Making success out of failures

A quantitative research in Failure culture and Quality improvement

Anna Bayramyan
Summary

Companies can draw valuable lessons from their failures, and use them for their improvement work. A positive failure culture is however a necessary precondition. High reliability organisations (HRO) are known for their effective way of tackling and using failures as potential for improvements. The aims of the study were to evaluate failure culture in an ISO 9001 certified company, after HRO standards and thereby withdraw improvement possibilities. The study was conducted through a deductive quantitative method using a survey for data gathering. With approximately 30 percent response rate, and using statistical tests, the failure culture of the company was evaluated. The results showed that the company is not currently reaching a failure culture within HRO standards, but nevertheless has positive tendencies rather than negative. Lastly, improvement possibilities were discussed and suggestions for further research were given.

Key words: Quality management, Failure, Errors, Failure culture, HRO, TQM
Acknowledgement

I wish to express my warmest gratitude to my supervisor, Daniel Carnerud, and the crew of Quality Management and Mechanical Engineering at Middle Sweden University, Östersund.

I would also like to pay my special regards to my supervisors in the investigated company.

Last but not least, I want to thank my family and friends for their support.

Anna Bayramyan
24 May 2020
# Table of Contents

1. Introduction  
   1.1 Aims  
2. Literature review (Theory)  
   2.1 Quality Management and continuous improvement  
   2.2 Self-designed values, beliefs and behavioural norms  
   2.3 High Reliability Organisations and failure culture  
   2.4 The five principles of HROs  
3. Design and Methodology  
   3.1 Research hypotheses and model  
      3.1.1 Hypothesis 1  
      3.1.2 Hypothesis 2  
      3.1.3 Hypothesis 3  
      3.1.4 Hypothesis 4  
4. Results and analysis  
   4.1 Rate of response and representativeness  
   4.2 Hypotheses analysis  
      4.2.1 Hypothesis 1  
      4.2.2 Hypothesis 2  
      4.2.3 Hypothesis 3  
      4.2.4 Hypothesis 4  
5. Conclusions  
6. Discussion  
7. List of sources
List of tables

Table 1, Procedure - Questionnaire........................................................................................................... 15
Table 2, Summary of results of Cronbach's Alpha ....................................................................................... 18
Table 3, Dependent and independent variables......................................................................................... 18
Table 4, Result 1 Hypothesis 1..................................................................................................................... 23
Table 5, Result 2 Hypothesis 1: One-Sample T Test .................................................................................. 24
Table 6, The measures of central tendency of Failure culture categories .................................................. 24
Table 7, Significance test Hypothesis 2 ....................................................................................................... 25
Table 8, Pearson's correlation test on Failure culture categories ................................................................. 27
Table 9, Measures of central tendencies divided between genders............................................................... 28
Table 10, Significance test Hypothesis 4 ..................................................................................................... 29
Table 11, Comparison between genders for Detailed approach to mistakes and Delegation..................... 30
Table 12, Results of the hypotheses, summary ......................................................................................... 31

List of figures

Figure 1, The Cornerstone Model of TQM developed by Bergman & Klefsjö ........................................ 3
Figure 2, Failure culture in HROs ............................................................................................................... 17
Figure 3, comparison of age and work experience from sample and population ....................................... 22
Figure 4, comparison Y1, Y3, Y5, Y6 between managers and workers ..................................................... 26
Figure 5, distribution for Reputation and Appreciation and Feedback ....................................................... 27
Figure 6, Simple Boxplot showing the answers distribution for the category Safety feeling..................... 28

Word definitions

The words failure, error and mistake are in this thesis used as synonyms.

The words mean and average are synonymously used as measure for central tendency.

ISO 9001:2015 - The quality management system developed by the International Organisation for Standardisation is commonly used as a tool to help improving the performance and sustainability of organisations, their products and processes (Bergman & Klefsjö, 2010).
1. Introduction

Research has shown that failures lead to higher learning values than successes (Madsen & Desai, 2010; Vogus, 2012). While success often leads to celebrations, failures on the other hand often lead to examination of what went wrong and lessons learned. Consequently, such events will be minimized in future (Sitkin, 1992).

Within quality development theories the topic of managing failures, errors and problems is commonly discussed. When discussing the theory of continuous improvement, Sörqvist (2004) mentions the importance of identifying and solving problems and mistakes in terms of possibilities to improve. Sörqvist describes improvements as breakthroughs, where organisations go through a relative change, from a current result level to a better, through decreasing the current level of accepted problems (ibid).

Many organisations use quality models such as Lean and Total Quality Management (TQM) as an approach in their improvement work (Sörqvist, 2004). However, Jeffrey Liker (2009) also highlights that many organisations do not succeed implementing the “Lean thinking”, which has a lot deeper cultural aspect. This perspective is lifted in Mike Rother’s book Toyota Kata (2013), where he describes the “katas” - the habits and routines that are being used in Toyota leading to their successful improvement work. The author distinguishes between the Improvement kata and the Coaching kata. One original way of thinking within the Improvement kata is “no problems = a problem” (Rother, 2013: 171). Toyota is more interested in the actions that do not go as planned rather than flawless work (ibid). The logic behind this is that, in a case where there are no problems, or when the workers make it to seem so, this results in the company standing still. Furthermore, the leadership of Toyota encourages the detection and use of small problems in order to capitalise on potential they reveal (ibid). In a case when people feel threatened by problems and failures, they will either try to hide them or try to come up with quick countermeasures leading to bad solutions without any deeper analyses (ibid). Also, TQM-programmes often fail to have the wanted effect on improvement work, since they do not provide structured solutions for the improvement work and mindful organisation (Sörqvist, 2004).

An interesting approach to research failure culture is the theory around High Reliability Organizations (HROs). Weick, Sutcliffe and Obstfeld (1999) argue that HROs are useful to study because of their processes and mindfulness that help the organizations to foster “effectiveness under trying conditions” (ibid: 31). The authors also state that improvement programs such as TQM often fail because of the underdevelopment of cognitive infrastructure and mindful organization, which on the other hand are characteristic for HROs (ibid). According to the authors there are some overlap and parallels between the improvement practices such as TQM, mindful error detection and correction, and the goals in HROs (ibid).
Given this background, the HRO theory can be used to advantage as a binding bridge between quality management and an efficient and positive failure culture in an organization. However, Weick and colleagues (Weick, Sutcliffe & Obstfeld, 1999) suggest that there are some discrepancies between the two frameworks. The authors state for instance that the total quality programs neutralize the mindfulness due to their formalities (ibid).

Resulting from this, the following purpose was formulated.

1.1 Aims

The aim of the thesis is first to investigate failure culture, using HRO standards, in an organisation that is certified with ISO 9001:2015, hence working with continuous improvement. Secondly, on the basis of the results, the aim is to find suggestions for improvement possibilities in the area of failure culture, which can be used in the improvement work of the organisation.

The aim will be reached by testing the following hypotheses. A more detailed description of how these were chosen is found under Design and Methodology.

**Hypothesis 1**
H0 - The mean values of the categories of HRO’s failure culture in the investigated company are equal to or bigger than 2.
H1 - The mean values of the categories of HRO’s failure culture categories in the investigated company are smaller than 2.

**Hypothesis 2**
H0 - Managers and workers work equal according to the categories of HRO’s failure culture.
H1 - Managers work more according to the categories of HRO’s failure culture than workers.

**Hypothesis 3**
H0 - Experienced and less experienced workers (of the company) work equal according to the categories of HRO’s failure culture.
H1 - Experienced workers (of the company) work more according to the categories of HRO’s failure culture than less experienced.

**Hypothesis 4**
H0 - Women and men work equal according to the categories of HRO’s failure culture than men.
H1 - Women work less according to the categories of HRO’s failure culture than men.
2. Literature review (Theory)

2.1 Quality Management and continuous improvement

Quality improvement is in organisation’s interest for several reasons, one of them being the relation between quality and profitability (Bergman & Klefsjö, 2010). Bergman and Klefsjö describe quality management as the aim to maximise customer satisfaction with minimised amount of resources (ibid). The authors present Total Quality Management (TQM) as a holistic concept that works as a framework for quality issues of the activities in an organisation - in other words for quality management (Bergman & Klefsjö, 2010). More precisely the authors describe TQM as “a constant endeavour to fulfil, and preferably exceed, customer needs and expectations at the lowest cost, by continuous improvement work, to which all involved are committed, focusing on the processes in the organization.” (ibid: 37).

To make the work of TQM successful, the authors suggest building it up through the following cornerstone model and support the cornerstones with relevant methodologies and tools for their operationalisation (Bergman & Klefsjö, 2010).

![Figure 1, The Cornerstone Model of TQM developed by Bergman & Klefsjö](from Bergman & Klefsjö, 2010: 38).

The cornerstone in the centre, “Focus on customers”, refers to meeting the customers’ wants and needs (Bergman & Klefsjö, 2010). This is done by actively studying what the customer wants in order to develop and produce the respective service or product (ibid).

“Focus on processes” underlines the importance of identifying the processes that obtain customer satisfaction, while using as little resources as possible (ibid). This is fulfilled by identifying all parts of a workflow, documenting them and through this visualising the flow, roles, responsibilities and the values that are being created as outputs of the activities in the flow (ibid). The cornerstone “Improve continuously” implies that continuous improvement is necessary and important, considering that surrounding world is in constant growth and development, while the customer demands are growing (ibid). The work of continuous improvement is based on for example using tools and methods from Lean and TQM to eliminate deviations in the processes (ibid). Such examples are the PDCA cycle or improvement boards (ibid).
“Let everybody be committed” refers to the employees’ inclusion and participation in decision-making and overall commitment to the improvement work (ibid). This part is focused on getting all employees involved, since the workers, and not the process owners, are the ones seeing potential improvements in their everyday work (ibid). Different methods are used to strengthen participation in the process work. The overall intention is that the improvement groups will, over time, be a continuous part of the work (ibid). By “Base decisions on facts” it is meant that relevant information is actively gathered to be analysed and applied (ibid). When deviations are being identified, it is important to break them down clearly and seek their roots. The authors suggest various tools and methods of cause-effect analysis, such as the Ishikawa diagram and the 5-Why method (ibid). In this way agreements on the problems and respective solutions can be reached, based on facts. Lastly, “Committed leadership” in the bottom of the model refers to the importance of a strong and committed leadership, which creates the right culture in the organisation for reaching quality improvements. The cornerstones are resting on the foundation of committed leadership (ibid).

Lars Sörqvist (2004) discusses continuous improvement as part of different quality models, such as TQM, Lean and ISO 9000. The author talks about different types of improvements that organisations can make, referring to them as either preventive or reactive improvements. The preventive improvements are realised before a problem occurs and the reactive improvements are made after a problem has arisen. The researcher further suggests, that the smaller breakthroughs, so called incremental solutions, bear the most impacts on the improvement work in the company because of their vast numbers (ibid). An important part in the improvement work therefore consists of solving problems by identifying and handling the system-related and individual/people-related problems (ibid). One cause for the mistakes made by individuals, which the author mentions is communication errors. Further, to achieve success and efficiency in an improvement work, the work must be based on the organisation as a whole. By this, Sörqvist argues that the organization must use a cross-functional process approach (Sörqvist, 2004).

While Bergman and Klefsjö as well as Sörqvist put a lot of focus on describing the different methodologies and tools that can be used in order to support a given set of values, Mike Rother (2013) in his book “Toyota Kata” describes the behaviour routines and patterns of Toyota. By this, the researcher attempts to describe how Toyota attains continuous improvement. This also shows the necessary behaviours and ways of thinking for reaching improvements also for companies outside Toyota (ibid). Rother does this by introducing the reader to the two Katas - two routines or ways of doing - the Improvement Kata and the Coaching Kata. According to the author those are the invisible aspects within Toyotas improvement work, unlike for example the different tools that Toyota works with (ibid). In the long run the Katas create the culture of the organisation (ibid).
2.2 Self-designed values, beliefs and behavioural norms

Schein’s (2004) description of culture consists of three elements. Firstly, culture offers structural stability in the form of meaning and predictability in the organisation. It stays the same even in the case of some employees leaving the organisation. Secondly, the culture resides in the subconscious part of a group and thirdly it influences how an organisation deals with its operations (ibid). Gary Yukl, inspired by Schein’s definition, suggests that an organisational culture consist of the shared assumptions, beliefs and values of the members (Yukl, 2013). These values and beliefs support members to deal with internal and external problems by helping people to understand the environment and know how to respond to it (ibid). In order for the organisation to prosper, these values and beliefs need to be consistent with and match to the organisational strategies, otherwise a strong corporate culture can be a weakness (ibid). For example, learning organisations, which learn and use this learning to become more effective, have the values of learning, innovation, flexibility etc. embedded in their culture (ibid). When solutions to problems are developed through experience, they are later spread to other members of the organisation in form of assumptions, which in the long run can turn into unconscious behaviour (ibid).

The leader’s role is also highlighted here. Yukl (2013) highlights the leaders’ role and influence on the organisational culture. Leaders need to model and communicate the values of the organisational culture (ibid). While the leaders in an organisation are themselves influenced by the organisation’s culture, they can influence it back over time (ibid). Leaders formulate visions, values and goals for the organization and how to achieve them, and thus communicate the values and ideals that are important. In order for these to be perceived as credible, the leader must not only write down the statements, but also show in action that they are important (ibid). In quality management theories (Liker, 2009; Bergman & Klefsjö, 2010), an example for this are the so called “Gemba walks”. The leaders go and see in person and in this way are able to better interact with the workers. This provides the opportunity to actively teach workers how the problems can be handled, where the leaders act as mentors for the workers. The obstacles are tackled with help of small PDCA-cycles, towards a vision using successive “target conditions”. This represents a leadership according to Toyota, in contrast to the traditional management systems where the leaders define goals on their own and control through hierarchical structures (Rother, 2013).

Another interesting position on culture combined with leadership is provided by Elmgart and Bloch (2017). The authors discuss the concepts “Culture by default” and “Culture by design”. The former is consistent with Yukl's and Schein's notions of a culture that emerges through norms and values as owners and managers, often unconsciously built into the organization. Culture by design instead means using purposeful work with culture, a strategy, as a way of promoting business development in the direction that management wants (Elmgart & Bloch, 2017). The latter would match the model suggested by Rother (2013) to change the organisational culture. Rother states that by practicing new behaviour patterns regularly, one can in the long run reach a norm that will build up the new culture. Through its Coaching
kata, Toyota is striving to reach a state of Improvement kata as the norm in the company (ibid).

2.3 High Reliability Organisations and failure culture

High Reliability Organizations (HROs) are the kind of organizations that succeed in coping with unexpected conditions, hence avoiding major accidents despite the constant high level of risk they are naturally facing in case of even a minor error (Weick & Sutcliffe, 2007). For instance, nuclear power plants and air traffic control systems belong to HROs (ibid). In order to mindfully manage the unexpected, Weick and Sutcliffe (2007) argue that culture is a key element.

Since the 1980s, HROs have been interested in organisational culture (Caldwell, 2018). Caldwell (2018) states that in the most literature “safety culture” is interpreted as an integrated part of the organisational culture. The term “safety culture” was coined by International Nuclear Safety Advisory Group with the aim of emphasizing the cultural and organisational factors that are as important for the safety of the organisations as the technologies they use (ibid). The definition of safety culture is, similarly to the definition of organisational culture, made up attitudes, perceptions and workplace behaviours (ibid). The most accepted and predominated viewpoint by psychologists, management consultants, and organizational theorists is that culture can be managed and changed to influence outcomes such as safety and reliability (Caldwell, 2018).

Creating a positive safety culture is fundamental to the reliability of the High Reliability Organisations (Weick & Roberts, 1993). It is commonly accepted that the leadership has the primary influence also on safety culture (Schein, 2004). Leadership is critical when establishing a just culture, organisational learning, mutual engagement and problem anticipation (Cardwell, 2018). Having a just culture means having “prerequisites of open communication and personal accountability for safety”. The role of the leadership is to create a culture where failures and information are not suppressed and hidden from managers, by instead trying to resolve them within their workgroup (ibid). Like that also the organisation’s ability to learn from failures would decrease (ibid). Leaders support organisational learning by creating transparency, balanced and unbiased processing of information and by developing high moral standards in the organisation. Mutual engagement covers the creation of teamwork and engaging with workers’ opinions (ibid). Lastly, the problem anticipated from the leaders’ side is about not ignoring warning signs that can lead to bigger problems (ibid).

Weick and Sutcliffe (2007) argue that culture shapes the actions and the way they are interpreted. The researchers see the culture through three views: “shared values”, which means what is important around here, “beliefs” which means how things work around here and “structures and control systems to produce behavioural norms”, which they describe as “the way we do things around here” (ibid). Further the authors suggest that organisations act their way into what they become (ibid). These norms and concrete actions and practices that
are present within HROs and construct their cultures are presented under the following subtitle, divided along the five principles that HROs work along.

2.4 The five principles of HROs

In a dynamic, unknowable, unpredictable world one might presume that organizing in a manner analogous to HROs would be in the best interest of most organizations. Hints of such moves are evident when traditional organizations graft TQM cultures onto a pre-existing preoccupation with efficiency, and aspire to the relatively error-free performance found in HROs. But many of these attempted changes fail because traditional organizations demonstrate little awareness of just what kind of infrastructure it takes to support reliable performance. Unfortunately, mainstream organizational theory isn’t much help in developing this awareness. (Weick, Sutcliffe & Obstfeld, 1999).

There are five key ideas of the mindful infrastructure, so called principles of Highly Reliable Organizations. The three first principles are focused on the HRO’s capability to anticipate “unexpected problems”, whereas the last two focus on the capacity to contain them (Weick & Sutcliffe, 2007: 9).

Principle 1: Preoccupation with Failure

In an HRO no failures are ignored or worked around, because any deviation could lead to a tragedy. The failures are therefore addressed immediately and wholly. To avoid failures, HROs “embrace them” to begin with (Weick & Sutcliffe, 2007: 46). By this the authors suggest that the HROs on one hand pay close attention to weak signals of failure that could be symptoms of bigger problems and on the other hand spell out mistakes, often strategic mistakes, that the employees do not dare to make (ibid: 46). By actively searching for symptoms of possible mistakes the organisations are better prepared to prevent them (ibid). In addition, the earlier the symptoms are found, the more alternatives there are to deal with the problem (ibid).

The authors stress that detecting failures is not a simple task and suggest for example using checklists that help to find situations with shaky practices or where unforeseen events might occur due to human-system interface (Weick & Sutcliffe, 2007: 48). Furthermore, the authors highlight the importance of an existing “safety culture”, which also allows the practice of the other four principles of high reliability (ibid). In addition, the HROs also clearly communicate who in the organisation is there for making observations, similar to “maintenance personnel” who can more easily spot unexpected problems and improve organisational learning (ibid 48-49), while at the same time noting the importance that everyone is open about their own failures. Lastly, it is important to see the whole organisation as a system in order to prevent silos caused by working in a function-oriented way instead of system-orientated (ibid).
To report failures, instead of ignoring or covering them up, the employees need a feeling of safety in the organisation. This is an essential part of an organisational culture that encourages and values reporting of failures. Here the managerial practices are prevalent. For instance, asking questions and reporting occurred incidents needs to be encouraged and rewarded, even in a case where a failure was committed by oneself (Weick & Sutcliffe, 2007: 50). An open climate is necessary to allow the employees to speak freely about the errors and work on correcting them, and therefore learn more about the processes of the system (ibid: 51).

Lastly, the authors point out that “preoccupation with failure” needs to be seen in terms of reliable performance in a system, in other words a what and not an individual issue connected to a who. The people in HROs are in contrast described as “sceptical, wary, suspicious of quiet periods” (Weick & Sutcliffe, 2007: 52). Moreover, they do not let success of the organisation to equate to competence, which could lead to drifting back to inattention and predictable routines in the future. This requires eyes to be kept open at all times (ibid).

**Principle 2: Reluctance to Simplify**

The second principle emphasizes the topic of reducing simplifications when faced with failure (Weick & Sutcliffe, 2007: 56). Here the authors mention the importance of interaction among different people with different experiences and expertise and their diverse expectations, within the complex system-wide organisations (ibid). The variety in expertise can be reached by selecting employees with non-typical previous education and skills, job rotations and training. This contributes positively to the registering and handling of complexity within the “same event” in the complex system (ibid).

Another simplification to keep in mind is the creation of linguistic categories. In this way the knowledge gathered by direct exploration is transformed to names and then further transformed to descriptions rather than acquaintance. “Details get lost in generics. And they get lost in vagueness” (ibid: 58). Therefore, it is preferable to avoid communications with vague phrases. To deal with such simplifications, the HROs try to preserve the details. This can be done by differentiating them into categories and to both believe and doubt in the categories (Weick & Sutcliffe, 2007).

**Principle 3: Sensitivity to operations**

By principle three, the authors refer to the HROs’ ability to respond to “the messy reality inside most systems” by being sensitive to real-time operations and make changes along the time when necessary (Weick & Sutcliffe: 59, 61). This is achieved by conducting interdisciplinary and interdepartmental meetings. The reasons behind this is to create credibility and trust one another when dealing with complex tasks and to understand the
interdependence of working in complex systems. This principle is about the work of responding to and monitoring the complicated system (ibid).

In HROs the leaders create the conditions for employees to openly and frequently communicate with them (Weick & Sutcliffe, 2007). Respect is shown to the employees and the information is taken seriously followed by feedbacks. Operations are being observed in real time and the outcomes are continuously evaluated to see if they meet the objectives of the organisation. They withdraw experience and therefore lessons from the operations in real time (ibid).

**Principle 4: Commitment to Resiliency**

This is the one of the two principles of containment, that help to confront already occurred unexpected events with practices that reduce unwanted outcomes (Weick & Sutcliffe, 2007). Despite the anticipation of unexpected events, the HROs also have to deal with them reactively after they have occurred and try to prevent them from causing even worse harm (ibid).

HROs prepare in advance for emergencies (Weick & Sutcliffe, 2007). They have clear means of communication and control while being comfortable improvising (ibid). Team members are trained on how to manage unexpected events and the system has the elasticity to maintain the work. Furthermore, they encourage flexibility in team members to accommodate changes in conditions or resources and value interactions and working together in cross-functional collaboration. An important part of this is encouraging people to learn about the system, transparent and widely known. By knowing the weaknesses of the system and the ways to manage them the problem can be corrected faster.

**Principle 5: Deference to Expertise**

According to the last principle, HROs create operations where the leadership shifts to people who have an answer to the particular problem. This is done by migrating the decision making to a person who has the “specific knowledge to the event” (Weick & Sutcliffe, 2007: 74), since the authority hierarchy does not match to the knowledge hierarchy of a particular question (Weick & Sutcliffe, 2007).
3. Design and Methodology

The current study is built on a deductive reasoning and a quantitative method (Creswell, 2014). It begins with a theoretical background used as a basis, followed by hypotheses where the hypotheses are proved with empirical tests (ibid). The chosen method for gathering relevant objective data is through a web survey with closed-ended questions and Likert scale. The research is then formed through statistical tests and correlational design, where correlational statistic is used to describe the levels of associations and relationships between the variables to test the hypotheses. The advantage of this method is that primary data is collected relatively quickly and inexpensively (Djurfeldt, Larsson & Stjärnhagen, 2010; Persson, 2016). In addition, the method builds in protection against prejudice by deductively testing the theories (Creswell, 2014).

Moreover, the study design was conducted in accordance with “Occam’s Razor” principle. The principle, also known as the principle of parsimony, is useful when “defining methods for empirical inquiry, selecting scientific hypotheses, and refining statistical models” (Motl, 2010). The logic behind the principle is to choose the tool with fewer working parts over another with many, provided that they both measure equally well (ibid). In the same way, a straightforward answer to a question should be preferred before a complicated one with many contingencies (ibid). A correlation design may therefore be more appropriate than an experimental design (ibid). Hypotheses tests can be conducted with a small number of variables, from which one can withdraw clear relations. This matches Occam's razor rule in the sense of keeping explanations simple. “Correct use of the Razor at each step of the research process will reduce any final product to its most efficient, elegant, and consumable form, maximizing accuracy and reducing error.” (ibid: 955).

The target population of the research consists of all members of the headquarter of a German middle-sized company, which amounts to 958 employees. The company is certified with ISO 9001:2015. Subsequently the survey was released on the intranet of the company, addressing the employees of the headquarter of the company, leading to a simple random sampling (Esaiasson, 2012). A follow up message was posted on the intranet after approximately half of the business days that the survey was available for replying. Together with the reminder, a message was posted to the forum used for internal communication. The strategic reasoning for this was to enlarge the chance of reaching the younger members of the company, since the second platform is more dynamic and can be easily handled with a smartphone. The total duration of the survey was two and half weeks, but only eleven business days, as a result of the public holidays in between. Together with the internet link to the survey, an information text about the research was addressed to the respondents. The original text can be found in Appendix 1. Thus, the survey was personalized in order to reach higher participation rate. Here it was also explained, that the respondents should answer according to their own personal views and not represent any groups (Persson, 2016: 58). The questions were constructed accordingly. To catch the respondent’s interest and to inform the respondents about the research an introduction text was created and presented at the beginning of the questionnaire. The introduction was kept short, with no additional description of how to fill in
the survey, since the tool is easy to use and it was not possible to skip questions. The original version of the introduction is placed in Appendix 2.

Scaling

According to Colman and Preston (2000) scales with 10 response categories are generally preferred by respondents, followed by seven and nine response categories. Furthermore test-retest reliability decreases for scales with more than 10 categories (ibid). Also, Persson (2016) emphasises the importance of not having too many answer alternatives, in order to avoid confusing the respondents, especially in a case where the answers are closed-ended and classified in a logical order, for example in a scale, “completely” to “not at all” (Prüfer, Vazansky & Wystup, 2003). This depends on the language limiting the scaling (ibid). In the handbook of the European Statistical System, five to nine categories of response options are recommended (Brancato, Macchia, Murgia & Signore, 2005). In response to this, it was decided that the current survey would include five answer alternatives within the ordinal scale. This means that there is a middle alternative, which is a potential risk for error in measurement, since it can be the “easiest alternative” to choose. However, it is still recommended to have a middle alternative and with five answer alternatives instead of only three, it is already seen to be less risky for measurement error (ibid). The goal was to create answer alternatives that will cover the full scale of possible replies and be exclusive (ibid). To avoid misunderstandings the answer alternatives were created to match to the scale dimension of the question (ibid: 67).

Validity and Reliability

It is important to reach accuracy in the tailoring of the indicators in the operationalisation of the theory. Esaiasson (2012: 64) highlights this issue in terms of “conceptual validity”. By this the author means the distance between the theoretical definition and the operationalised indicator (ibid: 65). In other words, the more complicated the concepts used in the theory, the harder it is to find indicators that can empirically reach an adequate measurement. For example, it is easier to find an accurate indicator for a concept like “voter participation” by measuring the amount of voting people than a concept like “power” or “democracy”, which does not have a direct measurement indicator (ibid). To address this issue the author is suggesting two strategies: “reasoning validity” and “empirical validity”. One possibility way to reason to a conceptual validity is through “face validity”, which means to choose the most obvious and straightforward way when creating indicators for less troublesome concepts (ibid: 67). Although it is hard to say that the concept of failure culture is easy to operationalise, this was done by creating the questionnaire based on a theoretical background, through breaking it down to relevant criteria and questions.

Triangulation is a way to increase validity (Bryman, 2011; Djurfeldt et.al., 2010). Triangulation can be used by combining different types of methods or combining different theories, to address the problem from different perspectives and by that avoid bias (ibid). For that reason, the theoretical background has been broadened to touch the topic of failure
culture from different perspectives and approaches. The methodology has however been limited to a survey study because of the scope of the thesis.

Reliability is reached through making sure that the chosen tool measures with precision, without random variations and mistakes (Esaiasson, 2012: 70). In the case of a quantitative research, such as survey, the reliability tends to be higher, due to the standardised questions (Djurfeldt et.al., 2010). However, the risk consists in that the questions could be misunderstood or poorly written (ibid). There are also some statistical measurements possible for proving the reliability, such as measurement of internal consistency of summated items and their scale (Esaiasson, 2012).

**Ethical perspectives**

The survey was created with the online application, Netigate, commonly used in the company. The application was chosen as the standard tool for anonymous and non-anonymous survey creations and regulatory agreement between the managing directors of the company and the works council was made. The application is owned and steered by a service provider and not the researched company itself. The questionnaires are processed and aggregated at the servers of the service provider. This information is provided on the website of the company (https://www.netigate.net/sv/legal/). Therefore, the repliers’ safety and anonymity can be considered as high.

The end result was consulted with and approved by the supervisors employed by the company. Moreover, previous to publishing the questionnaire, it was also approved by the works council of the company. This is a standardised procedure in the company before any surveys are conducted.

**Risks**

While creating the questions, attention was paid to the cognition of survey response process, through implementation of the four-step model within the Cognitive Aspect of Survey Methodology (CASM) approach (Tourangeau, 1984). The model was developed by Roger Tourangeau where he breaks down the mental processing of survey responses to the following four steps, where failures in the answers could occur (ibid):

- **Cognition** - the respondent must understand the instructions and questions in the manner intended by the researcher.
- **Retrieval** - the respondent must be able to remember the necessary information to answer the question.
- **Judgement** - the respondent must be able to form a judgement from the recalled information.
- **Response** - the respondent need to be able to make a choice, which matches best to his/her individual response, without misreporting because of different reasons.
It is important to be informed that mistakes can occur within the above mentioned four steps. To avoid potential mistakes to the maximum extent possible, the questions were constructed with simple wording to avoid misunderstandings. Furthermore, the questions did not require the recollection of particular events (Lavrakas, 2008; Persson, 2016; Tourangeau, 1984). The third step, from survey methodology perspective, according to the handbook of the European Statistical System, is connected to the decisions that are connected to the response categories (Brancato et.al., 2005). Choosing closed-ended reply alternatives is making the judgement process easier for the respondents, but entails the risk that some important categories may be forgotten, or that the respondents tend towards choosing the intermediate positions. Lastly, when the respondent chooses an option, it is important that the respondent feels safe to answer also the case that the questions touches on a sensitive topic. To avoid the insecurity, it was important to introduce the research as neutral and anonymous (Persson, 2016).

Another risk with this method is the potential low response rate. Two factors contributing to this risk are motivation to answer the questions and the burden that is being put on the respondents (Persson, 2016; Brancato, et.al., 2005). Both have a connection to an unnecessarily large number of asked questions (ibid). Regarding the current situation of an ongoing pandemic, which lead to home office working and therefore changed habits, a long questionnaire would lead to a burden. Together with the contact persons of the company, a decision was made to keep the questionnaire to a response time of five minutes. Further details are to be found in the next coming section.

Lastly, it is important to be aware of risks connected to biased sampling. Although the sample is random, it is important to keep in mind that people who decide to answer the questionnaire may have different interests than those who do not answer the questionnaire (Kovera, 2010).

**Procedure - Questionnaire**

When it comes to measuring the failure culture in an organisation the Error Orientation Questionnaire (EOQ) is an interesting approach to take. Rybowiak et al. (1999) have developed the questionnaire through a psychological perspective, by categorising it into eight scales that measure the attitude to and the coping with errors at work. The psychological approach is a valid starting point. The questions do not however cover important aspects of quality management. On this ground, the EOQ was not sufficient for the purpose of this thesis. Instead, a new questionnaire was developed from the HRO-theory. Weick and Sutcliffe (2007) suggest audits that they have constructed to help researchers to construct questions for evaluating their organisations and how well they work in accordance with HRO principles (ibid: 85). What is however important to keep in mind is that not every aspect of HRO theory is relevant to the non-HROs and in particular to the investigated organisation. For instance, nuclear power plants work according to the International Nuclear Events Scale (INES), which also requires near-miss events to be documented (International Atomic Energy Agency, 2014). In the investigated organisation, near-miss events are not relevant.
At first sight it can seem irrelevant to study HROs since the investigated company in this thesis is not facing crises of the same severe characters, and it can seem extreme to undertake such measurements. However, the HROs are an optimal benchmark, since they provide practices and routines, which can effectively cope with misspecifications, misestimation and misunderstanding (Weick & Sutcliffe, 2007: 18-19). Non-HROs can therefore learn from and adapt this approach, which by Weick and Sutcliffe is described as a “mindful infrastructure” affecting the culture of the company (ibid).

The creation of a home-grown questionnaire began by firstly exploring the theoretical background of HROs. The key elements of the principles in the theory were marked. After that, one of the seven management and planning tools, the tree diagram, was used to visualise the principles by breaking them down in categories. The tree diagram is a helpful tool for organising and visualising the outcomes of a particular event (Klefsjö, Eliasson, Kennerfalk, Lundbäck & Sandström, 1999). The titles of the principles were placed as headings in the tree diagrams. The next step was to develop questions relevant for the organisation to be examined, which fit to each of the key elements in the theory of HROs. The key elements of each principle together with belonging questions are presented in tree diagrams in Appendix 4.

The next step was to minimise the number of questions to only most essentials for the failure culture model. This was done by removing questions that were very similar to each other, that were self-evaluating, which is not recommended to be used in surveys (Persson, 2016), or did not correspond to values, beliefs and behavioural norms – which are the interesting parts for a model of failure culture (Yukl, 2013; Cardwell, 2018). The questions were also eliminated, when the topic behind the questions have other reasons than only values, beliefs and norms. Such a questions is “How common are job rotation in your team?”. Job rotations in teams is also impacted by the size and structure of the company. In the table below, the reason for eliminating the questions are written under column 4.

At this point, it should be noted that a decision was made to predominantly use the questions connected to the first principle of HRO theory - Preoccupation with failure - when creating the overall Failure culture model used in HROs. The decision was made pragmatically, steered by the need to keep the questionnaire short to achieve a higher reply rate. Furthermore, Principle 1 is also in substance more detailed, involving several different aspects of the organisational failure culture. While all five principles describe how to handle failures and how to avoid them, in terms of creating necessary environment Preoccupation with failure is seen as a “hallmark” of HROs (Weick & Sutcliffe, 2007: 53). Therefore, this principle describes the “safety culture” that is necessary in the HROs (Cardwell, 2018) best. However, relevant questions derived from the other principles were also added to create a full model that fully corresponds to the culture in HROs. The questions that were not eliminated correspond to the description of culture, which is shared values, beliefs and behavioural norms (Yukl, 2013; Cardwell, 2018).
<table>
<thead>
<tr>
<th>No.</th>
<th>Original questions to the key elements of the HRO principles</th>
<th>Cause for elimination</th>
<th>Reduced amount of questions used in the survey touching values, beliefs and behavioural norms (culture)</th>
<th>Categories for coming failure culture model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>To what extent do you feel safe to report errors that were caused by you?</td>
<td>Double-question to question 12 &amp; 18, where 12 &amp; 18 better correspond to values, beliefs and behavioural norms.</td>
<td>To what extent do you feel safe to report errors that were caused by you?</td>
<td>Safety feeling</td>
</tr>
<tr>
<td>2</td>
<td>To what extent do you feel safe to report errors that were caused by others?</td>
<td>To what extent do you feel safe to report errors that were caused by others?</td>
<td>Safety feeling</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>To what extent do you think that your reputation will be positively influenced if you communicate errors a) caused by yourself? b) caused by a close colleague? c) caused by colleagues from other departments?</td>
<td>To what extent do you think that your reputation will be positively influenced if you communicate errors a) caused by yourself? b) caused by a close colleague? c) caused by colleagues from other departments?</td>
<td>Reputation</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>To what extent is the identification of errors / problems part of your job?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>To what extent do you feel encouraged to discuss your errors/mistakes with your colleagues within the department?</td>
<td>To what extent do you feel encouraged to discuss your errors/mistakes with your colleagues within the department?</td>
<td>Safety feeling</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>To what extent do you feel encouraged to discuss your errors/mistakes with your colleagues across the departmental boundaries?</td>
<td>To what extent do you feel encouraged to discuss your errors/mistakes with your colleagues across the departmental boundaries?</td>
<td>Safety feeling</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>To what extent do you feel encouraged to ask questions about your workflow?</td>
<td>To what extent do you feel encouraged to ask questions about your workflow?</td>
<td>Safety feeling</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>To what extent do you think it is appreciated that you communicate mistakes?</td>
<td>To what extent do you think it is appreciated that you communicate mistakes?</td>
<td>Appreciation and feedback</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>To what extent are you aware of the contribution your work makes to the “name of company”</td>
<td>Self evaluation question, not recommended to be used in surveys.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>To what extent do you feel that you are being made personally responsible for your mistakes?</td>
<td>To what extent do you feel that you are being made personally responsible for your mistakes?</td>
<td>Safety feeling</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>To what extent do you feel that your mistakes are being systematically questioned?</td>
<td>Double-question to question 10 (counter question for check).</td>
<td>To what extent do you pay attention to system-related errors in the company?</td>
<td>Mistake recognition and reporting</td>
</tr>
<tr>
<td>12</td>
<td>To what extent do you pay attention to system-related errors in the company?</td>
<td>To what extent do you pay attention to system-related errors in the company?</td>
<td>Detailed approach to mistakes</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>How often are occurred errors being discussed in the team?</td>
<td>How often are occurred errors being discussed in the team?</td>
<td>Detailed approach to mistakes</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>How common are job rotations in your team?</td>
<td>Has too many other impacts, other than values, beliefs and behavioural norms.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>How common are participations in various trainings?</td>
<td>Has too many other impacts, other than values, beliefs and behavioural norms.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>How detailed are occurred errors being discussed in the team?</td>
<td>How detailed are occurred errors being discussed in the team?</td>
<td>Detailed approach to mistakes</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>How often do you participate in interdisciplinary/interdepartmental meetings to share information?</td>
<td>Not itself a value, belief or behavioural norm.</td>
<td>How often do you communicate system-related errors?</td>
<td>Mistake recognition and reporting</td>
</tr>
<tr>
<td>18</td>
<td>How often do you communicate system-related errors?</td>
<td>How often do you communicate system-related errors?</td>
<td>Mistake recognition and reporting</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>How often are you being encouraged from your managers to communicate them?</td>
<td>How often are you being encouraged from your managers to communicate failures/mistakes?</td>
<td>Safety feeling</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>How often do you get constructive feedback when you communicate a mistake?</td>
<td>How often do you get constructive feedback when you communicate a mistake?</td>
<td>Appreciation and feedback</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Are you well prepared to manage a mistake?</td>
<td>Self evaluation question, not recommended to be used in surveys.</td>
<td>Not itself a value, belief or behavioural norm.</td>
<td>Detailed approach to mistakes</td>
</tr>
<tr>
<td>22</td>
<td>To what extent do you in your company have a structured procedure for eliminating errors?</td>
<td>Not itself a value, belief or behavioural norm.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>To what extent are the work methods being adapted after an occurred failure, in order to avoid it in the future?</td>
<td>Has too many other impacts, other than values, beliefs and behavioural norms.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>How much expertise do you have, to be able making decisions within your work area?</td>
<td>Self evaluation question not recommended.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>How often are you (as the expert) being consulted before making decisions?</td>
<td>How often are you (as the expert) being consulted before making decisions?</td>
<td>Delegation decision making</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>To what extent do the decision maker follow your recommendations when making decisions?</td>
<td>To what extent do the decision maker follow your recommendations when making decisions?</td>
<td>Delegation decision making</td>
<td></td>
</tr>
</tbody>
</table>
To make one general model of positive and efficient failure culture a new tree diagram was created, containing the questions that were left after eliminating non-relevant ones, in the table above. “Failure culture of HROs” was put as a heading in this tree diagrams. By thoroughly examining the questions, they were put together in categories, shown in the last column in the table above. From the original questionnaire the following model for failure culture in Figure 2 was created.

Before running statistical tests, a decision was made to construct indexes from the different items created from the questionnaire in order to capture the complexity of the less concrete categories (Esaiasson, 2012; Rickards, Magee & Artino, 2012) of the model for a positive failure culture. This is also used as a technique to create “survey scale” to allow statistical calculations with the mean-value and tests like the Pearson’s correlation test, which measures if there is correlation between two variables (Rickards et al., 2012; Barnette, 2010). This was done by adding together the data gathered to the questions asked under the different categories. During this procedure however, some details were lost. After consideration, the decision was made to add the variable into the index, despite the loss in detail. The decision was based on the argument, that the failure culture is investigated wholly. The effect of answers to the individual questions would be reflected in the final values. The priority was instead put on delimiting the scope of the thesis to an adequate size.
Figure 2. Failure culture in HROs

- Safety feeling
  - To what extent do you feel encouraged to discuss your errors/mistakes with your colleagues across the departmental boundaries?
  - How often are you being encouraged from your managers to communicate failures/mistakes?

- Reputation
  - To what extent do you feel safe to report errors that were caused by you?
  - To what extent do you feel safe to report errors that were caused by others?

- Mistake recognition and reporting
  - To what extent do you think that your reputation will be positively influenced if you communicate errors a) caused by yourself?
  - b) caused by a close colleague?
  - c) caused by colleagues from other departments?

- Appreciation and feedback
  - To what extent do you think it is appreciated that you communicate mistakes?
  - How often do you get constructive feedback when you communicate a mistake?

- Detailed approach to mistakes
  - To what extent do you think it is appreciated that you communicate mistakes?
  - How often are occurred errors being discussed in the team?
  - How detailed are occurred errors being discussed in the team?

- Delegation (decision making)
  - How often are occurred errors being discussed in the team?
  - To what extent do the decision maker follow your recommendations when making decisions?
3.1 Research hypotheses and model

Research variables

Before adding the questions of the categories together a reliability test was necessary. Cronbach's Alpha was used to test the internal index scale consistency, where the reliability coefficient of 0,70 and higher is considered as acceptable (Esaiasson, 2012). The test was carried out in SPSS, like all the statistical analyses of this thesis. Before running Cronbach's Alpha, it was necessary to change the direction of answer alternatives of one of the questions. In this way it was ensured that all answers pointed in the same direction, where 1 represents the best value for failure culture model and 5 the worst (Djurfeldt et.al., 2010) in regard to HRO failure culture. The following table presents the results of the Cronbach's Alpha.

Table 2. Summary of results of Cronbach's Alpha

<table>
<thead>
<tr>
<th>Subscales</th>
<th>Cronbach's α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety feelings</td>
<td>0,707</td>
</tr>
<tr>
<td>Reputation</td>
<td>0,717</td>
</tr>
<tr>
<td>Mistake recognition and reporting</td>
<td>0,736</td>
</tr>
<tr>
<td>Appreciation and feedback</td>
<td>0,735</td>
</tr>
<tr>
<td>Detailed approach to mistakes</td>
<td>0,837</td>
</tr>
<tr>
<td>Delegation (decision making)</td>
<td>0,751</td>
</tr>
</tbody>
</table>

The result, being over 0,70 was satisfactory in order to create the indexes.

The created indexes could now be used in form of dependent variables Y, that are presented hereunder together with the independent variables X.

Table 3. Dependent and independent variables

<table>
<thead>
<tr>
<th>Dependent variables Y</th>
<th>Independent variables X incl. explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1: SafetyFeeling</td>
<td>X1: Gender - The gender alternatives were Male, Female, Other.</td>
</tr>
<tr>
<td>Y2: Reputation</td>
<td>X2: WorkExperience - This variable measures the amount of years spent in the investigated company. The variable does not capture the details of previous work experiences.</td>
</tr>
<tr>
<td>Y3: MistakeRecognitionReporting</td>
<td>X3: JobTitle - This variable distinguishes between managers and workers in the company.</td>
</tr>
<tr>
<td>Y4: AppreciationFeedback</td>
<td></td>
</tr>
<tr>
<td>Y5: DetailedApproach</td>
<td></td>
</tr>
<tr>
<td>Y6: Delegation</td>
<td></td>
</tr>
</tbody>
</table>

The following hypotheses were constructed with the above-mentioned principle of parsimony, Occam’s Razor. The approach was to build models with few numbers of parameters representing the collected data (Hegdé, 2010). The hypothesis will be tested with a 5% significance value. This means that a random error has with a 5% risk influenced the results of the sample that they are not reflecting the population. In the tests it is achieved with a p-value of 0,05 (Djurfeldt, 2009).
3.1.1 Hypothesis 1

Since the investigated company is certified with ISO 9001, it may be claimed that it is in the interest of the company to make organisational improvements (Hausmann & Scheel, 2013). The six created indexes, representing the categories, together describe the failure culture of HRO’s. Since the failure culture of HRO’s can be described as efficient and positive, it can be, as mentioned previously in this thesis, set as a benchmark. By statistically analysing the created indexes, it can be shown whether the failure culture of the investigated company is positive and efficient or not. In this measurement it is enough to, through a descriptive analysis, observe the measures of central tendency and dispersions, that are the mean and standard variation (Djurfeldt et.al., 2010). However, to statistically test the hypothesis, it is also necessary to conduct a One-Sample T Test, which shows a significance value that must be bigger than 0.05 in order to reject the null-hypothesis (ibid). For this a test value must be specified, with which the measured mean values must be compared. This means that a value of comparison must be given, which can be set as a limit for what represents a positive failure culture. Since the value 1 in the scale represents the best possible failure culture, a decision was made to set the limit at the value 2, since it represents a rather positive failure culture, also considering that the benchmark for the Failure culture model were HROs with extraordinary ways of working. Therefore, the following hypothesis was formed.

H0 - The mean values of the categories of HRO’s failure culture in the investigated company are equal to or bigger than 2.

H1 - The mean values of the categories of HRO’s failure culture in the investigated company are smaller than 2.

3.1.2 Hypothesis 2

Derived from the presented theory, the leaders’ role in the organisational and failure culture is highlighted. Also in the general quality management theories and approaches the leaders play an important role in the organisational improvements. It is therefore interesting to investigate whether there are differences between the manager’s and worker’s way of working. The following hypothesis was formed.

H0 - Managers and workers work equal according to the categories of HRO’s failure culture.

H1 - Managers work more according to the categories of HRO’s failure culture than workers.

The hypothesis 2 will therefore be tried, by testing if there are significant differences between the mean values of the two groups (managers and workers), by conducting a t-test within a bivariate analysis. (Djurfeldt et.al., 2010). This technique is appropriate to use when testing the difference between two categories of qualitative variables (ibid). The precondition for this test is to have a sufficient sample size (n>30). In this way it can safely be assumed that the distribution of the mean has a bell-shaped normal distribution (ibid), which is the case of the sample sizes used in this research.
Important values to observe when conducting a t-test are firstly the F-value together with its significance value. The F-value is based on the Levene's test and describes if the variance in the two different groups are alike or not. The null-hypothesis in this test is that the variances are alike. If the F-value is not significant (p≥0,05), then the null-hypothesis can be rejected. Depending on the result, the values can be read either from the row “Equal variances assumed” or “Equal variances not assumed” in the output box in SPSS (Djurfeldt et.al., 2010).

After deciding which row the t-value must be read in the output box, the significance, also known as the p-value, of the T-test must be observed. It can be read under Sig. (2-tailed) and has the limit of p≥0,05 for rejecting the null-hypothesis (Djurfeldt et.al., 2010).

3.1.3 Hypothesis 3

Further to this, it is relevant to investigate the interrelation between the grade of experience in the company, depending on the amount of working years in the company, and the work in regard to the failure culture. The theory mentions for example the importance of delegating some decision makings to relevant people of expertise (Weick & Sutcliffe, 2007). Therefore, the interrelation between work experience in the company (leading to more relevant expertise), independent variable X2: WorkExperience and the dependent variable Y6: Delegation is particularly interesting to measure. This can be measured by performing Pearson’s correlation analysis (Djurfeldt, 2009), which like the t-test must be tested for significance (ibid).

However, correlation analyses including all dependent variables are also interesting to perform in order to measure if increased years spent in the company have an effect on the work in terms of positive failure culture seen from all of the categories. An assumption is that a positive failure culture in the organisation would affect people’s way of working. Therefore, there should be a positive correlation between longer working time in the company and a positive failure culture.

Hence, the following hypothesis was constructed.

H0 - Experienced workers (of the company) do not work more according to the HROs failure culture than less experienced.

H1- Experienced workers (of the company) work more according to the categories of HROs failure culture than less experienced.

3.1.4 Hypothesis 4

Moreover, since there appears to be a lack of theoretical knowledge regarding the role of gender within failure culture, an attempt to enrich this knowledge was made by creating also the following hypothesis.
H0 - Women and men work equal according to the categories of HROs failure culture than men.

H1 - Women work less according to the categories of HROs failure culture than men.

Similar to Hypothesis 2, first the F-value based on Levene’s Test must be observed, followed by the significance (p-value) of the t-test (Djurfeldt et al., 2010). To reject the null-hypothesis the p-value of the t-test must be under 0.05 (ibid).
4. Results and analysis

4.1 Rate of response and representativeness

The overall number of responses by the end of the survey time were 290 people. From these, 265 have filled in the survey completely (the rest did not finish the survey). The response rate is therefore 30% at the most and 27% at the lowest. Despite the fact that simple random sample was chosen, instead of an alternative stratified random sampling (Esaiasson, 2012), the representativeness in terms of division of male and female, manager and non-manager was relatively representative for the organisation. The gender division in the company is 54,7% male and 45,3% female, while participants in the survey were 51,03% male, 48,62% female and 0,34% other. The participation of the male gender is slightly overly representative. 8,6% of the workers in the company are managers, while 12,76% of the survey participants were managers, also leading to a slight overrepresentation.

The age distribution and the distribution of how long people have worked in the company are shown in the following histograms. The histograms on the left side show the distributions in the company, while the ones on the right side show the distribution of the survey participants. The average age of the participants with 43,7 years is comparable with the age of the full organisation of 44,5 years. The mean of the experience level (1-5) is also comparable between the sample with 3,25 and the organisation with 3,08 while the distribution is a little different. The overall representativeness is therefore given.

![Histograms of age and work experience](image)

*Figure 3, comparison of age and work experience from sample and population*
4.2 Hypotheses analysis

4.2.1 Hypothesis 1

Table 4, Result 1 Hypothesis 1 shows the descriptive statistics of the dependent variables, the categories, of failure culture.

<table>
<thead>
<tr>
<th></th>
<th>Safety feeling</th>
<th>Reputation</th>
<th>Mistake recognition and reporting</th>
<th>Appreciation and Feedback</th>
<th>Detailed approach to mistakes</th>
<th>Delegation (decision making)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Valid</td>
<td>270</td>
<td>275</td>
<td>269</td>
<td>275</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>20</td>
<td>15</td>
<td>21</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>Mean</td>
<td>2.3148</td>
<td>3.4509</td>
<td>2.4665</td>
<td>2.7836</td>
<td>2.5243</td>
<td>2.5303</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.51983</td>
<td>.73295</td>
<td>.78155</td>
<td>.81361</td>
<td>.75969</td>
<td>.73675</td>
</tr>
</tbody>
</table>

The second and third rows, N, show the amount of valid and missing answers of the survey, for each of the categories of Failure culture. The rows that are coming underneath show measures of central tendencies, the mean and the standard deviation.

Throughout all six categories, this descriptive analysis shows mean values over 2.3. The scale has 5 as a maximum score, which is the worst score for the categories of the Failure culture model, while 1 represents the best value. It can already be observed that the results are close to the middle value, 3. All of the variables, except Reputation, are closer to 1 than 5, which would mean that the culture is rather on the positive side than the negative. To statistically test the hypothesis, a One-Sample T Test was conducted (Dјurfeldt et.al., 2010). For this the test value 2 was specified, representing the “standard” with which the measured mean values needed to be compared. This means that 2 is the border-value for what represents a positive failure culture.

The test resulted in the following output in Table 5.
Table 5, Result 2 Hypothesis 1: One-Sample T Test

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety feeling</td>
<td>9.951</td>
<td>0.000</td>
<td>0.31481</td>
</tr>
<tr>
<td>Reputation</td>
<td>32.827</td>
<td>0.000</td>
<td>1.45091</td>
</tr>
<tr>
<td>Mistake recognition and reporting</td>
<td>9.791</td>
<td>0.000</td>
<td>0.46654</td>
</tr>
<tr>
<td>Appreciation and Feedback</td>
<td>15.972</td>
<td>0.000</td>
<td>0.78364</td>
</tr>
<tr>
<td>Detailed approach to mistakes</td>
<td>11.297</td>
<td>0.000</td>
<td>0.52425</td>
</tr>
<tr>
<td>Delegation (decision making)</td>
<td>11.695</td>
<td>0.000</td>
<td>0.5303</td>
</tr>
</tbody>
</table>

The important values to observe here are under column three, Sig. (2-tailed) (Djurfeldt et al., 2010). The p-value, being lower than 0.05 shows that the differences between the mean values for all categories of failure culture observed above and the decided limit of 2 are significantly different. This means that the null-hypothesis is rejected. This means that the investigated company’s failure culture does not reach the limit of what was set for a positive failure culture.

4.2.2 Hypothesis 2

Table 6, The measures of central tendency of Failure culture categories

<table>
<thead>
<tr>
<th>Index</th>
<th>Job Title</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety feeling</td>
<td>Manager</td>
<td>34</td>
<td>2.0756</td>
<td>0.44657</td>
</tr>
<tr>
<td></td>
<td>Worker</td>
<td>236</td>
<td>2.3493</td>
<td>0.52139</td>
</tr>
<tr>
<td>Reputation</td>
<td>Manager</td>
<td>35</td>
<td>3.3810</td>
<td>0.74598</td>
</tr>
<tr>
<td></td>
<td>Worker</td>
<td>240</td>
<td>3.4611</td>
<td>0.73206</td>
</tr>
<tr>
<td>Mistake recognition and reporting</td>
<td>Manager</td>
<td>34</td>
<td>2.2206</td>
<td>0.69839</td>
</tr>
<tr>
<td></td>
<td>Worker</td>
<td>235</td>
<td>2.5021</td>
<td>0.78786</td>
</tr>
<tr>
<td>Appreciation and Feedback</td>
<td>Manager</td>
<td>35</td>
<td>2.7000</td>
<td>0.85061</td>
</tr>
<tr>
<td></td>
<td>Worker</td>
<td>240</td>
<td>2.7958</td>
<td>0.80919</td>
</tr>
<tr>
<td>Detailed approach to mistakes</td>
<td>Manager</td>
<td>34</td>
<td>2.2647</td>
<td>0.70962</td>
</tr>
<tr>
<td></td>
<td>Worker</td>
<td>234</td>
<td>2.5620</td>
<td>0.76076</td>
</tr>
<tr>
<td>Delegation (decision making)</td>
<td>Manager</td>
<td>34</td>
<td>2.0882</td>
<td>0.58349</td>
</tr>
<tr>
<td></td>
<td>Worker</td>
<td>230</td>
<td>2.5957</td>
<td>0.73555</td>
</tr>
</tbody>
</table>

The table above represents the averages of the Failure culture categories for managers and workers. At first glance it looks like there are differences in all the categories between both groups, where the mean values for managers are closer to 1. However, it is also evident under
column five, that the differences of standard deviation between the groups manager and worker are bigger for some of the categories than other. This means that the variance around the mean-values differ more or less.

The hypothesis was therefore also tested with help of t-tests for Equality of Means (see Table 7). First, Levene’s test for Equality of Variances was conducted, shown in the F-value. As mentioned above, the test measures whether the variances in the two groups are similar or not (Djurfeldt et.al., 2010). If the F-value is significant (with a p-value of ≤ 0.05, under the column Sig.), the null-hypothesis, which states that the variance between the two groups are similar, must be rejected. Depending on the result, the value on the row “Equal variances assumed” or “Equal variances not assumed” must be observed in the rest of the table. Only the p-value to Delegation (decision making) was under 0.05 in the Levene’s Test, and therefore showed a significant F-value. This means that the values on the row “Equal variances not assumed” must be read for the category Delegation (decision making). As a result, the following table was summarised with a part of the output from SPSS.

Table 7, Significance test Hypothesis 2

<table>
<thead>
<tr>
<th>Index</th>
<th>Levene's Test for Equality of Variances - F</th>
<th>Sig.</th>
<th>t-test for Equality of Means- t</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety feeling</td>
<td>Equal variances assumed</td>
<td>0.407</td>
<td>0.524</td>
<td>-2.909</td>
<td>0.004</td>
</tr>
<tr>
<td>Reputation</td>
<td>Equal variances assumed</td>
<td>0.670</td>
<td>0.414</td>
<td>-0.604</td>
<td>0.547</td>
</tr>
<tr>
<td>Mistake recognition and reporting</td>
<td>Equal variances assumed</td>
<td>0.658</td>
<td>0.418</td>
<td>-1.974</td>
<td>0.049</td>
</tr>
<tr>
<td>Appreciation and Feedback</td>
<td>Equal variances assumed</td>
<td>0.038</td>
<td>0.846</td>
<td>-0.65</td>
<td>0.516</td>
</tr>
<tr>
<td>Detailed approach to mistakes</td>
<td>Equal variances assumed</td>
<td>0.859</td>
<td>0.355</td>
<td>-2.146</td>
<td>0.033</td>
</tr>
<tr>
<td>Delegation (decision making)</td>
<td>Equal variances not assumed</td>
<td>6.184</td>
<td>0.014</td>
<td>-4.563</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The t-test could now be interpreted. Again, to see if the t-value is significant, the p-value must be observed, this time under the column Sig. (2-tailed). A p-value bigger than 0.05, under Sig. (2-tailed), means that the null-hypothesis must be confirmed, which means that there is no difference between the two groups.

The observations will be conducted from the upper row to the lowest. The p-value for Safety feeling in the first row is 0.004 and with this under 0.05. This rejects the null-hypothesis for this index. The shown mean of 2.07 of the managers for Safety feeling in the Table 6 is therefore significantly smaller than the one from the workers. Managers are feeling safer reporting failures.
The p-values belonging to the row Reputation, is 0.547, and with this over 0.05. This means that the null-hypothesis was proven to be correct. In other words, there is no significant difference between the mean values of the managers and workers.

The p-values for Mistake recognition and reporting, have a p-value of 0.049 and by that is smaller than 0.05. The null-hypothesis is thereby rejected, meaning that there is a significant difference in the mean values between the two investigated groups. According to this value managers are keener to recognise and report failures.

Appreciation and Feedback - has no significant difference between the mean values of the workers and managers, since the p-value of 0.516 is bigger than 0.05 and the null-hypothesis is proven to be correct.

The last two rows, Detailed approach to mistakes and Delegation (decision making) are all having a p-value under 0.05, with their 0.033 and 0.000. Both null-hypothesis are therefore rejected, meaning that there are significance differences between the means of the managers and the workers. That means that the managers feel more involved in detailed approach to mistakes and get more decision makings delegated to them.

This means that Table 7 shows that in only four out of the six categories (indexes) it can be said that there is a significant difference between managers and workers. The two differences that cannot be significantly confirmed are for the category Reputation and Appreciation and Feedback - there is no significant difference between the workers and managers for these categories in other words.

Visually the differences can be presented in the following histograms:
Since no significant difference could be proved for the categories *Reputation* and *Appreciation and Feedback*, these results are being presented without splitting them into the groups managers and workers.

### 4.2.3 Hypothesis 3

Hypothesis 3 was tested by using Pearson’s correlation test. The correlation is shown as a number between -1 and 1 and indicates the extent in which the variables are linearly correlated. The relevant outputs are summarised in the table underneath. The results show that there are no significant correlations between the dependent and independent variables that have been tested. This can be seen in the very low correlation values (R²-values) and the significance value, showing that there is no significance. This means that the null-hypothesis of hypothesis 3 is proved to be correct.

**Table 8. Pearson’s correlation test on Failure culture categories**

<table>
<thead>
<tr>
<th>Index</th>
<th>R²</th>
<th>ANOVA Significance</th>
<th>H0 (no significance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety Feeling</td>
<td>0.006</td>
<td>0.197</td>
<td>confirmed</td>
</tr>
<tr>
<td>Reputation</td>
<td>0.001</td>
<td>0.634</td>
<td>confirmed</td>
</tr>
<tr>
<td>Mistakes recognition and Reporting</td>
<td>0.007</td>
<td>0.175</td>
<td>confirmed</td>
</tr>
<tr>
<td>Appreciation and Feedback</td>
<td>0.002</td>
<td>0.488</td>
<td>confirmed</td>
</tr>
<tr>
<td>Detailed approach to mistakes</td>
<td>0.003</td>
<td>0.335</td>
<td>confirmed</td>
</tr>
<tr>
<td>Delegation</td>
<td>0.002</td>
<td>0.59</td>
<td>confirmed</td>
</tr>
</tbody>
</table>

To visualise the lack of correlation, the following box plot diagram was created using one of the dependent variables, Safety feeling, as an example.
The boxplot shows clearly that there is no particular correlation between experience in the organisation and the survey values.

![Simple Boxplot of Safety feeling by How long do you work in "name of company"?](image)

_Figure 6, Simple Boxplot showing the answers distribution for the category Safety feeling._

4.2.4 Hypothesis 4

Table 9, Measures of central tendencies divided between genders

<table>
<thead>
<tr>
<th>Index</th>
<th>Gender</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety feeling</td>
<td>Male</td>
<td>136</td>
<td>2,271</td>
<td>0,49952</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>133</td>
<td>2,3588</td>
<td>0,53988</td>
</tr>
<tr>
<td>Reputation</td>
<td>Male</td>
<td>139</td>
<td>3,4341</td>
<td>0,69927</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>135</td>
<td>3,4667</td>
<td>0,77073</td>
</tr>
<tr>
<td>Mistake recognition and reporting</td>
<td>Male</td>
<td>136</td>
<td>2,4485</td>
<td>0,69995</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>132</td>
<td>2,4811</td>
<td>0,86140</td>
</tr>
<tr>
<td>Appreciation and Feedback</td>
<td>Male</td>
<td>139</td>
<td>2,6906</td>
<td>0,81070</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>135</td>
<td>2,8778</td>
<td>0,81145</td>
</tr>
<tr>
<td>Detailed approach to mistakes</td>
<td>Male</td>
<td>134</td>
<td>2,4142</td>
<td>0,74975</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>133</td>
<td>2,6353</td>
<td>0,75911</td>
</tr>
<tr>
<td>Delegation (decision making)</td>
<td>Male</td>
<td>134</td>
<td>2,4254</td>
<td>0,73451</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>129</td>
<td>2,6395</td>
<td>0,72881</td>
</tr>
</tbody>
</table>

At first glance the table shows that there are differences between the means of the two groups. However, the table below, investigates the significance of these differences with a t-test. Similar to the testing of the Hypothesis 2, also here first Levene’s test for Equality of Variances was conducted, shown in the F-value, to see if the variances in the two groups are
similar or not (Djurfeldt et.al., 2010). The p-value under the column Sig. is over 0,05 for all the categories, except Mistake recognition and reporting. The null-hypothesis, which states that the variance between the two groups are similar, must be rejected, and the value must be read on the row “Equal variances not assumed”. As a result, the following table was summarised with the relevant parts taken from the output from SPSS.

Table 10. Significance test Hypothesis 4

<table>
<thead>
<tr>
<th>Index</th>
<th>Levene's Test for Equality of Variances - F</th>
<th>Sig.</th>
<th>t-test for Equality of Means - t</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety feeling</td>
<td>Equal variances assumed</td>
<td>1,242</td>
<td>0,266</td>
<td>-1,384</td>
<td>0,167</td>
</tr>
<tr>
<td>Reputation</td>
<td>Equal variances assumed</td>
<td>2,002</td>
<td>0,158</td>
<td>-0,367</td>
<td>0,714</td>
</tr>
<tr>
<td>Mistake recognition and reporting</td>
<td>Equal variances not assumed</td>
<td>5,827</td>
<td>0,016</td>
<td>-0,339</td>
<td>0,735</td>
</tr>
<tr>
<td>Appreciation and Feedback</td>
<td>Equal variances assumed</td>
<td>0,260</td>
<td>0,611</td>
<td>-1,909</td>
<td>0,057</td>
</tr>
<tr>
<td>Detailed approach to mistakes</td>
<td>Equal variances assumed</td>
<td>0,484</td>
<td>0,487</td>
<td>-2,395</td>
<td>0,017</td>
</tr>
<tr>
<td>Delegation (decision making)</td>
<td>Equal variances assumed</td>
<td>0,484</td>
<td>0,487</td>
<td>-2,373</td>
<td>0,018</td>
</tr>
</tbody>
</table>

Now the t-test could be interpreted. To see if the t-value is significant, the p-value must be observed under the column Sig. (2-tailed). A p-value bigger than 0,05, under Sig. (2-tailed), means that the null-hypothesis must be confirmed, which means that there is no difference between the two groups.

The observations will be presented from the upper row to the lowest. The p-value for Safety feeling in the first row is 0,167 and with this over 0,05. The null-hypothesis for this index is therefore proven to be correct. The difference shown in the mean values between males and females is therefore not significantly proved. This means that it cannot be stated that males are feeling safer reporting failures than females.

The p-values belonging to the row Reputation, is 0,714, which also is bigger than 0,05. This means that the null-hypothesis was proved to be correct. In other words, there is no significant difference between the mean values of males and females.

Mistake recognition and reporting has a p-value of 0,735. Being bigger than 0,05 this means that the null-hypothesis for this category is thereby confirmed, meaning that there is no significant difference in the mean values between the males and female.

Appreciation and Feedback - has no significant difference between the mean values of the males and females, since the p-value of 0,057 is bigger than 0,05, and the null-hypothesis is thereby proved to be correct.
The last two rows, *Detailed approach to mistakes* and *Delegation (decision making)* are the only categories having a p-value under 0.05, with their 0.017 and 0.018. Both null-hypothesis are therefore rejected, meaning that there are differences that are significant between the mean values of the males and females. In these cases, men feel more that they are participating in detailed discussions of occurred mistakes, and that decision makings are delegated to men more than to women.

To visualise these, the following histograms present the differences.

*Table 11, Comparison between genders for Detailed approach to mistakes and Delegation*
5. Conclusions

The results of the hypotheses tests are summarised in the following table.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Proven to be correct for:</th>
</tr>
</thead>
<tbody>
<tr>
<td>H0 - The mean values of the categories of HRO’s failure culture in the investigated company are equal to or bigger than 2.</td>
<td>Safety feeling, Reputation, Mistake recognition and reporting, Appreciation and Feedback, Detailed approach to mistakes, Delegation (decision making)</td>
</tr>
<tr>
<td>H1 - The mean values of the categories of HRO’s failure culture categories in the investigated company are smaller than 2.</td>
<td></td>
</tr>
<tr>
<td>H0 - Managers and workers work equal according to the categories of HRO’s failure culture.</td>
<td>Reputation, Appreciation and Feedback</td>
</tr>
<tr>
<td>H1 - Managers work more according to the categories of HRO’s failure culture than workers.</td>
<td>Safety feeling, Mistake recognition and reporting, Detailed approach to mistakes, Delegation (decision making)</td>
</tr>
<tr>
<td>H0 - Experienced and less experienced workers (of the company) work equal according to the categories of HRO’s failure culture.</td>
<td>Safety feeling, Reputation, Mistake recognition and reporting, Appreciation and Feedback, Detailed approach to mistakes, Delegation (decision making)</td>
</tr>
<tr>
<td>H1- Experienced workers (of the company) work more according to the categories of HRO’s failure culture than less experienced.</td>
<td></td>
</tr>
<tr>
<td>H0 - Women and men work equal according to the categories of HRO’s failure culture than men.</td>
<td>Safety feeling, Reputation, Mistake recognition and reporting, Appreciation and Feedback,</td>
</tr>
<tr>
<td>H1 - Women work less according to the categories of HRO’s failure culture than men.</td>
<td>Detailed approach to mistakes, Delegation (decision making)</td>
</tr>
</tbody>
</table>
The following conclusions can be made:

1. Hypothesis 1 helped to measure the failure culture in the investigated organisation after HRO standards, and thereby fulfilled the first part of the purpose of the current thesis. The hypothesis for all categories of the Failure culture model was rejected. This means that none of the categories reached the decided border-value of 2 for a positive HRO failure culture, which was set as a limit to call the failure culture positive, on a scale from 1- best score, to 5 - worst score. The biggest difference is detected for the variable Reputation, while the lowest is for Safety feeling, which is an interesting mismatch. While the workers feel safer asking questions, reporting failures etc., they are more concerned about their reputation when reporting the failures. The rest of the categories vary. In summary this means that an organisation which works with continuous improvements, does not necessarily meet the standards of a failure culture in an HRO, which was also stated by Weick and Sutcliffe. Nevertheless, some categories match the failure culture of HROs more closely than others. According to Yukl (2013) the values and beliefs of an organisation need to match to their strategies. Therefore, the full model Failure culture in HROs needs to match to the strategies of the organisation, because the organisation works with continuous improvements, and thereby with elimination of deviations.

When viewing the results in light of the theory of continuous improvement and the Cornerstone Model from Bergman and Klefsjö (2010) shown in Figure 1, important conclusions can be drawn. The cornerstone “Improve continuously”, as well as Sörvits description of continuous improvement, are directly affected by the failure culture of the organisation. By reporting failures, the organisation has the chance to analyse them in order to eliminate deviations from workflows and processes. Moreover, this allows the possibility for smaller breakthroughs that are essential for continuous improvement (Bergman & Klefsjö, 2010; Sörvist, 2004). All categories of the model Failure culture in HRO’s are therefore relevant to the achievement of good results for continuous improvement. For all categories except one, Reputation, the values were above 2 but below 3. Reputation is 3,45 and therefore lies on the negative side of the scale. It is questionable whether continuous improvement is fully working when reputation is negatively affected by reporting mistakes.

Open communication of failures and thereby a positive failure culture is also a precondition for the cornerstone “Let everybody be committed”. Decision making and overall participation in improvement work is more easily reached when employees recognise and report mistakes and are encouraged to openly communicate and discuss them. Moreover, Delegation (decision making) was a category of the failure culture of HRO’s, being an essential part of it as well as being essential for the cornerstone. This was measured by the questions which ask how often employees are consulted before decision making and to what extent the decision maker follows recommendations of the employees. It is essential for the cornerstone since it states
that the involved employees and not the project owners are the ones seeing potential for improvements. The result for this category was third worst. Nevertheless, with a value of 2.53 the category has positive tendencies, since it is on the positive side of the scale.

“Base decisions on facts” also benefits from a positive failure culture of HROs, since it is important to break down information about deviations to their root causes. This is reached by Detailed approach to mistakes. “Focus on customers” together with “Focus on processes” benefit from positive failure culture of HROs because the reported and discussed mistakes can be removed from the processes to make them more efficient, making it easier to attain customer satisfaction and lower costs. Lastly, the cornerstone in the base, “Committed Leadership”, is realized through a positive failure culture when the leaders encourage and appreciate the communication of mistakes (Bergman & Klefsjö, 2004). The leadership’s role was also highlighted in Rother’s (2013) interpretation of Toyotas quality work, stating that leaders should act like mentors. The category Appreciation and Feedback has the second worst result, with its 2.78 value. An improvement in this area would therefore benefit the quality work of the organization. In general, improvement in all categories of failure culture model will have a positive effect on quality work of the organization.

2. The results of Hypothesis 2 show that there are significant differences found between the values of managers and workers in four out of six categories. These are the Safety feeling, Mistake recognition and reporting, Detailed approach to mistakes and Delegation (decision making). The biggest difference is detected for the categories Delegation (decision making), with the mean difference of -0.50742, which is the difference between the means 2.0882 of the managers and 2.5957 of the workers, seen in Table 6. Safety feeling, Mistake recognition and reporting and Detailed approach to mistakes all have around an -0.27 to -0.30 difference in the mean values. The conclusion is that the managers work more in accordance with the failure culture of HROs than the workers in the company. In addition to the different jobs and tasks managers and workers have in an organisation and dealing with failures there is a difference in the culture towards failure. Rother (2013) states that leaders should act as mentors for workers. On the one hand it could therefore be seen as positive that managers have gotten better values. However, it could also be concluded that managers do not fully transfer the same circumstances to workers. Leadership is perceived as important for creation of culture in the organisation according to Cardwell (2018) as well as Weick and Sutcliffe (2007). The discrepancies between workers and managers show that prerequisites of accountability for safety and transparency in the organisation can be improved, in order for everyone to live a positive failure culture.

3. The results of Hypothesis 3 have shown that there are no significant differences between experienced and less experienced employees of the company, and no correlations could be found. In particular the missing correlation shown between work
experience and *Delegation (decision making)* is a potential for improvement, since migrating decision making to people with expertise is an important factor for HROs (Weick & Sutcliffe, 2007). However, it also demonstrates that belonging to the company for longer has no effect on the failure culture of the individual.

4. The test-results of Hypothesis 4 showed that there are significant differences between the values of males and females in two categories: *Detailed approach to mistakes* and *Delegation (decision making)*. The differences in the means of these two categories are -0,22 and -0,21, where males have the smaller values (closer to 1). This means that females feel less that decision makings are delegated to them and are also less involved in detailed discussions of mistakes. It is interesting to point out that both categories are also those with the biggest differences in the means under Hypothesis 2, where significant differences were detected between workers and managers.

Among the respondents of the survey, it was shown that only one of the managers was female. This is a clear improvement possibility, since the failure culture is lived by all workers of the company.

Based on the results of the statistical analysis, some suggestions for improvements of the failure culture in the investigated organisation can be made, reaching the second part of the aims of this study.

Improvements in the organisation can be detected and processes developed by involving the workers more – this can be done by delegating decision makings more. There were also differences detected between workers and managers. Through practicing coaching kata the managers can show and transfer their ways of work to the workers.

Reporting mistakes could be done more systematically using checklists explicitly designed for this and the discussion of mistakes could be part of weekly meetings.

Differences between gender can be eliminated through different empowerment programs. Further research could also analyse why there are differences.

A starting point for improvement in failure culture could be the topic of reputation, having the worst score of all categories. This attitude can be changed when the employees know the system and its weaknesses and the consequences of failures, leading to reports of failures without bias.

By actively choosing to act in particular ways, and practicing new behavioural patterns, the organisation can reach a wished failure culture “by design”, mentioned by Elmgart and Bloch (2017).
6. Discussion

The current study was conducted using a deductive approach and a quantitative analysis. The advantage of this approach was the possibility to explain correlations between concrete variables, which favoured validity of the work. Creating indexes helps to draw an interpretable picture. The statistical measurements were allowed conclusions to be drawn which could be generalised to the whole organisation. The drawback of using statistical measurements which required indexes is the loss of some details. For example, by creating the “SafetyFeeling” variable, seven questions were added together to form one variable. As a result, the details of silo thinking were lost. The indexes were created in order to keep the thesis to a reasonable size. The overall picture is not altered for the items themselves in each category. This research could therefore be improved by creating indexes that catch more details of failure culture which could generate more detailed improvement suggestions.

The reliability of the work is high due to the conducted T-test with the 5% significance value. Random errors have with a 5% risk influenced the results of the research. The Cronbach’s Alpha shows a high consistency within the created indexes.

It is also important to note that the measured data can be biased. As stated in the results, a slight discrepancy was detected in the age and experience distribution between the surveyed people and the full number of employees in the company. This could mean that the overrepresentation of some groups could lead to biased results. However, the benefit of using statistical analyses was that despite the lower response rate (27-30%), it was still possible to generalise the results to the full organisation through significance tests.

Also the conclusions of the study can be discussed further. The conclusions to Hypothesis 1 could mean that the reporting and discussion of failures is encouraged by leaders to more extent, shown in better values in Safety feelings, than it is accepted as a value in the organisation, shown in the value of Reputation. The second worst result was for the category Appreciation and feedback. This is an interesting observation, given that the results for Safety feeling were better than for this category. The employees feel safer and more encouraged to report and discuss failures, but the score for appreciation is worse. That the Safety feeling is better than the others could be a result of employment contracts and the workers’ union that supports the employees.

The overall picture shows a not yet fully developed failure culture according to HROs. It is however important to mention that the investigated organisation is not an HRO, and mean values under 3 in the tests are relatively positive. Since the organisation is certified by ISO 9001:2015, it is in the organisation’s interest to maintain continuous improvement. The general conclusion is therefore that the topic of failure culture could be prioritised more in the organisation. The detection and solution of problems are highlighted as an important part in organisational improvement work and TQM. The solutions to smaller problems, together with the sum, lead to the biggest number of improvements in the organisations, as stated in the theory (Sörqvist, 2004). This is done by detecting and finding solutions to deviations and
working them through in small PDCA-cycles, which can get more effective by a positive failure culture (Rother, 2013).

Hypothesis 2 - 4 generate and offer more detailed views for improvements. The conclusion of the analysis of Hypothesis 2 is that the company can work on delegating decision making, encouraging the reporting of mistakes and involving the workers more in discussions of problems. Safety feeling, Mistake recognition and reporting and Detailed approach to mistakes have smaller differences between the mean values than the one for Delegation (decision making). However, it provides the important information that the workers in the company are less keen to report failures, feel less safe doing so and feel less involved in discussions around occurred mistakes. It is not surprising that the managers in the organisation feel more that decision makings are delegated to them. However, by not reaching out to the workers, the leaders can miss valuable advice. Also, it could be an indicator that people with advanced expertise hold the management roles in the organisation, which would match the traditional management system. By involving the workers more, improvements in the organisation can be detected and processes developed. The reporting of mistakes could be done by using checklists explicitly designed for this and the discussion of mistakes could be part of weekly meetings, as suggested above. By using these kinds of practices, the organisation can act their way into what they become, which was stressed by Weick and Sutcliffe (2007) and Rother (2013). By using these tools, the norms in the company will be affected, indirectly affecting the other categories of failure culture, such as the Safety feeling and Delegation (decision making). Although the managers and workers have different functions and ways of working, the culture needs to be homogeneous in the organisation. Future research could be conducted to investigate if the managers of the company are interested in improving the failure culture of the company.

Hypothesis 3 showed that belonging to the company for longer has no effect on the failure culture of the individual. This means that people who work for longer in the organisation do not have a better average than the rest. On one hand this could imply that people are quickly adapting to the failure culture of the company, which would be a good result. On the other hand, it also offers the possibility that there is no difference in the failure culture within the company compared to the previous occupation/education of the employees. Working longer in the company results in neither a positive nor negative impact on employees’ failure culture. The cause for this cannot be identified with the collected data and offers the opportunity for additional research.

For Hypothesis 4, where females have gotten worse values than men for the categories Delegation (decision making) and Detailed approach to mistakes, there are possibilities for improvements through for example empowerment programs, as suggested above. This could be connected to the very traditional role model between men and women in Western Germany for a long time. In combination with the relatively high average age in the company, this could have an impact on the results. Further research could also analyse in more depth the differences between the genders within the company. The current analysis shows that it is worth going into more detail.
7. List of sources


Appendix 1. Information text posted on the intranet and communication platform of the company

Liebe Kolleg*innen, ich (Anna Bayramyan, Masterandin im Bereich OI) schreibe aktuell meine Masterarbeit beim *name of the company* zum Thema Fehlermanagement als Teil der Qualitätsentwicklung. Dafür benötige ich Ihre Hilfe. Zusammen wollen wir untersuchen, wie wir im *name of the company* mit Fehlern umgehen – betreut werde ich hierbei vom zentralen Qualitätsmanagement.

Die Umfrage richtet sich an Mitarbeiter*innen des *headquarter of name of the company*. Ich wäre Ihnen sehr dankbar, wenn Sie sich einen kleinen Moment nehmen würden, um daran teilzunehmen. Selbstverständlich ist die Umfrage anonym und die Daten werden nicht weitergegeben. Zum Ausfüllen der Umfrage benötigen Sie ca. fünf Minuten.

*Hier geht es zur netigate Umfrage*

Vielen Dank, dass Sie mich bei meiner Masterarbeit unterstützen.

Dear colleagues, I (Anna Bayramyan, Master student in the department OI) write in the moment my master thesis in *name of the company*. The topic is Failure management as part of quality development. For that I need your help. Together we want to investigate, how we in *name of the company* handle failures - I am supervised by the central quality management. The questionnaire is aimed for the employees of the headquarter of *name of company*. I would be very grateful if you would take a little time to participate. Of course, the survey is anonymous and the data will not be passed forward. It takes approximately five minutes to complete the survey.

* Click here for the netigate survey *

Thank you for supporting me in my master thesis.
Appendix 2. Introduction text for the survey

Liebe/r Teilnehmer/in
dies ist eine Umfrage von mir - Anna Bayramyan, Studentin im Fach Qualitäts- und Führungsentwicklung in Mid Sweden University, Schweden.
Mit dem folgenden Fragebogen sollen im Rahmen meiner Masterarbeit Fragen zu Fehlermanagement als Teil der Qualitätsentwicklung untersucht werden. Die Bearbeitung des Fragebogens nimmt ungefähr fünf Minuten in Anspruch. Der Fragebogen ist Anonym und die Daten werden nicht an Dritte weitergegeben. Herzlichen Dank für Ihre Teilnahme!

Dear participant
dthis is a survey of mine - Anna Bayramyan, student in Quality and Leadership development at Mid Sweden University, Sweden. The following questionnaire intends to answer questions related to the topic of my master thesis, failure management as part of quality development. It takes approximately five minutes to complete the questionnaire. The questionnaire is anonymous and the data will not be passed on to third parties. Thank you very much for your participation!
Appendix 3. The survey questions

1 General questions:
You are: male, female, other
Your age:
How long have you worked in the company?: 0-2 years, 2-5 years, 5-10 years, 10-20 years, more than 20 years
Are you a manager?: Yes, No

2 Failure Culture:
(Please choose one of the options)
1. To what extent are you aware of the contribution your work makes to the *name of company*?
   Very aware
   Strongly aware
   Medium
   Little
   Not aware at all

2. To what extent is the identification of errors / problems part of your job?
   Very strong
   Strong
   Medium
   Little
   Not at all

3. To what extent do you feel safe to report errors …
   a) that were caused by you
   b) that were caused by others?
   Completely
   Very
   Somewhat
   Barely
   Not at all

4. To what extent do you feel encouraged to ask questions about your workflow?
   Completely
   Very
   Somewhat
   Barely
   Not at all

5. To what extent do you feel encouraged to discuss your errors(/mistakes?) with your colleagues
   a) within the department?
   b) also across departmental boundaries?
   Completely
   Very
   Somewhat
   Barely
   Not at all
6. To what extent do you think it is appreciated that you communicate mistakes?

Completely
Very
Somewhat
 Barely
Not at all

7. How often do you get constructive feedback when you communicate a mistake?

Always
 Often
Sometimes
Rare
Never

8. To what extent do you think that your reputation will be positively influenced if you communicate errors.
   a) caused by yourself
   b) caused by a close colleague
   c) caused by colleagues from other departments?

Completely
Very
Somewhat
 Barely
Not at all

9. How...
   a)...often are occurred errors being discussed in the team?
   b)... detailed are occurred errors being discussed in the team?

always/ very strong
often / strong
sometimes/medium
rare / little
never/ not at all

10. To what extent do you feel that
    a) ...you are being made personally responsible for your mistakes?
    b) ...that your mistakes are being systematically questioned?

Very strong
Strong
Medium
Little
Not at all

11. To what extent do you pay attention to system-related errors in the company?

Very much
 Much
Medium
Little
Not at all
12. How often
   a) … do you communicate system-related errors?
   b) ... are you being encouraged from your managers to communicate them?

   Always
   Often
   Rarely
   Little
   Not at all

13. Are you well prepared to manage a mistake?

   Fully applies
   Largely applies
   Partially applies
   Does rather not apply
   Does not apply at all

14. To what extent do you in your company have a structured procedure for eliminating errors?

   Very much structured
   Much structured
   Medium
   Little
   Not structured at all

15. To what extent are the work methods being adapted after an occurred failure, in order to avoid it in the future?

   Very strong
   Strong
   Middle
   Little
   Not at all

16. How much expertise do you have, to be able making relevant decisions within you work area?

   Very much
   Much
   Medium
   Little
   None at all

17. How often are you (as the expert) being consulted before making decisions?

   Always
   Often
   Sometimes
   Rare
   Never

18. To what extent does the decision maker follow your recommendations when making decisions?

   Follows fully
   Rather follows
   Follows partially
   Rather not follow
   Does not follow at all
Appendix 4. The tree diagrams conducted for each of the five HRO principles.
Commitment to Resiliency

Readiness to manage unexpected events

Are you well prepared to manage a mistake?

To what extent do you in your company have a structured procedure for eliminating errors?

Readiness to manage the correction

To what extent are the work methods being adapted after an occurred failure, in order to avoid it in the future?

Deference to Expertise

Decision-making migrated to people with the right knowledge

How much expertise do you have, to be able making decisions within you work area?

How often are you (as the expert) being consulted before making decisions?

To what extent does the decision maker follow your recommendations when making decisions?