicated that the lists contained different numbers of engines. To help visualize the next step, an automatic formatting rule was applied to highlight the rows where the engine numbers didn’t match (Figure 3). The last step was to go through the list manually, and every time there were a mismatch between the two columns, the lines with the corresponding engine number were deleted from the Data 1. Doing so removed it from the “Data1” column (Figure 3) shifting the numbers under it upwards, after which the next mismatch would be highlighted. After addressing all the mismatches, the two columns were exactly the same length including exactly the same engine numbers.

Before

Data2	Data1
20731	20731
25220	25220
26154	25569
26509	26154
35173	26509
58839	35173
61078	42581
72009	58839
	61078
	72009

After

Data2	Data1
20731	20731
25220	25220
26154	26154
26509	26509
35173	35173
58839	58839
61078	61078
72009	72009

Figure 3: Visualization: removing unwanted engine numbers from Data 1 (fake engine numbers).

Lastly, a couple of helper columns were added to the Data 2 to make the analysis easier. The following columns were added: “Model”, “Engine class” and “Slot duration” (Figure 2). These columns show simple information about the specific engine based on the other columns on the same row. For example, “Slot duration” calculates the duration of the reservation by first subtracting the start date from the end date and then adding one. At this point, the data was ready for analysis.

3.2.3 Problems during the data preparation

Getting the data ready for analysis required some problem solving. The biggest challenge was selecting the correct data. Like said before, there was a lot of data and a lot of different systems to choose from. The author not knowing beforehand which systems were in active use and which were not, meant that it was quite difficult to know which data sources could be trusted. It was also acknowledged that some sources could, for example, display the engine test scheduling plans, but it wasn't clear whether this data reflected reality. This was the case with the Data 2. To account for this, the author did ask various employees their opinions about this matter and the consensus was that the data could be trusted. Another way the data was verified was to pick random samples from the data and see if the dates matched the dates in the Teams chats. These samples also showed that the dates were correct or the difference was minimal, so the author decided to trust the data and use it in the analysis.

Other issues that arose during the work included the previously mentioned export problem with the control software and difficulties with the Excel formulas. There wasn't any specific problem that stood out during the generation of the formulas. A typical approach was to build functions from small, simple components and then combine them into one comprehensive formula. Occasionally, problems occurred when integrating the components, but these were resolved by searching solutions for similar issues from the internet.

3.3 Hands-on training and interviews

At the meeting on the 13th of November with the stakeholders, it was agreed that the author would take part in two hands-on training sessions at the factory. One at the engine assembly and one at the engine testing. Scheduling the sessions was left for the author to decide. The scheduling started by seeing what the best time from the side of the study would be and then checking from a planning software to see when the desired engine models were built and tested. The goal was to arrange these as soon as possible to avoid wasting unnecessary time in the beginning of the work. The engine that the author ended up following was one variation of the LB3 engine with multiple of them starting their assembly in the beginning of 2026. Given that assembling one of these engines typically takes approximately four to six weeks, and that the objective was to follow the same engine through assembly and testing, the sessions were held on 19–23 January and 16–20 February, respectively.

There were three distinctive reasons why these trainings were chosen to be especially useful for this study. First, being able to be a part of the day-to-day work at the factory, gave the author a unique perspective and opportunity to learn the inner workings and processes of the factory. Just being able to walk around the factory and see the different processes and hear the different conversations of the employees, gave a huge amount of silent information for the author that couldn’t have been learned from any internal system within the company. This helped the author immensely in understanding the research topic.

Second, since management chose that a more detailed study would be conducted about the LB3 engine, these trainings were a great opportunity to learn more about the specific engine. That is why the training was chosen to be carried out in an assembly team that was building one of these LB3 engines.

Third, in the planning phase of the work, it was decided, that interviews of the personnel should be carried out to find root causes for the issues found from the data. Prior to these training sessions, the data had already indicated that assembly work played a major role in issues observed during engine testing. Being present on the factory floor therefore provided an opportunity for the author to engage with assembly workers about these issues, to elicit their perspectives, and to explore the causes as they perceived them. This is why large portion of the interview data used in this work, was gathered during the hands-on training sessions at the factory.

3.3.1 Training at large bore engine assembly

The training at engine assembly started by checking that the author had the necessary security clearances and safety trainings completed to be able to move around the factory. After this, suitable work clothes were provided by the company. Next the author was taken to the assembly cell where the training would take place.

The engine in question was counterclockwise rotating LB3V3 inline engine. Next step was for the author start to get to know the personnel working with the engine and the engine itself. In practice this meant that the author would follow the team in their work and then ask relevant questions simultaneously. When the author joined the team, the engine was still in its early stages. The operation at hand was to install the pistons into the engine block. It was interesting to see how these massive pistons were installed utilizing a crane and multiple employees working on top of and under the engine simultaneously. The pace of the assembly was calm to assure that mistakes are avoided. For example, if the piston were to get jammed in the cylinder, the recovery from that could take a long time and possibly risk damaging the components in the process, because the forces involved are so large. During the training week, the team installed the pistons to the engine, roller tappet carriers, some piping and the engine was timed (camshaft was aligned with the crankshaft).

The atmosphere in the assembly was relaxed, and the author could see that they built the engine very thoroughly. The workers didn't speak much to the author but once spoken to, they were kind and open to talk about the issues the engines are having at the testing. During these discussions quite a few problems regarding the assembly were also brought up. This was interesting and should be investigated further to determine whether these problems have any effect on the issues encountered in T&V. Since this was a hands-on training, the author got to participate to the actual building of the engine. The tasks the author got to do included installing stud bolts, tightening the bolts on the roller tappet carriers and installing oil jets inside the cylinders. The guidance from the workers for these tasks was minimal, but since the author had previous experience working with machinery and tools, it wasn't a problem.

3.3.2 Training at Testing and Validation

Second hands-on training session was arranged in the T&V where the FATs are carried out. Unfortunately, the author could not attend the FAT for the same engine observed during the previous training session. However, there were two LB3 engines undergoing testing and one of them was the same kind that the author had been building. This time, the training wasn't as simple as in the assembly. In the testing there are multiple cells running different engines in different phases of their tests. This meant that the author wasn't assigned any particular test cell to follow, but instead, all the cells were followed to see where something was happening that could be useful regarding the study. Prior to going to the factory, it was known that the beginning of the week would probably be very quiet, because the engines were running their tests and the problems for those ones were already solved. According to the schedule, one of the LB3 engines should have been removed from the test cell and other one moved to its place on Wednesday, and for the other LB3 engine the same operation should have been on Thursday. Unfortunately, there were some issues with the first LB3 engine, which meant it could not be removed from the cell until Thursday.

On Thursday when the engines eventually got moving, there was a lot to follow for the author. The author chose to observe the LB3V4 changeover procedure, which proved to be useful, as the data showed that this engine variant had the most problems within the LB3 engine line-up.

The T&V process started with the installation of engine mounts to the test cell. These mounts sit between the engine stands and the engine absorbing the vibrations generated by the engine. Once those were in place, the main crane of the factory, with a lifting capacity of 400 tons, was used to lift the engine from its waiting position to the test cell. After the engine was successfully lowered onto the engine mounts, the preparation of the engine started. This involved attaching all necessary piping for oil, cooling water, fuel, and compressed air to the engine; connecting the generator to the engine flywheel to act as a load; and making the electrical connections required to control the engine and monitor its systems. Once everything was installed, fluids were pumped into the engine, and the running-in and testing period could begin.

In the factory, the workers run in three shifts, morning, evening and night shift. This meant that the author couldn’t be present during all the phases of the work. On Thursday, the author got to see how the engine was lifted into the cell, and the beginning of the installation of the piping. On Friday morning, the engine had been fully hooked up to the facility and the task at hand was to align the generator with the engine’s crankshaft. After that the distance between each pair of counterweights were measured to assure that the bend of the crankshaft was within specifications.

Similarly, as during the assembly training, the author spoke with multiple workers during the visit, gaining an overview of operations and the challenges they face. The employees were again kind and forthcoming when asked about these issues.

3.4 Interviews

The original plan was to have separate structured interviews with different employees to learn the root causes of the issues. However, as the work progressed, it became apparent that a more fitting approach would be to integrate most of the interviews to the hands-on training sessions. The reason for this was that the people facing the actual issues are the ones working with the engines physically, so the author assumed that they have a good, if not the best understanding about the scale of the issues. The author concluded that the most natural approach would be to interview participants while they were performing their duties, allowing them to discuss the issues as they arose. The slow-paced nature of the work made it possible for the workers to explain the task at hand and reflect on the bigger picture as well. This made it the perfect combination for the author to learn the operations effectively while gathering important data from the experiences of the employees.

Prior to these training sessions, the author had conducted the data analysis using Excel, which provided a preliminary understanding of where the issues were, and which types of issues were most common. This approach provided a basis for asking the employees a set of guiding questions.

Examples of the questions asked based on the findings from the data:

- How can pipe connections be loose in the testing, when they have been marked with a marker, that they have been tightened?
- How are the pipe connections tightened?
- Are the parts checked for any defects before installation?
- How the workers feel about the software used to guide the engines assembly? Are there any problems?
- How can an O-ring be damaged during installation?

Naturally, while the author was on site, many practical questions about the engines and installation procedures arose that could not have been prepared in advance. These questions proved even more useful than the prepared ones. After each conversation, the author took notes and asked participants for consent to use the findings in this study. These notes were later refined into clear and uniform transcription of the events.

In addition to these chats during the hands-on periods, there were four separate meetings with different employees from different sides of the organization. The structure of these meetings was to first introduce the topic and the author to the employee and then ask questions from them based on their area of expertise. Most of these were held before the hands-on periods so aside from learning about the issues, a major goal for these conversations was to give the author more information about the topic. Several valuable findings were already obtained from these meetings and therefore those will be included in the results as well.

It is important to explain how the interviewees were selected. For the individual meetings, employees were chosen based on their role within the organisation, informed by the company supervisor’s existing knowledge, with the aim of making the meetings as useful as possible for the author. Selection did not reflect any prior assumptions about how the participants would respond to the topic. By contrast, the interviewed employees in the hands-on training sessions were chosen purely at random, with no prior expectations. This approach was adopted to enable the author to build a general picture of the operations, and cherry-picking the teams for observation would have undermined that goal. It should be noted that findings from the training sessions might have differed had the teams under observation been different. However, this was a necessary compromise to make, to be able to conduct the study.

3.5 Thematic analysis

Thematic analysis was used to analyse the data gathered from the interviews [40]. Due to the nature of the data, the methodology had to be adapted little bit to fit the purpose. However, the basic structure of the analysis was kept the same (see Chapter 2.3.5 for more details). The analysis started with Phase 1 during which the notes from the meetings were revisited multiple times to build a general understanding as well as consolidate all of them into one document to make the Phase 2 easier. In Phase 2, initial codes were created. A new table was created in which the columns would represent the different codes. The notes were then categorized one by one by the author. Each note was reviewed to determine whether it fit any existing category, if so, it was added to the corresponding column; if not, a new category was created to accommodate it. Once all the notes had been categorized, the author moved on to the next phase. The Phases 3 and 4 were combined during which the original table was duplicated and the new version modified. Modifications involved modifying the original codes to be descriptive, as well as dividing the codes into themes. Additionally, some codes were merged to reduce unnecessary complexity that didn’t bring any new information to the study. Lastly, in the Phase 5, the naming of the codes was fine-tuned and descriptions for each theme were added. Eventually, there were twenty codes which were divided into eight themes. The final result can be found from Appendix 2.

3.6 Root cause analysis: Fishbone and Five Why’s method

The last step in the work was to conduct the root cause analysis. The author chose to use a combination of two methods for this. A cause-and-effect (Fishbone) diagram was used to build a general picture of the operation and its issues while Five Why’s diagrams were used to investigate the actual root causes [17], [19].

The Fishbone diagram was created first. The diagram was created by going through the notes and categories made in the thematic analysis phase and identifying different causes for the delays in T&V. These causes were then listed, after which they were categorized by the author into five categories: people, processes, issues, facility and engine. These categories became the branches of the diagram, and the causes were placed onto the branches. The finished figure can be found from Chapter 4.3.1 as Figure 11. [17]

The final step was to create Five Why’s diagrams. The most critical part of this process was selecting the issue statements for investigation. The author generated three issue statements from the “issue” branch, based on what was observed during the study, what the author considered most beneficial to investigate, and what was likely to yield the best results with the Five Why’s technique. Once the statements had been formulated, each of them was analysed using the technique, building the diagram one step at a time. For this, the author referred to previous notes from the study to answer the questions as accurately and truthfully as possible. Once the diagrams were complete, they were reviewed for grammatical consistency to ensure they adhered to the “Why?” and “Therefore” rules. The finished figures can be found from Chapter 4.3.2. [19]

3.7 Questionnaire

A questionnaire was conducted to validate the findings from the hands-on training, interviews, and root cause analysis. The purpose was to provide additional evidence for the identified issues so that the results would not rely solely on the author’s own interpretation. The questionnaire also helped to compensate for the limited number of people observed and the limited range of situations encountered during the earlier research phases.

The questionnaire was designed to be as short as possible to maximize the response rate, as the employees needed to be able to answer it during their workday. The questions were focused on validating the most significant findings from the research. Ultimately, the questionnaire was centred on leaking Ermeto fittings. This focus was chosen for three reasons: many of the identified problems revolve around these connections, their impact on lead time had been clearly indicated during the investigation, and the author had already gathered substantial domain knowledge about them during the hands-on trainings, which helped in formulating effective questions.

The target group for the questionnaire was the assemblers and assembly team coordinators, as they are the ones who actually install the parts. Because of this, the questionnaire was written in Finnish, which the author believed would yield the highest response rate. The questionnaire was built using Microsoft's MS Forms application and distributed to the entire assembly organization via email [44]. The full questionnaire can be found in Appendix 3.

The questionnaire was structured around three main questions: what are the most important factors causing leaks, how are the fittings tightened, and whether the workers have had to remove already installed parts due to poor instructions. Each main question (excluding last one) was followed by a specifying open-ended question, allowing the assemblers to elaborate on their answers if they wished.

In the first question, the assemblers were asked to select up to two out of six factors that, in their opinion, have the biggest impact on leaking connections. The options were: installation error, lack of training, manufacturing error, part reinstallation, hard-to-reach installation location, and wrong tightening method. The second question asked whether the assemblers usually tighten the connections using official manufacturer procedures or based on hand feel. The third question was a yes-or-no question asking whether the assemblers have had to remove already installed parts due to incorrect PMCS instructions.

4 Results

This section of the work talks about the findings from the data analysis as well as the observations made based on the hands-on training periods and the interviews. Some of the data in this section has been classified to protect the case company. This was done by removing certain axes from figures, replacing actual engine models with generic ones, and replacing values in tables with “********”.

4.1 Results from the data

The first thing this section will show, are the results gathered from the data. Here the analysis answers the questions that management set in the beginning of the work. These include the need to justify why the focus of the work is in the LB3 engine, and to identify the most frequent issue categories with Pareto analysis.

4.1.1 Why did the work focus on LB3?

To justify the focusing on LB3, it was necessary to show that this engine was indeed the biggest bottleneck, meaning that it had the biggest impact on how many engines went through the testing on yearly basis. To be able to conclude this, it was necessary to consider the constraints the factory posed to the operations. At the time of making this study, there were eight test cells which were divided based on the engine types they could support. The cells differed based on what fuels they could use to run the engines, but the most important difference, for this study, was the size of the engine they could handle. From those eight cells, there were two SB cells, four MB cells, and two LB cells. One of the MB cells could also support “small” (inline engine with eight or nine cylinders) LB engines so in principle there were 3 LB test cells.

The first results gathered from the data were basic statistics like how many tests were conducted during the year 2025 for different engine types. This was done by first extracting all different engine models in the data (utilizing the added “Model” column) and then counting every line in the Data 2 that included the model in question. This produced the number of tests per engine type which can be seen visualized in Figure 4.

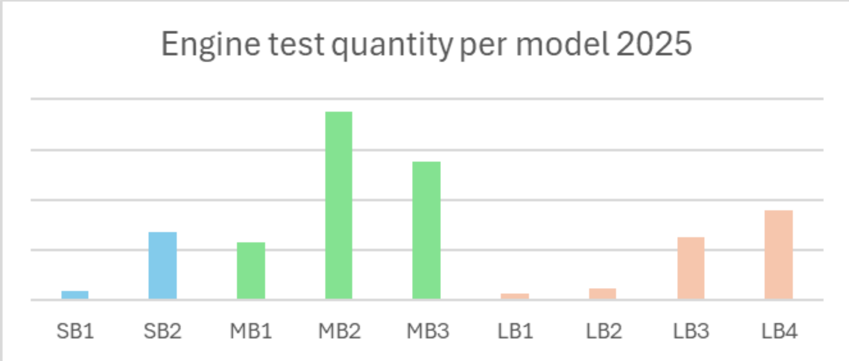

Figure 4: Engine test quantity per engine model 2025.

The Figure 4 shows, that majority of the tests in 2025 were MB engine tests, after which came LB and lastly SB. In the generic engine model naming scheme, the numbers after the class of the engines are used to identify different bore sizes within the same class. Here the first observation regarding the managements question is, that the LB3 engine wasn't the most frequently driven engine in 2025. However, it is necessary to take into consideration the duration of the tests. The Figure 5 shows the average test duration (days) for the different engine models. The figure shows that the even though there were fewer tests, the LB tests took significantly longer. This, combined with the fact that there were more MB testing cells than LB ones, gives a strong indication that there was a bottleneck in the LB testing capacity. Figure 5 also provides the first indication of why the LB3 engine was the main focus of the study. The figure shows that the average lead time of an LB3 engine was noticeably higher than that of the other MB and LB engines. It is important to note that while the test duration of SB engines was high, this was mitigated by the low number of engines.

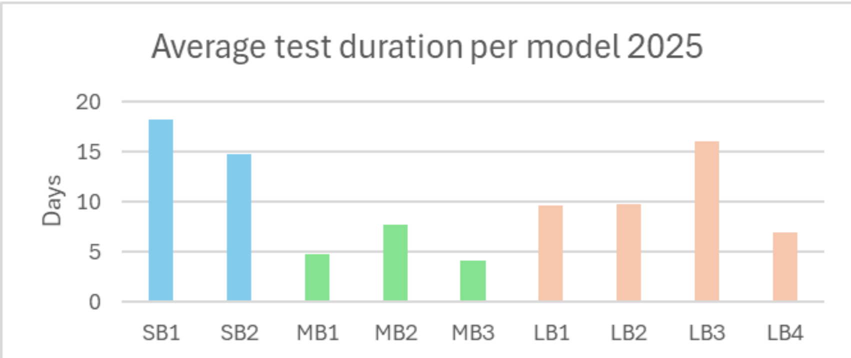

Figure 5: Average engine test duration per engine model 2025 (days).

To confirm the bottleneck, cell utilization was measured to determine whether the test cells were continuously in use. If utilization was below 100%, faster testing would not increase throughput – it would simply increase idle time. The resulting utilization metric is shown in Figure 6.

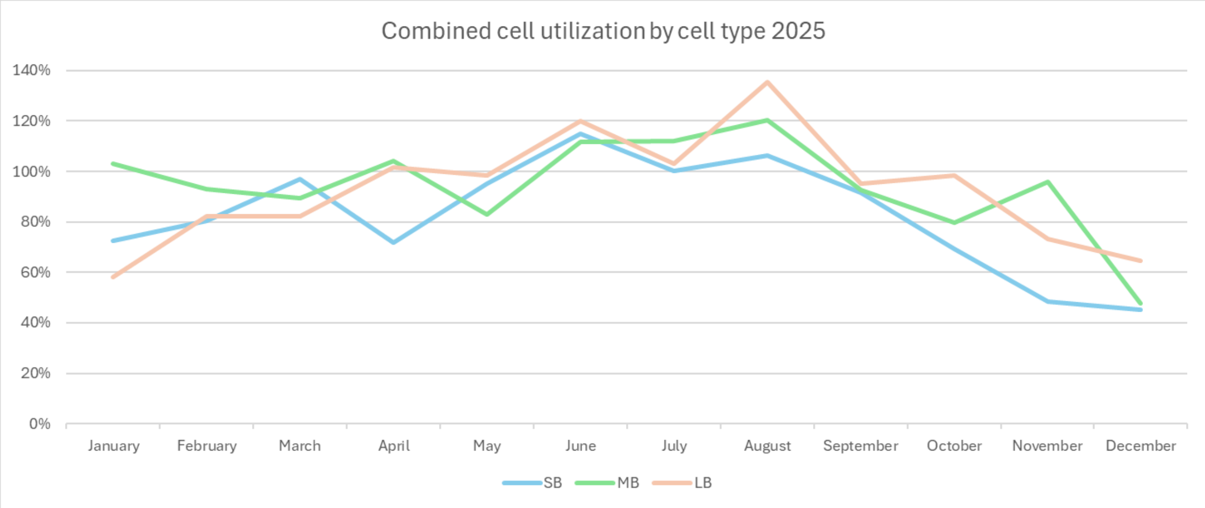

Figure 6: Cell utilization rate by cell type 2025.

A couple of notes from the Figure 6. First, the figure shows the sum of testing and maintenance reservations, that's why it is called "Combined cell utilization…". Second, there were some renovation works going on in the end of the year which wasn't marked in the data explaining the dip in the utilization in November and December.

The utilization metric was calculated by counting the number of testing days in each month based on the reservations recorded in Data 2. The testing days were then divided by the total number of days in that month, yielding a utilisation percentage. However, due to the resolution of the data, this metric doesn't depict reality perfectly. Since the reservations are marked with only start and end dates, the function must add one day to every reservation to avoid reservations with a length of zero (same start and end dates). As a consequence, if the reservations are short or their dates overlap (which they do), they may be counted twice, resulting in the utilization metric being too high leading to it going over 100% at some points.

Nevertheless, useful observations can still be made from this graph (Figure 6). The question about the utilization can be answered by using this. As shown in Figure 6, there has not been any more downtime in the LB cells compared to the MB ones. Therefore, it can be concluded that the only way to increase engine throughput in testing is to reduce the duration of the tests or to run the cells around the clock with full staff.

Getting back to the Figure 4. The graph shows the number of tests, but it doesn't consider the fact that some engines can be tested more than once due to re-runs or other special circumstances. To get the actual number of engines, it was necessary to remove duplicate engine numbers from this equation so that the same engine wouldn't be counted twice. This was done by utilizing "UNIQUE" function to remove the duplicates after which the "XLOOKUP" function was used to find the corresponding engine models. After that the quantity of every engine model could be counted.

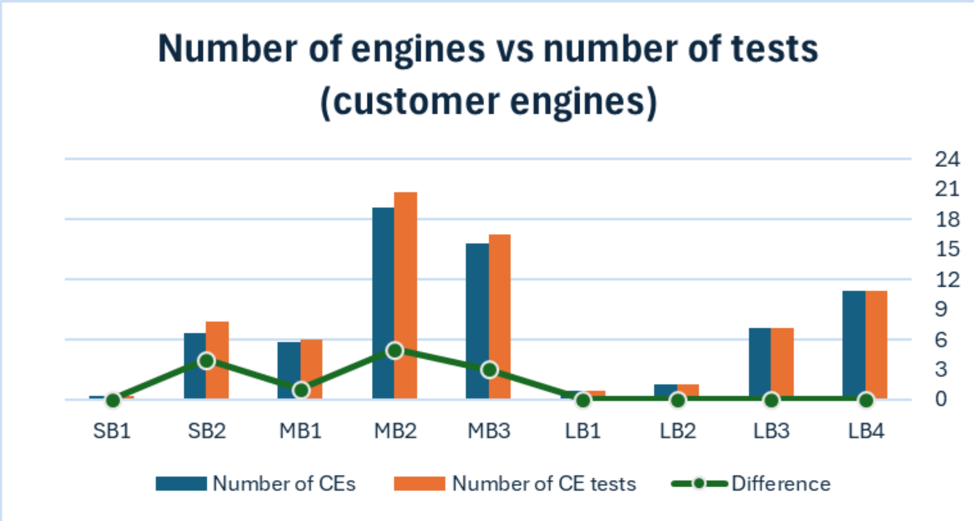

Figure 7: Number of engines vs number of tests (customer engines).

The Figure 7 above displays the difference between the number of tests and the number of engines. To help this visualization, a "Difference" line graph was added to the figure, which scale can be seen on y axis on the right side of the figure. The "customer engines" means that this analysis has excluded the reservations for engines that are purely for research and development (R&D) purposes. The reason for this was that those are not the focus of this study and those engines are naturally tested multiple times with different parts and parameters. The optimal situation would be that the difference between the variables would be zero effectively meaning that every engine is tested once and only once. However, this was not the case, the worst situation being with the MB2 and MB3 engines, with a delta of five and three, respectively. Although, the important observation from this data was that the LB3 engine didn't have any unnecessary tests contrary to the expectations made based on the first meeting with the management. It was mentioned that re-runs are a significant problem and that suggestions would be needed to address them. However, now that the scope has been set to focus on LB engines, re-runs do not appear to be as relevant to this study.

The management was asked about this finding, and the reason for this result is probably that the data just doesn't show the re-runs even though they have happened. For re-runs to be recorded in the data, the engine needs to be removed from the cell for repairs and then returned. Due to the sheer size of the LB engines, it is likely more practical to leave the engine in the test cell to be repaired rather than to remove it entirely. Unfortunately, this practice does not leave any trace of it in the data. This is still bad for the lead time, since a lot of time is wasted when the engine just sits in the cell waiting for repairs occupying the cell.

Next, Data 1 will be examined to determine whether any factors contributing to the extended lead times of the LB engines, and specifically the LB3 engines, can be found there. To investigate this, a table listing all distinct customer engine models was created. The table was generated by first extracting the unique engine numbers, which were then used to retrieve the corresponding models from Data 2. The models were then sorted by bore size and variation. The result can be seen in Figure 8 where all different customer engine models from 2025 are listed. Here the models have gained a variation indicator (Vx) after the bore size. This expresses different cylinder and fuel configurations that the same bore size engine can have.

Figure 8 lists all distinct engine models, their quantities, the number of issues per model, and the corresponding averages. On the right side of the Figure 8, there is the most common issue type for every model as well as their quantity. Using the average column, it was possible to compare the different models by the number of issues encountered. From this, it was evident that the LB3 engine had clearly the most problems on average compared to the other engine models. Based on this finding, combined with the longer lead time of the LB3 engine and the limited number of LB test cells, it was concluded that the most significant impact on testing lead times could be achieved by investigating the LB3 engine.

Found issues per engine type (customer engines)				Issues	
Model	Engines	Issues	Average	Most common issue	Amount
SB1V1	******	******	1.0	Electrical connection	******
SB2V1	******	******	6.1	Supplier quality	******
SB2V2	******	******	7.1	Tightening	******
SB2V3	******	******	3.0	Software	******
SB2V4	******	******	12.0	Test run error	******
MB1V3	******	******	6.0	Tightening, pipes	******
MB1V4	******	******	9.5	Supplier quality	******
MB1V5	******	******	8.3	Supplier quality	******
MB1V6	******	******	9.5	Tightening, pipes	******
MB2V1	******	******	6.8	Tightening	******
MB2V2	******	******	5.8	Supplier quality	******
MB2V3	******	******	8.3	Missing part	******
MB2V4	******	******	10.0	Tightening	******
MB2V5	******	******	3.0	Test run error	******
MB2V6	******	******	5.9	Missing part	******
MB3V1	******	******	5.2	Tightening	******
MB3V2	******	******	4.3	Tightening	******
MB3V3	******	******	5.0	Installation error	******
MB3V4	******	******	8.2	Tightening	******
MB3V5	******	******	3.8	Installation error	******
LB1V1	******	******	14.0	Tightening	******
LB2V1	******	******	9.0	Electrical connection	******
LB2V2	******	******	15.5	Installation error	******
LB3V2	******	******	17.0	Installation error	******
LB3V3	******	******	12.6	Missing part	******
LB3V4	******	******	19.0	Tightening	******
LB3V5	******	******	18.7	Tightening	******
LB4V1	******	******	14.5	Tightening	******
LB4V2	******	******	8.1	Tightening	******

Figure 8: Average number of issues found per engine type.

4.1.2 What are the most frequent issue categories?

The issue categories were extracted from the Data 1 and sorted by frequency to improve readability. This was accomplished by using a combination of Excel functions, though the same result could have been achieved with a pivot table. The resulting table, shown in Figure 9, excludes entries with no assigned category, which accounted for approximately 11% of the total dataset. The formula used can be found in Appendix 1.

Reported issue types and their amounts (LB)	
Issue	Amount (%)
Valve	0.3 %
Instructions	0.4 %
Cylinder head	0.8 %
Software	1.0 %
Wrong part	1.0 %
Rust / debris	1.3 %
Design	3.0 %
Test run error	3.5 %
Electrical connection	9.2 %
Supplier quality	12.4 %
Missing part	12.9 %
Tightening, pipes	13.3 %
Installation error	18.3 %
Tightening	22.6 %

Figure 9: Reported issue categories and their relative amounts for LB engines.

In Figure 9, the categories are sorted from least common (top) to most common (bottom). The most common issue types are dominated by errors involving assembly work like tightening errors and installation errors. The table was then turned into a Pareto graph which is displayed in Figure 10.

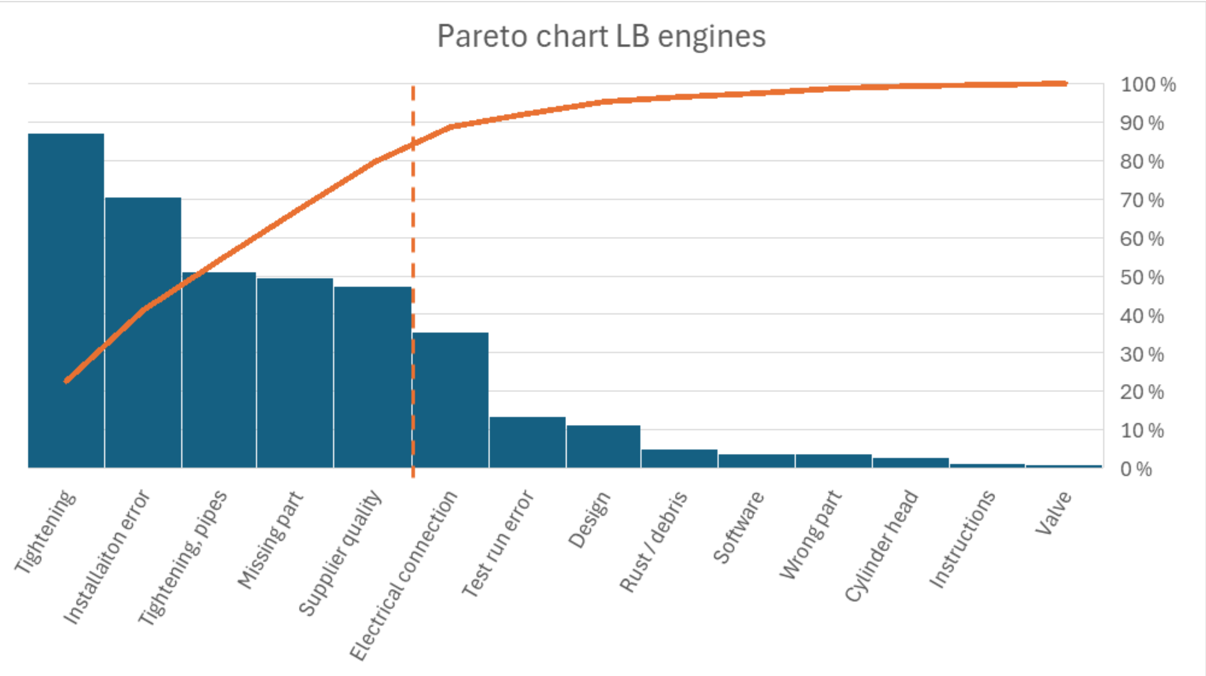

Figure 10: Pareto chart of the relative amounts of issues regarding LB engines.

The Figure 10 shows the relative amounts of the issues in the data. The distinct characteristic in the Pareto chart is the line graph showing the cumulative percentage of issues across the different categories. A dashed line was added between the two categories where the cumulative percentage crosses the 80% threshold, indicating that the five categories to the left of the line account for 80% (79.5% to be exact) of the total number of categorized issues. The Pareto analysis revealed that the largest issue categories in 2025 were: “Tightening”, “Installation error”, “Tightening, pipes”, “Missing part”, and “Supplier quality”.

When analysing these results, it was important to consider the quality of the categorization, as it had a great impact on the outcome of the analysis. Like said in Chapter 3.2.1, the categorization was made by another employee in the case company and not by the author. Therefore, the categorization was not necessarily optimal for this use case. While analysing the data, the author considered redoing the categorization but decided not to for a couple of reasons. First, the categorisation already looked quite good and logical, and the only thing concerning the author was that some issue lines were missing the categorisation entirely. Second, redoing the categorisation would have required a significant amount of manual labour, given that the database contained nearly 2000 entries. Due to the combination of these two factors, the author decided not to risk spending a considerable amount of time on work that might not yield any significant improvements in the results. Additionally, the missing lines seemed to appear at random and because the amount of them was only 11% of the LB entries, a decision was made to just exclude them from the analysis.

4.2 Discoveries from interviews and participant observation

This chapter presents the findings gathered from the interviews conducted during the hands-on training sessions as well as from the other interviews. The aim was to identify the causes behind the most common issues observed in the data analysis. However, during this deeper investigation, it became apparent that the issues identified in the data are not the only sources of delays. For this reason, this chapter brings up all the lead-time-related observations by beginning with broader, upper-level issues and then gradually narrowing in on more specific problems and causes for the issues observed during the data analysis phase. Additionally, the chapter will provide further rationale for why the study focuses on the LB3 engine.

4.2.1 Operations control

During the thematic analysis, the findings were divided into eight themes of which five will be addressed here. The first theme to be addressed is “Operations control”, which includes employees’ comments on how day-to-day operations are controlled and monitored at the factory. There were three main topics that were brought up: communication between teams and organizations, Teams as a communication platform, and the usage of the case company’s proprietary manufacturing control software (PMCS). Already during the preliminary meetings at the beginning of research, it was brought to author’s attention that communication inside the factory might be lacking. Case company’s Chief Technical Expert (CTE) said that there is a lack of open communication between different departments hindering knowledge sharing across the organization [45]. This was further observed by the author during the visit to the factory, where it was noted that even the different teams handling the test cells were not always up to date on what was happening in each cell. The CTE also mentioned that, in their opinion, the quality organisation is excellent in principle but does not function effectively in practice [45]. This leads to a situation where the problems they identify are not documented in reports and consequently do not lead to improvements [45].

During the interviews, the author’s decision to use the Teams channels as the source of issue data was further reinforced. This was confirmed by the General Manager (GM) overseeing the engine testing and supported by another GM, who stated that even though more official sources of data exist, the data would most likely be the same across the other systems, which is why it would not make sense to spend time on those [46], [47]. However, an important topic that arose was whether Teams channels are actually an effective way to communicate during testing. Advantages were mentioned to be: search function, real time communication, quick general view of the T&V, and the ability to quickly check who is currently working [47]. However, the downside of Teams is that it isn’t a structured system, so when an issue occurs, an official fault notification isn’t generated automatically. This makes analysing operations at T&V more difficult, as the data has to be logged manually from Teams. [45], [48].

One suggestion to replace the Teams channels, or at least to work besides them, has been to use the company’s proprietary manufacturing control software. The PMCS is the case company’s own system that includes all the logistics and manufacturing steps needed to build the engines. At the time of the study, this software was used during every step of the manufacturing process until the testing phase [45]. If this were to be adopted for testing as well, all engine issues could be logged in the same system, making it easier to identify their root causes. However, the problem with using this software in T&V is that it does not work on mobile phones the way Teams does, and the data is harder to find compared to Teams [47].

The interviews revealed that the PMCS also had another issue affecting operations, particularly regarding assembly errors. The assembly order of parts was not always correct in the system. The author was informed of situations where parts were installed onto the engine only to be removed later in order to install another part first. This was told to be caused by an incorrect assembly order in the PMCS [49]. The most notable example of this was provided by an assembly team Coordinator, who told the author that even though the engine assembly was currently at Phase 0 (out of a total of 12 phases), they had already installed a pipe that was not mentioned in the PMCS until Phase 5 [49]. These situations expose connections to tightening errors, increasing the possibility of leaks later during testing.

During the hands-on training, the author also observed that there were some clear gaps in the workers' knowledge, which could indicate poor instructions or a lack of training. One example was seen when the generator was being lined up with the engine, and the employees were unsure how to correctly use the instrument for measuring the orientation of the generator. It seemed like they had not been given training on how to use that instrument. Additionally, the workers did not know where this instrument was located in the factory, so some time was spent searching for it. Another area lacking a proper procedure was when testing personnel encountered a fault in the engine that they could not diagnose themselves. The CTE stated that there should be a clear process for determining who should be informed in these cases, rather than simply calling some colleague who might know something about the issue [45].

4.2.2 Existing processes effect on factory operations

The next category under investigation is the effect of internal processes on the factory. During the study, two suggestions regarding the processes emerged that would reduce the workload at the cells and consequently shorten the lead time. A suggestion was made that the engines' oil system could be flushed already in the assembly, rather than in the test cell [47]. This would completely remove one step from the testing process reducing the time it takes to get the engine ready for running, while still assuring that the engine is clean before starting it for the first time. During this flushing operation, "oil umbrellas" are installed on top of the oil jets inside the cylinders of the engine. The function of these is to prevent the potentially dirty oil from reaching the piston pin and piston rings. The installation of these umbrellas also takes time, and removing them is especially unpleasant for the workers, as it involves going inside the engine, which is by then heavily dripping with oil. An employee working at the cells stated that the entire umbrella operation is useless and a waste of time [50]. However, suggesting the removal of this step would require significantly more research to determine whether the umbrellas are actually needed [46].

Another suggestion concerned the drying of the engines after testing. Currently, the coolant used in the engines is water-based, meaning that if not dried correctly, the engines would rust. This is why a significant amount of time is spent emptying and drying the engines to mitigate this effect. If the coolant were to be changed to a glycol-based one, similar to those used in the automotive industry, this problem would not exist, and the engines could be removed from the cells more quickly. One reason for using water-based coolants is that it conducts heat better than a glycol-based alternative [45]. However, the CTE said that this difference is negligible and it is not relevant anymore. [45]

While following the assembly workers, an important topic arose regarding the organization's reaction to quality deviations. Currently, the official policy is that if a part is broken, there is a mistake in the drawings, or there is some other clear mistake affecting the assembly team, a deviation report should be created in the system so that the issue can be addressed. However, the experience of the assembly workers was that the deviation process is too slow, potentially taking multiple years before corrections are made, if they are made at all [49], [51], [52]. For example, if there is one mistake in an engineering drawing, it usually isn't even fixed [53]. Only when there are multiple mistakes in one drawing, it will be addressed and even then it does not happen quickly [53]. Additionally, creating the deviation tickets takes a significant amount of time for the workers. The organization's lack of response to deviations, combined with the time required, has led to a situation where employees feel it is not worthwhile to spend their time creating these reports anymore when they have no effect regardless [54]. The author also heard that if a part has, for example, a minor casting or machining defect that can be resolved by sanding, the deviation report is simply skipped and the workers just fix the part themselves and proceed with the assembly [55].

Another popular topic that arose from multiple people, was the fact that who fixes the engines when they need repairing in the test cell. Currently, the process is that the same assembly team that built the engine comes to do the necessary major repairs into the test cell. Loose connections are usually simply tightened by the testing personnel. The idea behind this practise is that it creates accountability for the assembly teams while also serving as a learning opportunity for them. However, from the side of the lead time, this isn’t efficient. It was mentioned by multiple people that it would be faster if the T&V just fixed the issues themselves [45], [47]. First, the assembly workers don’t like to come to the test cells since the environment there is hot and dirty [45]. The assembly workers don’t also have permission to go to the test cells, which means that when they need to go there, they need to go and get a special approval for that, which based on the authors understanding isn’t very practical.

The author also noted during the hands-on training at T&V that when the assembly workers came to the cell, they were unsure of what they had come to fix and why. While the author was observing the repair crew, the workers even asked the author about what the issue with the engine was and which part they were supposed to replace. Eventually they opened the box containing the new part that had already been delivered to the cell, and based on that, reasoned what needed to be replaced on the engine.

A further consequence of this practise is that when the workers come to the test cells, they need to stop their own work on the assembly side, which in turn may delay the assembly of a new engine. However, it is difficult to evaluate the pros and cons of this way of working with the available data, so it is hard to make a case for any changes.

4.2.3 Facility’s effect on factory operations

Like seen already, there are many aspects other than the issues affecting the lead time, one of which is the facility’s restrictions. On the assembly side, it was seen that sometimes the workers couldn’t do anything because they didn’t have parts to install to the engine. This wastes time and might lead to rushing when the parts are eventually delivered. Another observation related to in-plant logistics was that workers might end up installing parts in the wrong order because a part was lying somewhere on the factory floor, getting in the way of others. As a result, they go and retrieve the part and install it on the engine without considering whether some other part should have been installed before it [55].

There were several observations on the T&V side regarding the factory’s constraints. Delays can occur already when the engine is lifted into the test cell, if the crane used to lift the engine is occupied by another cell or team. There are two main cranes in the factory, each serving four of the eight cells. Additionally, if the engine needs to be transferred from one crane to another, it takes time because a separate moving platform is required to move the engine sideways [47]. The crane controller also had an unusual issue where the most precise speed setting did not work when the crane was operated inside cell 16 [50]. This made it more difficult and time-consuming to lower the engine onto the engine stands, as the adjustments made with the faster speed settings were too imprecise.

The author also observed that the tools within the cells were disorganized and that some tools were even missing when the assembly workers attempted to carry out the repairs. Additionally, when connecting the piping to the engine, the author noticed that the sizes of the gaskets were not documented anywhere. The gaskets had to be collected from a common warehouse. The author witnessed a couple of instances where a gasket was fetched from the warehouse just to see that it wasn't the right size, again wasting time. Since the piping in the cells does not change, the author suggests that it would be beneficial to have the sizes clearly displayed in the cell and to have, for example, a shelf where all the necessary gaskets and bolts would be readily available inside the cells.

The last major restriction from the factory's side was the limited loading capacity. When the engines are tested, they need to be loaded to simulate the real running environment. The engines can be loaded using either resistors or the city grid as a load. However, this capacity is limited. Currently, the facility's loading capacity cannot support all cells running simultaneously. As a result, engines need to wait for other engines to complete their tests before they can be tested. During the visit, a situation was witnessed where a FAT in cell 16 was aborted because cell 17 required the loading capacity. This is not an easy problem to solve, and additional loading capacity is already under construction. In the meantime, one solution could be to schedule the tests more precisely so that the engines would run at different times according to the plan.

4.2.4 Personnel's effect on factory operations

This chapter will cover the observations made regarding personnel and their effect on operations and lead time. Regarding the option of increasing working hours, the testing GM mentioned that if the testing capacity were to be increased by recruiting more employees and adding more shifts, it would pose a challenge, as a large number of new workers would need to be trained simultaneously and the amount of available workers is limited [47]. Another observation regarding staffing was made while lifting the engine into the test cell, when the crane operator informed the author that there was no employee on the evening shift who had the required certification to operate the crane [50]. This amplifies the previously mentioned crane capacity problem (Chapter 4.2.3), as there is a possibility that a delay of a few hours caused by the crane being occupied could turn into a full-day long delay if there are no qualified workers at the factory when the crane becomes available. This was seen first-hand by the author during the training and in that case, the engine had to be rushed into the cell to avoid this issue. Doing so meant that one step had to be skipped: installing an adapter plate on the flywheel, which should have been done before moving the engine to the test cell.

While showing the facilities to the author, the GM told that in the cell 17, the quality of assembly work was really good whilst in cell 16, almost 100% of issues were caused by errors in the assembly [47]. It is important to note here that even though cells 16 and 17 are both LB cells, cell 17 mainly tests LB4 engines while cell 16 mostly tests LB3 engines. Therefore, the most likely reason for this is a difference between the engine types rather than the personnel building them.

Further insight into the causes of installation issues was provided by the CTE, who identified a lack of motivation among assembly workers as the primary contributor to errors [45]. The CTE remarked: *"Everyone has tyres in their cars and shoes on their feet, so it shouldn't be a matter of skill"* [45] (author's translation), implying that the required tasks are within the competence of all workers. This assertion provided a useful framework for the author's subsequent evaluation during hands-on training sessions, where the relationship between worker motivation and error frequency could be directly observed.

This question ended up being more nuanced than anticipated. While observing the work in the assembly, the author noted that the workers did their best to build the engine as well as they could, and there were no clear signs of a lack of care. Actually, the contrary was observed. The team did pay attention to every part they attached to the engine. If they were installing a covering plate, the seal and the sealing surfaces were visually inspected for any defects and by feeling the surface by hand. Also, O-rings were checked for defects. The author was told that every O-ring should be checked, even if it is a new one [53]. Like the seals, the parts themselves were also inspected for any defects or machining residue. For example, the oil jets located inside the cylinders were checked with a flashlight and pressurized air to blow out any residues. The author heard that previously, there had even been cutting shavings in the oil passages inside the jets [53]. The interesting part regarding this was that this activity was not, at least not clearly, instructed by the company, but the workers still carried it out based on tacit knowledge that had been accumulated within the community over years, if not decades. This shows that building these engines, with the current level of instructions, requires a lot of experience. This was further verified by an assembler who had been working at the company for a long time, who stated that *“you cannot just take anyone from the street to do this job”* [52] (author’s translation).

However, the work of the assembly team was not perfect either. There were quite a few instances where the workers took shortcuts from the official instructions. The first thing the author noticed was that the torque specifications for different bolts were not usually checked from the PMCS but were instead remembered by heart. It was told to the author that standard bolt connections did not have any special torque specification, but they were tightened based on a standardized torque table that shows the required torque according to the size and strength class of the bolt in question. It would make sense that these values are quickly memorized for the most common bolt sizes. However, there are special torques for certain connections, and those are checked from the PMCS.

After seeing the loose pipe connections in the data, the author wanted to observe how the connections were tightened. The author enquired about this matter and was informed that there is an official procedure for tightening them. The official procedure for Ermeto PSR connectors, according to Parker, the manufacturer, is to first tighten the nut finger-tight and then tighten an additional 540 degrees [56]. However, this was not usually followed by the assemblers, who instead tightened the connections based on feel. *“You will feel it when the connection is tight enough”* [49] (author’s translation). It was also noted that some of the connection locations are so difficult to reach that following the official procedure would be nearly impossible regardless [49].

One careless error was also spotted during the hand on training. There were some plugs missing from the front main seal that were supposed to have been installed when the seal was fitted to the engine block. This was noticed by one of the more experienced employees, who recalled that these particular plugs had been forgotten on a previous engine as well. The employee stated that the plugs are marked in the drawing, but they are so small in comparison to the seal that they can be easily missed [53]. One other aspect that might lead to these small mistakes is that the tasks are not assigned by the team leader; instead, every assembler needs to independently find tasks to work on. The goal, however, is to divide the work so that everyone on the team knows how to perform every step in the engine building process [49].

How about the employees at T&V? The author noted that there was a considerable amount of waiting at T&V. This is likely a combination of multiple factors, but it provides further indication that increasing staffing and working hours may not be the most effective solution to the long lead times. As mentioned in Chapter 3.3.2, the author had the opportunity to observe the installation of a new LB3 engine in the test cell and, during that process, was able to make notes about how the team worked.

The first thing worth mentioning was observed when lifting the engine from its waiting spot on the factory floor. During the lifting, the engine needs to be balanced so that it is perfectly level while being moved. In practice, this is done by using a lifting tool attached to the crane that can adjust the lifting position, thereby affecting the orientation of the engine. There were clear markings on the side of the tool indicating where it should be set so that this engine model would be level. However, there is apparently some variation between engines even though they are the same model, so the tool was not set to the mark but close to it [50]. The engine was then lifted slightly to check the orientation and then lowered to make an adjustment. On the first attempt, the correction was made in the wrong direction, which meant that the engine was even more tilted than initially. This was surprising given that this was not the first time the employees had lifted an engine and the operation was not particularly complicated. After some trial and error, the engine was eventually level and the lifting could commence.

Next observations were made when the generator was lined up with the engine. Like mentioned in Chapter 4.2.1, the orientation of the generator was measured using an instrument that the workers weren’t quite sure how to use. Based on those measurements, the huge generator should then be moved to correct the orientation. This was done by using hydraulic jacks positioned between blocks fastened to the cell floor and the generator. The difficulty in this process was that the adjustments needed were measured in tenths of millimetres, which is so fine that when the static friction eventually gave way under the pressure, the entire generator moved too far in the other direction. This back and forth continued for quite some time until it was finally agreed that the alignment was good enough. The author tried to ask what the acceptable margin of error was, but the workers did not seem to know. The author also noted that this process occupied two or three people, even though the job really only required one person. Meanwhile, the engine waited for the pipes and electrical connections to be made, which could have been done simultaneously by the other employees at the cell.

4.2.5 Observations made about engines' design

While speaking with different personnel, the author heard the LB3 engine being compared to the LB4 on multiple occasions. For instance, the GM of testing stated that the LB4 is their best product, while the LB3 is their worst [47]. The same comparison was made by the assemblers in the test cells who described the LB4 as reliable, simple and better in any way [50]. This is understandable, as the LB4 platform is considerably older and has had significantly more time to be perfected. Additionally, it does not utilize as advanced technology as the newer LB3.

One of the key differences of the LB3 compared to previous engines is that it has a lot of piping. This is due to two reasons. First, there are not as many internal fluid passages in the engine block, so fluids need to be transported externally using pipes. Second, the LB3 has more features and systems, each requiring its own piping. This results in a design with a large number of pipes and connections. Another note regarding the engine design was that some installation clearances are so small that it is difficult for the assembly workers to install the parts correctly without damaging or misplacing components such as O-rings [53].

Another difference between the LB4 and LB3 was noted when the engine was lifted into the test cell. The worker mentioned to the author that the LB4 is much easier to lower into the cell, as that model sits on top of spring racks, while the LB3 uses engine mounts. These mounts have a stud bolt that must be aligned precisely with the holes in the engine block to secure the engine in place. With the LB4, it is sufficient to simply land on the spring racks without the need for ultra-fine adjustment. [50]

While showing the test cells to the author, the GM also mentioned a lack of markings on the engine components. The engines have multiple pipes connecting to them, and some of them look identical. The GM stated that the workers need to check the drawings to determine which pipes to connect where and noted that it probably takes more time to retrieve and interpret the drawings than to actually install the piping. It was suggested that if the pipes would be marked already in the design phase, this step could be skipped. Additionally, customers have requested this information to be marked onto the parts on multiple occasions, which means that T&V need to make these markings to the pipes themselves regardless. [57]

Lastly, while observing the work at T&V, the author noticed that there were some clearly recurring faults in the engine models. For example, certain plugs were immediately checked or replaced when the engine arrived at the cell without even verifying whether they were faulty.

4.3 Root cause analysis of identified issues

This chapter presents the results obtained from the Fishbone and Five Why's analyses. The results combine findings gathered throughout the hands-on training and interview phases of this study. Conclusions are drawn based on these findings to identify the key root causes and contributing factors causing the issues at FAT. Together, these analyses provide a structured foundation for understanding the problem areas and inform the recommendations presented in later part of the work.

4.3.1 General overview of causes using a Fishbone diagram

At the beginning of the study, it was stated that the delays in FAT were mainly caused by engine issues such as oil and coolant leaks, the fixing of which consumed time in the test cell. However, after a closer inspection into the whole operation the author concluded that the issues, while being a major contributor, were not the only cause for the delays. To illustrate this, a Fishbone diagram shown in the Figure 11 was created to displays different causes for the delays found during the work [17]. The figure includes also the top five issue categories identified previously with the Pareto analysis in Chapter 4.1.2.

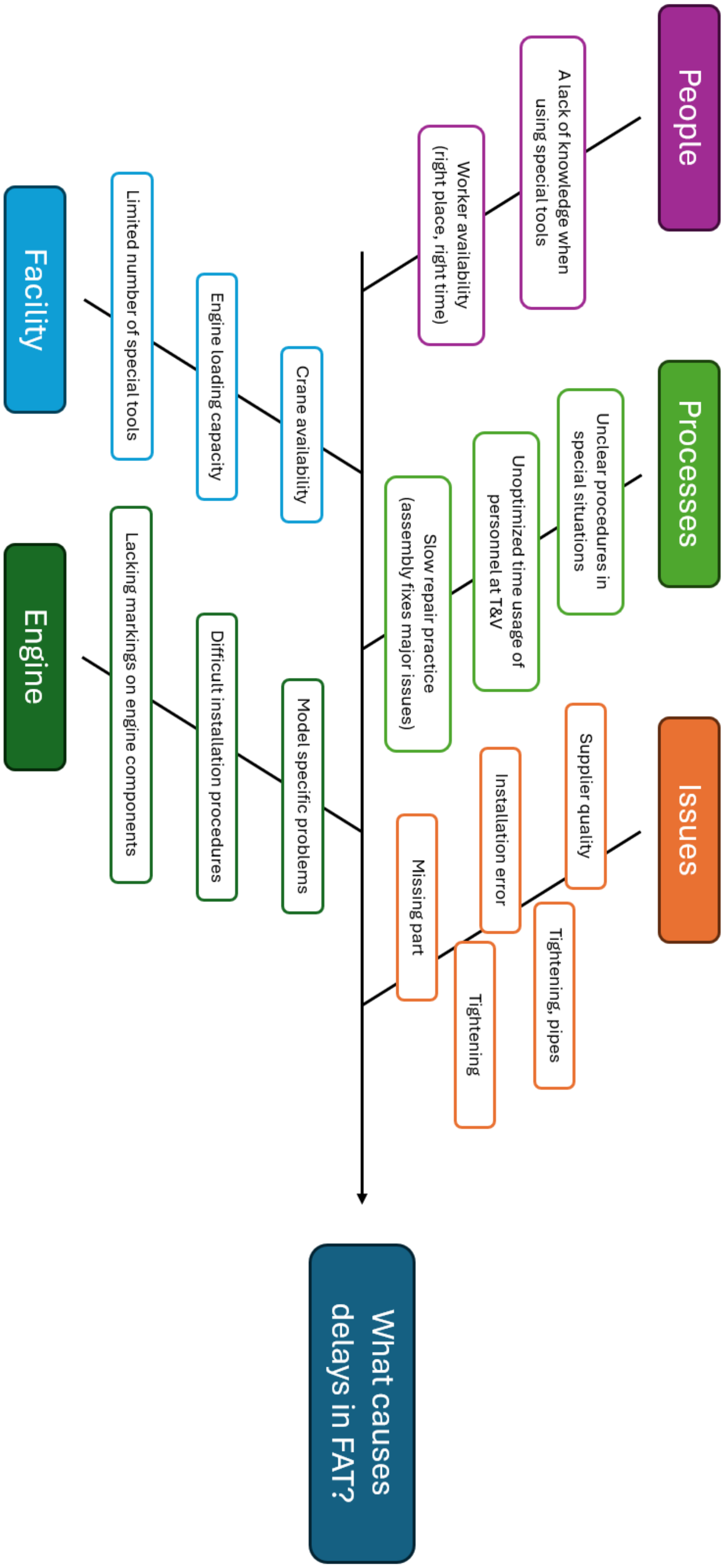

Figure 11: A Fishbone diagram displaying the causes for delays in FAT.

Although the focus of this study was on LB engines, Figure 11 includes all observations made across the factory to provide a general overview. Furthermore, the figure does not quantify the extent to which each cause affects the lead time; rather, it is intended to illustrate the complexity of the challenge of reducing lead time as a whole. It is important to note, that the causes listed in the Fishbone diagram are not necessarily the root causes, but the things that have an immediate impact on the lead time of the testing. The reasoning behind this was that when analysing any of these issues, the analysis quickly extends to different parts of the organisation, and some root causes may lead to multiple causes mentioned in the diagram. This would have made creating a clear and easily understandable diagram nearly impossible, so this approach was chosen instead.

4.3.2 Root cause analysis utilizing Five Why's method

Each of the causes in Figure 11 can be analysed as its own case, and a root cause analysis method can be applied to it. The method chosen for this was the Five Why's technique.

The root cause analysis was carried out only for entries in the "Issues" branch of Figure 11. At the beginning of this work, management identified engine defects as the most significant factor extending lead times – an assumption that guided the scope of the assignment. Since the study was unable to quantify how much the other causes (other branches in Figure 11) found during the interviews actually affect the lead time, and since the Pareto analysis already indicated the major role of engine defects in the delays, proceeding with management's initial assumption was deemed the most reasonable approach. Additionally, because the assignment was primarily scoped around the defects, enough information on the other topics was not gathered to perform a quality root cause analysis on them. The impact of these other findings on lead time is, however, something that could be studied in the future.

Another problem encountered during the analysis was that the Five Why's method usually requires specific cases where the causes can be investigated with high accuracy. However, this study talks more about general patterns than clear one-off cases, which means that the root causes are also general patterns and not absolute values. For example, a leaking pipe connection can be caused by multiple factors, ranging from human error to design error. This is why the analysis needs to be applied somewhat differently when analysing these issue categories. In practice, this is done by first specifying the issues beyond the original categorisation and then allowing for the possibility of splitting the causal trees to include multiple causes for each statement. The generalisation of the data also means that completely different issues can fall under the same category. For example, a leaking pipe connection can either be completely loose with a tightening marking or tightened and be still leaking. In these two cases the root causes would be completely different. For this analysis, the author has chosen these three specific cases based on what the author has learned during the research regarding their frequency, severity, and relevance to this study.

The specific cases chosen for Five Why's analysis:

- A marked pipe connection or bolt is only finger tight. Figure 12.
- An O-ring is damaged or completely missing. Figure 13.
- A plug is missing from some part or sub-assembly. Figure 14.

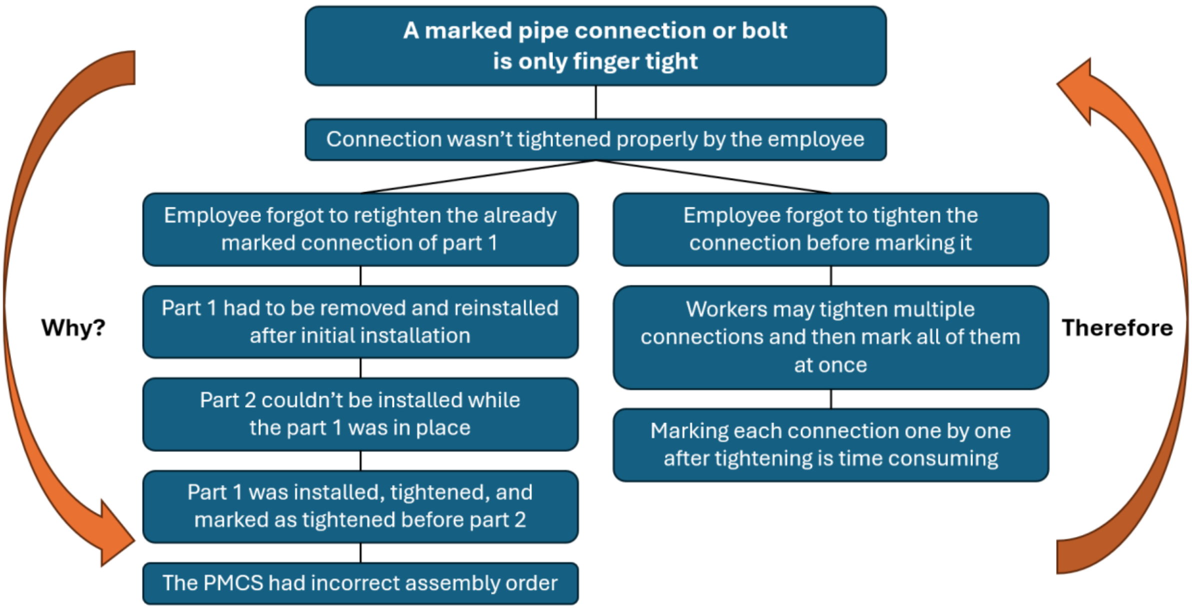

Figure 12: Five Why's analysis 1: A marked pipe connection or bolt is only finger tight.

A common pattern in all three cases is that ultimately, they are all errors made by the assembly workers. However, it is not beneficial for management to simply state that the problem lies with the employees, as rectifying that would be difficult. This has also been mentioned in the literature while defining a root cause: "Root cause: the most basic cause that can be reasonably identified and that management has control to fix" [58]. The more important thing is to find the patterns or processes that allow these mistakes to happen.

Figure 12 shows the analysis made about loose connections. This refers to situations where the fasteners are clearly loose but still have the marker pen marking on them, indicating that they have been tightened correctly. This issue is surprisingly common, as evidenced by both the data and the Teams conversations. In the Data 1, these appear under the “Tightening” and “Tightening, pipes” categories, which, according to the Pareto analysis, represent the first and third most significant categories, respectively.

Figure 12 was created by first stating the problem at the top of the figure and then asking “**Why?**” until a root cause was found. Each block in the graph is a cause statement, and beneath it is the cause for that statement. So, for example: **Why** “a marked pipe connection or bolt is only finger tight”**?** Because “connection wasn’t tightened properly by the employee”. When there was a situation where there was more than one substantial reason for a cause statement, the tree was split into the appropriate number of branches to account for all the causes. The decision of which reasons were substantial and which were not, was made based on the author’s knowledge on the matter gathered during the study. A way of verifying the analysis is to go from bottom to top while stating “**Therefore**” between the statements, instead of going from top to bottom asking “Why?”. This way, a clear causal linkage can be observed.

The first analysis showed that while the employees did forget to tighten the connections, the risk of this error was much higher when parts had to be removed and reinstalled after the fasteners had already been marked during the initial tightening process. This back and forth essentially nullifies the function of the markings on the fasteners if the workers cannot trust that a bolt with a marker pen line has already been tightened. This shows that more focus should be directed towards the PMCS system, which guides the engine building process. The building instructions should be designed so that when a part is attached to the engine, it is not necessary to remove that part again. This would greatly reduce the likelihood of these errors.

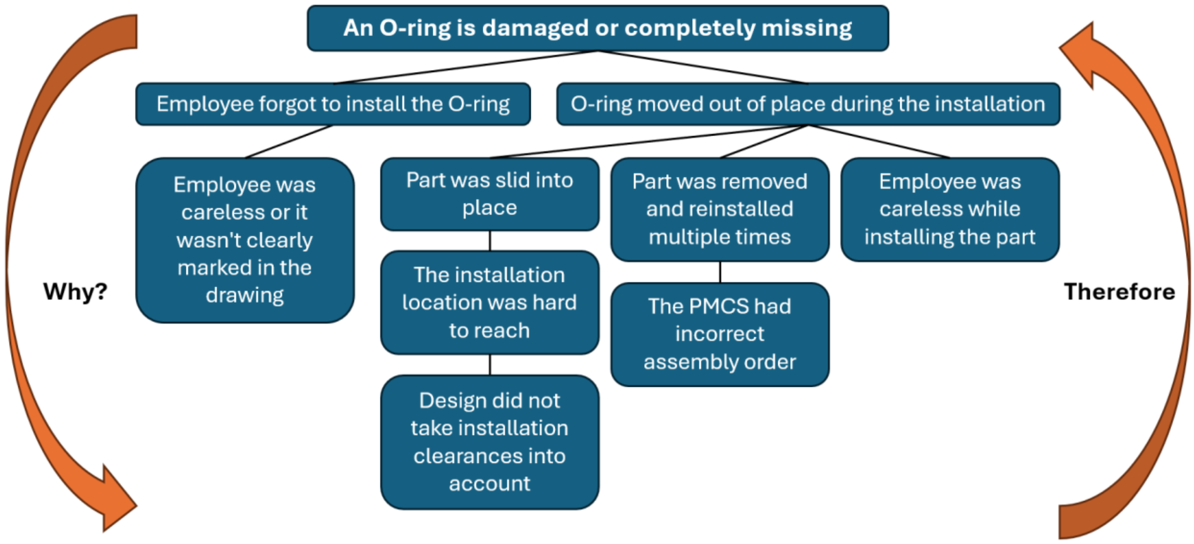

Figure 13: Five Why's analysis 2: An O-ring is damaged or completely missing.

The second analysis (Figure 13) covered the "Installation error" and "Missing part" categories, and more specifically the errors involving O-rings. Here, the same pattern was observed as in the previous analysis. The analysis showed that O-ring damage typically happens during the installation of the parts. While human error is a contributing factor, the analysis also identified several other underlying issues. Again, the PMCS came up. As with the connections, the removal and reinstallation of parts due to an incorrect assembly order can increase the risk of O-ring damage by rubbing the O-ring against the surfaces or by providing more opportunities for the O-ring to shift during installation and become lodged in the wrong position, destroying the ring when the bolts are tightened. Another root cause that could affect the O-rings, was the design of the engine. If there was not enough installation clearance, the parts were hard to install, which increased the chance for the O-ring to be misplaced. A worst-case scenario was described to the author, where there were so little room, that the part had to be slid into place so that the O-ring was rubbing against the other part during the installation. An important note here is that the O-rings should be intact before installation, as the employees check the rings prior to installation even if they are new.

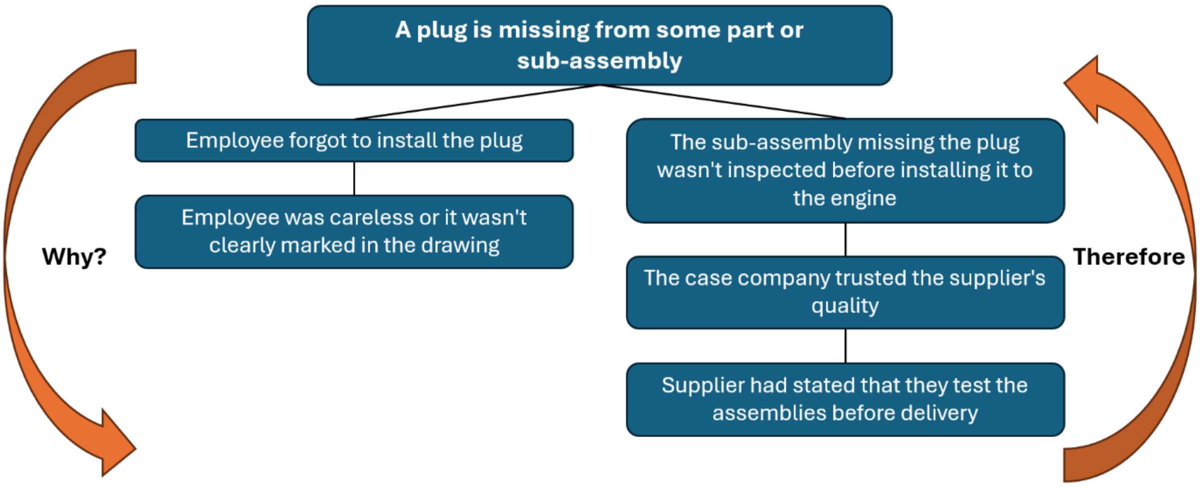

Figure 14: Five Why's analysis 3: A plug is missing from some part or sub-assembly.

The third analysis (Figure 14) covered the "Missing part" and "Supplier quality" categories and the specific case under investigation was missing plugs. The engine's oil and coolant passages have different size plugs all over them for multiple reasons. Some are for bleeding or draining the systems and some are just plugging holes that had to be drilled in the machining phase. It is essential that these plugs are properly installed before filling the engine's various systems, as failing to do so results in significant leaks that take a considerable amount of time to clean up.

Like in the previous analyses, the employees were at the centre of the issue. The installation of the plugs could be forgotten, but again there were other factors that affected how likely it was. It was mentioned that the plugs could be marked more clearly into the drawings so that they wouldn't be so easy to miss. Additionally, clear instructions on what plugs should be installed in which phase could make this process clearer and more structured. While visiting the assembly, the author observed a situation where a worker noticed that some plugs were missing and said something along the lines of: *"We should install those plugs next so that they aren't forgotten"*.

Another aspect of this issue was the sub-assemblies that came directly from suppliers, which were expected to work but were actually missing plugs. The suppliers have stated that the assemblies are tested before delivery, so they are not double-checked at the case company before being attached to the engine. One way to mitigate this would be to test every sub-assembly that is delivered to the factory, but this might prove logistically challenging. Another option would be to create a sanction system so that the suppliers would get a financial incentive to deliver only fully working assemblies.

4.3.3 Questionnaire results

The questionnaire received 39 responses out of a possible pool of slightly over 100 employees. Of the respondents, 30 were assemblers and 9 were assembly team coordinators. Regarding engine type, 2 respondents worked with small bore engines, 16 with medium bore engines, and 21 with large bore engines. This distribution provides a good representation of the large bore personnel which were in the focus of this study.

The responses to the question where the respondents were asked to select up to two of the most important factors causing leaking connections are shown in Figure 15.

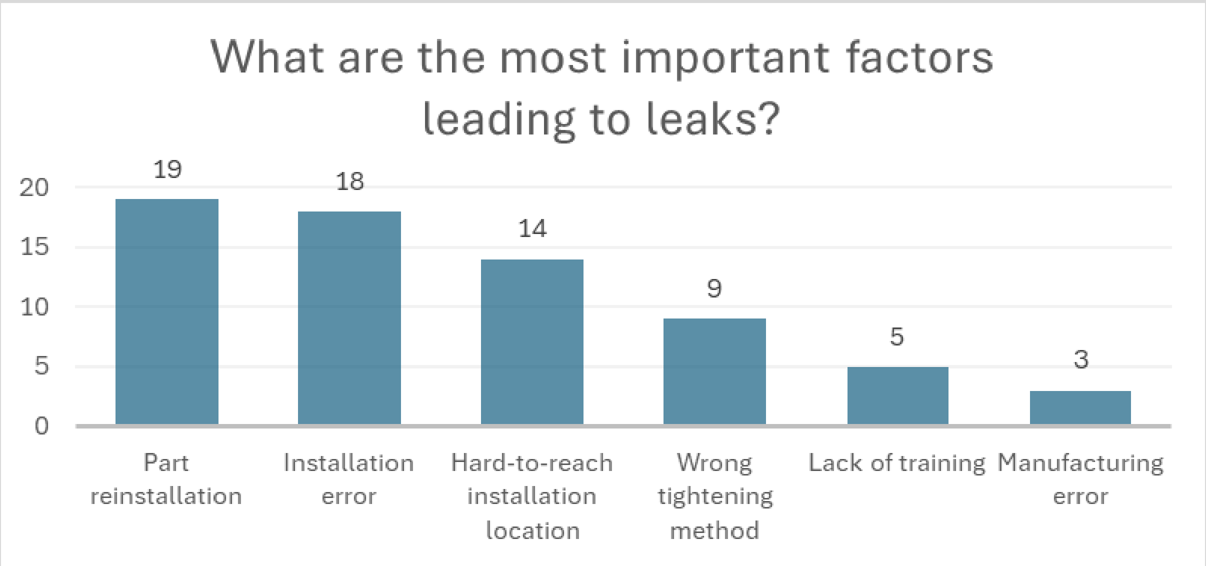

Figure 15: Answers to the question regarding the most important factors leading to leaking connections.

These results validate the earlier findings from the interviews and root cause analysis that part reinstallation is considered a major contributor to faults according to the workers themselves. Every time a connection is opened and reassembled, the risk of a leak increases, which makes it logical that the assemblers rank this as the most significant factor.

The high ranking of hard-to-reach installation locations is also noteworthy. This factor can stem from two different root causes: the engine design itself may place connections in locations that are physically difficult to access, or an incorrect assembly order may force workers to install fittings after surrounding components are already in place, making the work unnecessarily difficult. In both cases, the tight working conditions make it harder to achieve a proper connection, increasing the likelihood of a leak. This also connects to the tightening method issue, as assemblers working in confined spaces are less likely to be able to follow the official tightening procedure.

The open-ended follow-up question provided further insight into these issues. The respondents brought up employees rushing through the work, which naturally increases the chance of installation errors. An increased number of connections in the engines was also mentioned, meaning there are simply more potential points of failure. Additionally, respondents highlighted the inability to complete assemblies fully due to missing parts or an incorrect order of assembly. This last point directly supports the finding that reinstallations are a significant factor, as incomplete assemblies inevitably lead to connections being opened and redone later.

It is also worth noting that manufacturing errors received the fewest votes, indicating that the assemblers do not consider the quality of the parts themselves to be a major issue. Instead, the problems appear to be rooted in the assembly process, instructions, and working conditions.

The respondents were also asked whether they have had to remove already installed pipes due to lacking instructions in PMCS. Out of 39 respondents, 27 answered yes and 12 answered no (Figure 16). This confirms that insufficient PMCS instructions are a real problem with a direct impact on the leaking issues. A follow-up question asked how often this occurs: 14 respondents said less than once a week, 12 said 1–2 times a week, and 2 said 3–5 times a week. These results indicate that unnecessary reinstallations caused by poor instructions are not a rare occurrence but rather a weekly problem for most assemblers.

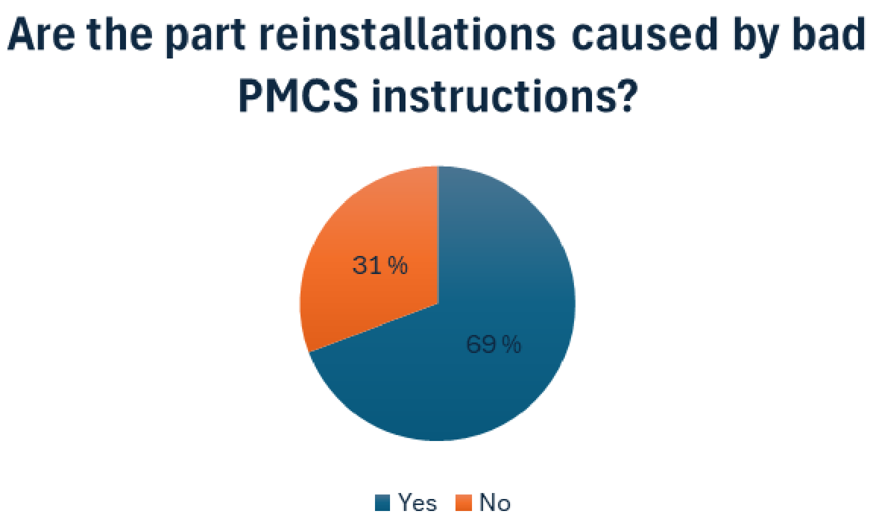

Figure 16: Answers to the question regarding the PMCS instructions.

The final main question concerned how the personnel tighten the Ermeto connections. Of the respondents, 10 said they use the official finger-tight plus 540 degrees method, while 29 said they rely on hand feel alone (Figure 17). When asked why they do not use the official method, the respondents mentioned the following reasons.

Overview of reasons given by the respondents:

- It is more efficient to tighten by hand feel.
- Years of experience and common sense make the official method feel unnecessary.
- Lack of knowledge about what the correct tightening method actually is.
- Many noted that the connections physically cannot be turned 540 degrees, especially in tight locations. In these cases, they simply turn the fitting as much as their strength and tools allow.
- A belief that if done strictly according to the instructions, the connections leak with certainty.

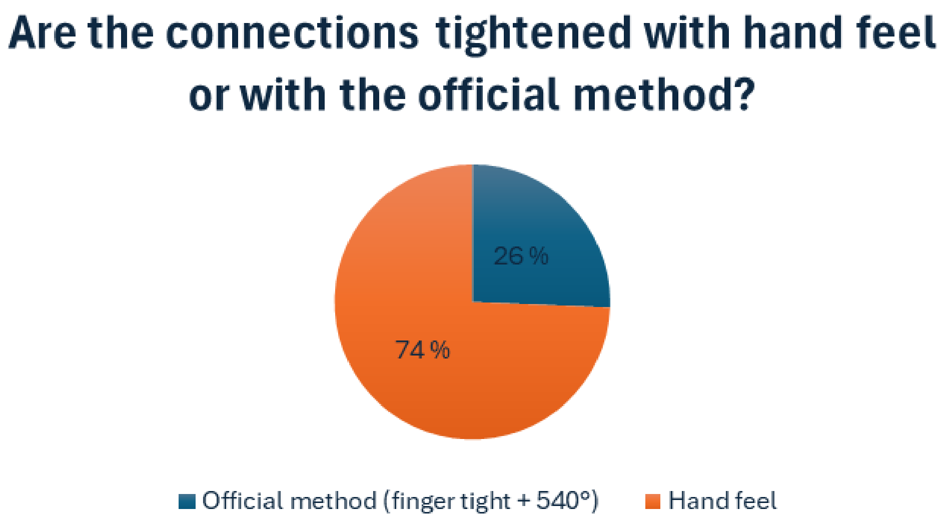

Figure 17: Answers to the question regarding the tightening method.

One response highlighted a major underlying problem. The respondent described a recent attempt to tighten a stainless-steel connection according to the official 540-degree method. Despite using full strength, they could only achieve approximately 500 degrees of rotation. After this, the connector's bead and threads were damaged, and the connection became extremely difficult to undo. This is a significant finding because the official instructions state that stainless steel connections must not be tightened in place [56]. Instead, a special hardened tool must be used to pre-assemble the bead, and lubricant must be applied to the threads to prevent stripping (if using a pre-assembled bead, the final installation procedure is wrench-tight plus 30 degrees) [56]. This single response effectively showcases two key issues: the lack of knowledge among assemblers regarding the correct use of these connectors, and a prevailing mentality of forcing connections rather than following proper procedures.

5 Conclusion and discussion

This thesis investigated the delays occurring in the factory acceptance testing process at the case company, with the goal of identifying the root causes behind the engine issues causing extended lead times. The work was motivated by the need to increase the throughput capacity of the existing test cells without major infrastructure investments. The study followed a mixed methods approach within a case study design, combining quantitative data analysis with qualitative interviews and participant observation during hands-on training sessions at the factory.

The quantitative phase of the study utilized two datasets: issue descriptions extracted from cell-specific Teams channels and engine test slot reservations from the company's production control software. The data analysis first established that the LB3 engine represented the most significant bottleneck in the testing process, based on its longer average test duration, higher number of issues per engine, and the limited number of available LB test cells. A Pareto analysis of the issue categories then revealed that five categories accounted for approximately 80% of all categorized issues: tightening errors, installation errors, pipe tightening errors, missing parts, and supplier quality problems.

The qualitative phase consisted of hands-on training sessions at both the engine assembly and Testing and Validation, interviews with employees from different parts of the organization, and a questionnaire distributed to the assembly personnel. These methods provided a deeper understanding of the mechanisms behind the issues identified in the data and revealed that the delays in FAT are not only caused by a single factor but rather by a combination of interrelated causes spanning multiple areas of the operation. To illustrate this complexity, a Fishbone diagram was created, organizing the identified causes into five categories: people, processes, issues, facility, and engine.

To answer the first research question – what the reasons for the delays in FAT are – the study focused on the engine issues and found, that the causes extend beyond the issues themselves. While faults such as leaking connections, damaged O-rings, and missing plugs are the most visible contributors to extended lead times, the root cause analysis showed that these issues are enabled by underlying factors in the assembly process, the manufacturing control system, engine design, and operational practices.

To answer the second research question – does assembly work contribute to these delays – the answer is clearly yes. The data showed that the dominant issue categories are directly related to assembly errors. The Five Why's analyses traced these errors back to factors such as incorrect assembly order in the PMCS system leading to unnecessary part reinstallations, insufficient tightening procedures, tight installation clearances in the engine design, and unclear markings in drawings. The questionnaire validated these findings, with the majority of respondents confirming that part reinstallation and hard-to-reach installation locations are the most significant factors contributing to leaking connections. Furthermore, 27 out of 39 respondents reported having had to remove already installed parts due to lacking PMCS instructions, and 29 out of 39 confirmed that they tighten Ermeto connections based on hand feel rather than the official procedure. These results demonstrate that assembly work is not only a contributor to the delays but is in fact the single largest source of issues encountered during FAT.

However, it is important to emphasize that the assembly workers themselves are not the root cause of the problem. Even though many of the found defects could be avoided by doing the assembly more carefully, the study showed that the workers generally perform their tasks with care and attention, and many of the errors are enabled by systemic factors outside their direct control, such as an incorrect assembly order in the PMCS, insufficient training on specific procedures, tight design clearances, and a deviation reporting process that does not react to the workers feedback. Addressing the delays therefore requires changes at the process and system level rather than placing responsibility solely on the individuals performing the work.

5.1 Proposals for the case company

The third research question guiding this study was: how can we mitigate the issues causing delays in FAT? This section addresses the remaining deliverables set by the management: proposals for reducing lead times, proposals for reducing lead time variation, and proposals for avoiding the need to repeat FAT tests.

It is important to recognize that reducing the number of issues encountered during engine testing contributes to all three of these deliverables simultaneously. This is especially true for lead time variation and repeated FAT tests. The defects create an unknown variable that affects engines randomly, introducing variation into the testing lead times that is difficult to predict or plan for. In worst cases, the issues are severe enough to require re-running the engines entirely. This uncertainty also extends the lead times indirectly, because to account for the possibility of issues, the target times for test slot reservations need to be set pessimistically. Therefore, any reduction in the number of issues would tighten the distribution of lead times, reduce the need for re-runs, and allow the company to plan test slots more accurately and with shorter durations.

The actual lead time can, however, also be reduced through other means. Investigating the efficiency of personnel at Testing and Validation to reduce the amount of time engines spend sitting idle in the cells would be a worthwhile area for further study. This was outside the scope of this thesis, but the observations made during the hands-on training indicated that there is potential for improvement in this area.

5.1.1 Proposal 1: Correcting the PMCS assembly order

One of the most impactful issues identified during this study is the incorrect assembly order in the PMCS system. As shown in the root cause analysis, an incorrect assembly order leads to parts being installed and then removed to install other parts first. This reinstallation process directly increases the risk of leaking connections, damaged O-rings, and other assembly-related faults by nullifying tightening markings and exposing components to additional handling. The questionnaire confirmed that this is a widespread and frequent problem, with the majority of respondents reporting that they have had to remove already installed parts due to lacking PMCS instructions.

To combat this, a project should be started where the PMCS instructions are thoroughly examined while an engine is being built. Any deviations between the instructions and the actual building process should be noted and evaluated to determine whether the system should be updated. This would require dedicated personnel to follow the assembly process and systematically document the discrepancies. The goal would be to bring the PMCS instructions in line with the optimal assembly order so that once a part is installed and marked as tightened, it does not need to be removed again. This is not a quick fix, but a sustained effort that would yield significant benefits across multiple issue categories.

5.1.2 Proposal 2: Ensuring correct installation of Ermeto fittings

A more specific and immediately actionable proposal concerns the Ermeto pipe fittings. The data showed that tightening errors and pipe tightening errors are the first and third most common issue categories, and the questionnaire revealed that the vast majority of assemblers tighten these connections based on hand feel rather than the official procedure. Furthermore, the study uncovered that there is a potential lack of knowledge among assemblers regarding the correct handling of different fitting materials.

It should be ensured that every assembler knows the correct installation method for these fittings. Specifically, it should be communicated that stainless steel connections must always be pre-assembled using the dedicated pre-assembly tools provided by the manufacturer, and that the required lubrication must be applied to the threads to prevent stripping. It should also be made clear that after this pre-assembly, the correct way to make the final connection is to tighten wrench-tight and then only 30 degrees more, because the bead has already been deformed during the pre-assembly step. For connections made from other materials, it should similarly be emphasized that the official installation method of finger-tight plus 540 degrees must be followed.

The author argues that the number of leaking connections would reduce drastically if these two proposals were to be implemented. Correcting the PMCS assembly order would remove the systemic factor that increases the risk of errors, while ensuring correct Ermeto installation practices would address the most common type of issue at its source. Together, these changes would target the root causes identified in this study and deliver improvements in FAT lead time, lead time variation, and the avoidance of repeated tests.

5.2 Discussion

The findings of this study are both expected and surprising. On one hand, it is not surprising that assembly errors are the dominant source of issues in engine testing – the engines are complex products with thousands of parts and connections, and human error is inevitable. On the other hand, the nature of these errors was more nuanced than initially anticipated. The errors are largely caused by a lack of care or knowledge regarding specific procedures, such as the correct installation method for Ermeto fittings. However, this incompetence is not entirely the fault of the workers themselves. Since the issue seems to be more systematic than individual based, the cause can probably be traced back to insufficient instructions and training. Furthermore, even where competence gaps exist, their impact on the final product could be significantly mitigated with better systems guiding the assembly process. If the PMCS provided a correct assembly order at every step, the error-enabling process of removing and reinstalling parts could be avoided, lowering the overall probability of mistakes.

The study also revealed that the delays in FAT cannot be attributed to a single cause. While the engine issues are the most visible and measurable contributor, factors such as facility constraints, crane availability, loading capacity limitations, and inefficiencies in communication and operational practices all play a role. However, the engine defects stand out as the primary cause of delays, with limited loading capacity occasionally causing additional unpredictability. If these were to be resolved, the most significant sources of delays could be eliminated. The lead time could then be further reduced by addressing the remaining factors such as crane availability, communication gaps, and idle time in the cells. These secondary factors do not cause delays on their own to the same extent but removing them would allow the testing process to approach its theoretical minimum duration.

5.2.1 Contextualization within existing literature

The findings align well with established production management theory. The concept that bottlenecks determine the throughput of an entire system is central to operations management, and this study confirmed that the issues present in LB testing are indeed the constraining factor in the factory's overall testing capacity.

The root cause analysis methods used in this study – the Fishbone diagram and Five Why's technique – proved to be suitable for this type of investigation. However, the generalized nature of the issues required some adaptation of the Five Why's method, as the technique is traditionally designed for specific, individual cases rather than recurring patterns. Despite this adaptation, the method was effective in tracing the issues back to their underlying causes.

The use of Pareto analysis to prioritize the issue categories was straightforward and effective, confirming the well-known principle that a small number of categories accounts for the majority of problems. This provided a clear focus for the qualitative investigation and ensured that the most impactful issues were addressed first.

5.2.2 Limitations of the study

Several limitations should be acknowledged. First, the data used in this study was not originally collected with this research in mind. The Teams channel messages were logged by employees during their daily work without consistent categorization rules. While the data was sufficient to identify patterns and prioritize issues, it is not perfectly accurate. The approximately 11% of entries without a category could, in theory, change the ranking of the issue categories if they were disproportionately concentrated in one type.

Second, the categorization of the issues was performed by another employee and not by the author. While the categorization appeared logical and consistent, it was not possible to verify every entry against its original message. Some ambiguity in the category definitions could also have led to inconsistencies. For example, the boundary between “Tightening” and “Installation error” was not always clear, and some entries could arguably belong to either category.

Third, the qualitative findings are based on observations and interviews conducted during the two one-week training sessions and a limited number of separate meetings. While the findings were validated with the questionnaire, the observations are inherently limited to the teams, engines, and situations that happened to be present during those weeks. Different teams or different engine variants might have revealed additional or different issues. This is a necessary compromise when conducting participant observation, but it should be kept in mind when generalizing the results.

Fourth, the study focused mainly on the LB3 engine based on the data showing it as the biggest bottleneck. While this was a justified decision, it means that the specific findings and proposals may not apply directly to other engine types. The broader observations regarding PMCS and tightening practices are likely relevant across the factory, but this has not been verified for other engine models.

Fifth, the questionnaire, while providing useful validation, had a response rate of approximately 39%. This is reasonable for a workplace survey, but it means that the views of the majority of the assembly organization were not captured. Additionally, the questionnaire focused specifically on Ermeto fittings and PMCS, which means that other issue categories were not validated in the same way.

5.2.3 Reliability and validity

The mixed methods approach strengthened the overall reliability of the study. The quantitative data provided an objective basis for identifying and prioritizing the issues, while the qualitative methods provided depth and nuance to the analysis. The questionnaire served as a third source of information, validating the findings from the other two approaches. This triangulation means that the central conclusions of the study are supported by multiple independent sources of evidence.

However, the author's position as a participant observer introduces a degree of subjectivity into the qualitative findings. The author's interpretations of events and conversations were influenced by the knowledge gained from the data analysis, which was conducted prior to the training sessions. While this prior knowledge enabled more focused observation and better questions, it also means that the author was not entirely neutral when entering the factory. Efforts were made to mitigate this effect by asking open questions and allowing workers to express their own views before presenting the data findings.

5.2.4 Future research

Several areas for future research emerged from this study. First, the efficiency of personnel at T&V was identified as a potential area for lead time reduction but was not investigated in depth. A time study or workflow analysis of the testing process could reveal how much of the test slot duration is spent on value-adding activities versus waiting or idle time.

Second, the impact of correcting the PMCS assembly order should be measured after implementation to quantify the actual reduction in issues. This would also validate the causal link established in this study between reinstallations and assembly errors.

Third, the deviation reporting process was identified as a systemic issue that discourages employees from reporting problems. Investigating how this process could be made faster and more responsive could have broad benefits beyond the scope of this study.

Fourth, the study identified that some engine design choices contribute to the difficulty of assembly and consequently to the errors found during testing. A design-for-assembly review of the LB3 engine, focusing on installation clearances and the number of external pipe connections, could provide longer-term solutions to issues that cannot be eliminated through process improvements alone.

Lastly, a similar study could be conducted for the MB engines to determine whether the same patterns and root causes apply across the factory, or whether different engine classes have fundamentally different problem profiles.

References

[1] 'Esittelyssä GPT-5'. Accessed: Apr. 14, 2026. [Online]. Available: https://openai.com/fi-FI/index/introducing-gpt-5/

[2] 'Claude Opus 4.6'. Accessed: Apr. 14, 2026. [Online]. Available: https://www.anthropic.com/claude/opus

[3] 'Demand: Global electricity use to grow strongly in 2025 and 2026 – Electricity Mid-Year Update 2025 – Analysis', IEA. Accessed: Apr. 28, 2026. [Online]. Available: https://www.iea.org/reports/electricity-mid-year-update-2025/demand-global-electricity-use-to-grow-strongly-in-2025-and-2026

[4] R. and M. ltd, 'Marine Internal Combustion Engines Market Size & Competitors'. Accessed: Apr. 29, 2026. [Online]. Available: https://www.researchandmarkets.com/reports/6006114/marine-internal-combustion-engines-market

[5] H. Krause *et al.*, 'Factory Acceptance Test for Serenity-S1, a CMS Phase-2 back-end card', *J. Inst.*, vol. 21, no. 03, p. C03035, Mar. 2026, doi: 10.1088/1748-0221/21/03/C03035.

[6] H. Yamashita and T. Altiok, 'Buffer capacity allocation for a desired throughput in production lines', *IIE Transactions*, vol. 30, no. 10, pp. 883–891, Oct. 1998, doi: 10.1080/07408179808966542.

[7] S. Tyagi, A. Choudhary, X. Cai, and K. Yang, 'Value stream mapping to reduce the lead-time of a product development process', *International Journal of Production Economics*, vol. 160, pp. 202–212, Feb. 2015, doi: 10.1016/j.ijpe.2014.11.002.

[8] J. Keilson and L. D. Servi, 'A distributional form of Little's Law', Oct. 1988.

[9] 'Internal document, case company'.

[10] 'ISO 15550:2016', ISO. Accessed: Apr. 28, 2026. [Online]. Available: https://www.iso.org/standard/70030.html

[11] D. A. M. Doggett, 'A Statistical Comparison of Three Root Cause Analysis Tools', *Journal of Industrial Technology*, vol. 20, no. 2, 2004.

[12] D. Mahto and A. Kumar, 'Application of root cause analysis in improvement of product quality and productivity', *JIEM*, vol. 1, no. 2, pp. 16–53, Dec. 2008, doi: 10.3926/jiem.2008.v1n2.p16-53.

[13] J. R. Dew, 'In Search of the Root Cause', *Quality Progress*, no. 24, pp. 97–102, Mar. 1991.

[14] B. Sproull, *Process Problem Solving: A Guide for Maintenance and Operations Teams*. New York: Productivity Press, 2001. doi: 10.4324/9780367807733.

[15] P. Beard, 'Canadian Root Cause Analysis Framework', 2005.

[16] L. Liliana, 'A new model of Ishikawa diagram for quality assessment', *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 161, p. 012099, Nov. 2016, doi: 10.1088/1757-899X/161/1/012099.

[17] A. Kumah *et al.*, 'Cause-and-Effect (Fishbone) Diagram: A Tool for Generating and Organizing Quality Improvement Ideas', *Global Journal on Quality and Safety in Healthcare*, vol. 7, no. 2, pp. 85–87, May 2024, doi: 10.36401/JQSH-23-42.

[18] M. Barsalou and B. Starzyńska, 'Inquiry into the Use of Five Whys in Industry', *QIP Journal*, vol. 27, no. 1, pp. 62–78, Mar. 2023, doi: 10.12776/qip.v27i1.1771.

[19] P. Gangidi, 'A systematic approach to root cause analysis using 3 × 5 why's technique', *IJLSS*, vol. 10, no. 1, pp. 295–310, Mar. 2019, doi: 10.1108/IJLSS-10-2017-0114.

[20] A. J. Card, 'The problem with "5 whys"', *BMJ Qual Saf*, vol. 26, no. 8, pp. 671–677, Aug. 2017, doi: 10.1136/bmjqs-2016-005849.

[21] J. De Mast, 'Diagnostic Quality Problem Solving: A Conceptual Framework and Six Strategies', *Quality Management Journal*, vol. 20, no. 4, pp. 21–36, Jan. 2013, doi: 10.1080/10686967.2013.11918362.

[22] C. Christensen, K. M. Betz, and M. S. Stein, *The Certified Quality Process Analyst Handbook, Second Edition*. Quality Press, 2013.

[23] G. F. Smith, 'Determining the Cause of Quality Problems: Lessons from Diagnostic Disciplines', *Quality Management Journal*, vol. 5, no. 2, pp. 24–41, Jan. 1998, doi: 10.1080/10686967.1998.11918852.

[24] R. Lenort, D. Staš, D. Holman, and P. Wicher, 'A3 Method as a Powerful Tool for Searching and Implementing Green Innovations in an Industrial Company Transport', *Procedia Engineering*, vol. 192, pp. 533–538, 2017, doi: 10.1016/j.proeng.2017.06.092.

[25] A. George, S. Ranjha, and A. Kulkarni, 'Enhanced problem solving through redefined 8D step completion criteria', *Quality Engineering*, vol. 33, no. 4, pp. 695–711, Oct. 2021, doi: 10.1080/08982112.2021.1969665.

[26] S. J. Benjamin, M. S. Marathamuthu, and U. Murugaiah, 'The use of 5-WHYs technique to eliminate OEE's speed loss in a manufacturing firm', *Journal of Quality*

in Maintenance Engineering, vol. 21, no. 4, pp. 419–435, Oct. 2015, doi: 10.1108/JQME-09-2013-0062.

[27] B. Flyvbjerg, ‘Five Misunderstandings About Case-Study Research’, *Qualitative Inquiry*, vol. 12, no. 2, pp. 219–245, Apr. 2006, doi: 10.1177/1077800405284363.

[28] M. D. Fetters, L. A. Curry, and J. W. Creswell, ‘Achieving Integration in Mixed Methods Designs—Principles and Practices’, *Health Services Research*, vol. 48, no. 6pt2, pp. 2134–2156, Dec. 2013, doi: 10.1111/1475-6773.12117.

[29] A. Bryman, ‘Integrating quantitative and qualitative research: how is it done?’, *Qualitative Research*, vol. 6, no. 1, pp. 97–113, Feb. 2006, doi: 10.1177/1468794106058877.

[30] C. Schröer, F. Kruse, and J. M. Gómez, ‘A Systematic Literature Review on Applying CRISP-DM Process Model’, *Procedia Computer Science*, vol. 181, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.

[31] S. K. Kwak and J. H. Kim, ‘Statistical data preparation: management of missing values and outliers’, *Korean J Anesthesiol*, vol. 70, no. 4, p. 407, 2017, doi: 10.4097/kjae.2017.70.4.407.

[32] J. Van Den Broeck, S. Argeseanu Cunningham, R. Eeckels, and K. Herbst, ‘Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities’, *PLoS Med*, vol. 2, no. 10, p. e267, Sep. 2005, doi: 10.1371/journal.pmed.0020267.

[33] Microsoft Corporation, ‘Microsoft Excel’. Accessed: Mar. 03, 2026. [Online]. Available: https://www.microsoft.com/en-us/microsoft-365/excel

[34] M. Thelwall and T. Nevill, ‘Is research with qualitative data more prevalent and impactful now? Interviews, case studies, focus groups and ethnographies’, *Library & Information Science Research*, vol. 43, no. 2, p. 101094, Apr. 2021, doi: 10.1016/j.lisr.2021.101094.

[35] M. J. McIntosh and J. M. Morse, ‘Situating and Constructing Diversity in Semi-Structured Interviews’, *Global Qualitative Nursing Research*, vol. 2, p. 2333393615597674, Jan. 2015, doi: 10.1177/2333393615597674.

[36] S. Elo and H. Kyngäs, ‘The qualitative content analysis process’, *Journal of Advanced Nursing*, vol. 62, no. 1, pp. 107–115, Apr. 2008, doi: 10.1111/j.1365-2648.2007.04569.x.

[37] A. Grosfeld-Nir, B. Ronen, and N. Kozlovsky, ‘The Pareto managerial principle: when does it apply?’, *International Journal of Production Research*, vol. 45, no. 10, pp. 2317–2325, May 2007, doi: 10.1080/00207540600818203.

[38] P. Raghuram, A. Srikanth, and P. Rithan Mandesh, 'Eliminating End of Line Rejections-a quality filter mapping approach', *IJSOM*, vol. 1, no. 1, p. 1, 2021, doi: 10.1504/IJSOM.2021.10044246.

[39] V. Braun and V. Clarke, 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, Jan. 2006, doi: 10.1191/1478088706qp063oa.

[40] S. K. Ahmed *et al.*, 'Using thematic analysis in qualitative research', *Journal of Medicine, Surgery, and Public Health*, vol. 6, p. 100198, Aug. 2025, doi: 10.1016/j.glmedi.2025.100198.

[41] C. Antaki, M. Billig, D. Edwards, and J. Potter, 'Discourse analysis means doing analysis: a critique of six analytic shortcomings', 2003.

[42] R. E. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development*. SAGE, 1998.

[43] 'Microsoft Teams'. Accessed: May 07, 2026. [Online]. Available: https://www.microsoft.com/fi-fi/microsoft-teams/group-chat-software

[44] 'Microsoft Forms'. Accessed: Apr. 30, 2026. [Online]. Available: https://www.microsoft.com/en-us/microsoft-365/online-surveys-polls-quizzes

[45] Employee 1, 'Chief Technical Expert', Dec. 2025.

[46] Employee 2, 'General Manager (LB testing)', Feb. 2026.

[47] Employee 3, 'General Manager (Testing)', Dec. 2025.

[48] Employee 4, 'Development Engineer', Dec. 2025.

[49] Employee 5, 'Coordinator', Jan. 2026.

[50] Employee 6, 'Assembler (FAT)', Feb. 2026.

[51] Employee 7, 'Assembler', Jan. 2026.

[52] Employee 8, 'Assembler', Jan. 2026.

[53] Employee 9, 'Assembler', Jan. 2026.

[54] Employee 10, 'Assembler', Jan. 2026.

[55] Employee 11, 'Assembler', Jan. 2026.

[56] 'BUL-4015-1-UK'. [Online]. Available: https://www.parker.com/content/dam/Parker-com/Literature/HPCE/New/BUL-4015-1-UK.pdf

[57] Employee 3, 'General Manager (Testing)', Feb. 2026.

[58] M. Paradies and D. Busch, 'Root cause analysis at the Savannah River Plant', *Conference on Human Factors and Power Plants*, vol. 1988, pp. 479–483.

Appendices

Appendix 1: Excel formulas

List all engine numbers:

- =SORT(UNIQUE(Data2[Engine number]))

Model (Data 2):

- B2=LEFT(C2,3)

Engine class (Data 2):

- D2=LEFT(B2,2)

Slot duration (Data 2):

- I2=H2-G2+1

Extract number of tests per engine model:

- =COUNTIF(Data2[Model],A3)
    - A3 = model of the engine we are looking for

Extract the number of engines per model while including only customer engines:

- =SUMPRODUCT(N(XLOOKUP(UNIQUE(FILTER(Data2[Engine number],(Data2[Test slot type]<>K$5)*(Data2[Test slot type]<>K$9))),Data2[Engine number],Data2[Model],”-”,0)=A45))
    - A45 = model of the engine we are looking for
    - K$5 and K$9 are test slot types that are excluded

Extract all models in sorted order excluding the lab engines:

- =LET(values,UNIQUE(Data2[Model]),model,LEFT(values,2),n,RIGHT(values,1),SORTBY(values,model,-1,n,1))

Extract and sort the issue categories based on their appearance quantity (LB engines):

- =LET(d,FILTER(Data1[Category],(Data1[Category]<>““)*(ISNUMBER(MATCH(Data1[Engine number],UNIQUE(FILTER(Data2[Engine number],Data2[Engine class]="LB")),0)))),s,UNIQUE(d),v,HSTACK(s,MAP(s,LAMBDA(x,SUMPRODUCT(N(d=x))))),SORTBY(v,INDEX(v,,2),1))

Appendix 2: Thematic analysis themes and codes

The theme is in the top cell, and the corresponding codes are under it.

Study background and related infromation		
Factory capacity	Data info	Lead times

Observations made about engines		
Design notes	Model specific opinions	Typical faults

Operations control			
Scheduling	Communication	PMCS	Unclear instructions

Existing processes effect on factory operations	
Comments about the factory processes	Addressing faults at T&V

Personnel's effect on factory operations		
Problems in staffing	Employee related observations	Activities not alligning with official guidelines

Facility's effect on factory operations	
Logistics of assembly	Facility restrictions

Conclusions made by the employees	
Reasons for issues	Suppliers

Other notes
Other

Appendix 3: Questionnaire

1. Työnimike *

- ◯ Asentaja
- ◯ Koordinaattori

2. Minkä kokoisten moottorien kanssa työskentelet pääsääntöisesti? *

- ◯ Small Bore
- ◯ Medium Bore
- ◯ Large Bore

3. Arvioi millä tekijällä/tekijöillä on suurin vaikutus vuotaviin Ermeto-liitoksiin, jotka esiintyvät koeajossa. Valitse 1–2 merkittävintä. *

Valitse enintään 2 vaihtoehtoa.

- ☐ Asennusvirhe. Esim. putkea ei asetettu liittimen pohjaan asti tai kiristys unohtunut
- ☐ Kiristämiseen liittyvän koulutuksen/tiedon puute
- ☐ Liittimen valmistusvirhe (liitin on jo valmiiksi viallinen)
- ☐ Osien uudelleenasennus (jo kiristetty osa on jouduttu irroittamaan)
- ☐ Vaikea/ahdas asennuslokaatio
- ☐ Väärä kiristysmetodi (liitos on kiristetty ranteella virallisen ohjeen sijaan)

4. Jokin muu syy, jota ei ole mainittu yläpuolella? (vapaaehtoinen)

Kirjoita vastaus

5. Kiristetäänkö Ermeto-liitokset pääsääntöisesti käyttäen virallista kiristysmenetelmää vai rannetuntumalla? *

- ◯ Virallisella menetelmällä (sormikireydelle + 540°)
- ◯ Ranteella

6. Jos kiristetään ranteella, onko siihen jokin tietty syy?

Kirjoita vastaus

7. Joudutaanko jo asennettuja putkia irroittamaan [redacted] virheellisten kasausohjeiden vuoksi?
(Esim. tilanne jossa kasausjärjestyksessä ei ole otettu huomioon, että jotain osaa ei pystytä asentamaan paikalleen jos jokin aikasempi putki on jo laitettu kiinni) *

- ◯ Kyllä
- ◯ Ei

8. Jos vastasit edelliseen kyllä, kuinka usein?

- ◯ Alle 1 kertaa viikossa
- ◯ 1-2 kertaa viikossa
- ◯ 3-5 kertaa viikossa