University of Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program/Specialization: Risk Management	Spring semester, 2021 **Open** / Restricted access
Writer: Bruna Atamanczuk	…… (Writer's signature)
Faculty supervisor: Terje Aven External supervisor(s): Bernt Eriksen	
Thesis title: A holistic approach to developing safety-related systems in compliance to the IEC61508	
Credits (ECTS): 30 ETC	
Key words: IEC61508 SIL Design Risk acceptance criteria QRA	Pages: …58…… + enclosure: …4…… Stavanger, …15/06/2021.. Date/year

ABSTRACT

In this report, several aspects related to the design of safety instrumented systems (SIS) have been presented and discussed. Some challenges related to the methods used to demonstrate functional safety were highlighted, especially when adopting the risk-based approach proposed by the IEC61508 (2010). Two issues were identified when using the standard approach for designing purposes. Firstly, the methods used do not highlight features of risk such as uncertainty and knowledge. This is particularly alarming given that risk acceptance criteria (RAC) and quantitative risk assessments (QRAs) are used to set the necessary performance requirements of the system and further referred throughout the process to verify if the requirements are met. If the risk is misrepresented, the final product can be a system that is eighter too complex or too simple to provide the necessary risk reduction. Secondly, the standard lacks a detailed approach to follow from a product development perspective. Based on these issues, this work presents a discussion that tries to shed some light on how to overcome them during the design of SIS. Furthermore, depending on the technology adopted the standard approach might not be suitable to perform safety demonstration. Hence, new safety philosophies might be needed to support the development of systems that violate the traditional requirements.

***Key words*:** IEC61508, *SIL, design, uncertainty, knowledge, QRA, RAC*

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Prof. Terje Aven for his continuous support, patience and invaluable advice throughout this project. I would also like to thank Bernt Eriksen and all SPEO Products' team for their technical support on my study. Finally, I would like to express my gratitude to my family, friends, and colleagues for the unconditional support during this very intense academic year.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS

ALARP	As Low as Reasonably Practicable
CCF	Common Cause Failure
DD	Dangerous Detected (failure)
DHSV	Down Hole Safety Valve
DU	Dangerous Undetected (failure)
E/E/PE	Electrical/Electronic/Programmable Electronic
ESD	Emergency Shut Down
EUC	Equipment Under Control
FMEA	Failure Modes and Effects Analysis
FMEDA	Failure Modes, Effects and Diagnosis Analysis
FTA	Fault Tree Analysis
HAZID	Hazard Identification
HAZOP	Hazard and Operability
HFT	Hardware fault tolerance
LOPA	Layers of Protection Analysis
MTTR	Mean Time to Repair
NCS	Norwegian Continental Shelf
NPD	New Product Development
PFD	Probability of Failure on Demand
PFH	Probability of Failure per Hour
PHA	Process Hazard Analysis
PMV	Process Master Valve
PWV	Process Wing Valve
QRA	Quantitative Risk Assessment

RAC	Risk Acceptance Criteria
RBD	Reliability Block Diagram
SC	Systematic Capability
SD	Safe Detected (failure)
SFF	Safe Failure Fraction
SIF	Safety Instrumented Function
SIL	Safety Integrity Level
SIS	Safety Instrumented System
SRS	Safety Requirements Specification
SU	Safe Undetected (failure)

1. INTRODUCTION

1.1. Background

New product development (NPD) can be motivated by different factors such as: changes in market or technology, increased competition or even changes in regulation. There are many ways of classifying new products and the definitions will vary depending on whose perspective is adopted (producer or customer) or in relation to what is new (Murthy et al., 2008). For instance, in relation to the producer's/customer's perspective, a product can be new to the world, to the industry or new to the customer. If a product is classified with respect to what is new, it is said to be a new technology, or new design, new use, etc. New products can be also classified with respect to the changes they bring. These can be minor or major depending on whose perspective is adopted. For example, some changes (e.g. changes that reduce the cost of producing the final product) will be major from the producer's perspective and will not be perceived from the customer's perspective. Changes can also be denoted as incremental or radical. Incremental changes are related to advances in existing technology, while radical changes are due to new technology (Murthy et al., 2008).

The process of creating a new product can be long and complicated, depending on the complexity of the product. To create a successful product, producers need to find balance between the customers' expectations and their own. Moreover, for products intended to be used in harsh environments, such as the one in the Norwegian Continental Shelf (NCS), there is an extensive focus on safety and reducing the risks associated with these activities. In addition, as the industry matured over the years, new products are usually intended to reduce operational cost while providing safety. Furthermore, depending on the application of the product, producers need to comply with several requirements, such as regulatory, safety and technical requirements to name a few (Lundteigen et al., 2009).

In order to reduce the risks associated with the oil and gas industry, safety barriers are implemented. Safety barriers can be defined as physical and/or non-physical and their main purpose is to prevent, control, or mitigate undesired events or accidents. In an analytical

context however, it is common to refer to barrier function, systems, or elements. According to Johansen and Rausand (2015):

- A *barrier function* is designed to prevent (or mitigate) the consequences of a specific hazardous event. It determines the role of the barrier, and can be specified by an action, e.g. “stop flow”. Sometimes, it is convenient to break down a barrier function into sub-functions. That is, “smaller” functions necessary to perform the main function, such as to detect, verify, and relieve high pressure.
- A *barrier system* is designed to perform one or more barrier functions and it can be constituted by several *barrier elements*. These are classified as *technical*, *operational*, and *organizational* barrier elements. According to Liu (2020):

    - Technical barriers elements are engineered systems that perform one or more barrier function.
    - Operational barrier elements are tasks carried out to perform a barrier function
    - Organizational barrier elements are the personnel responsible for carrying the activities out.

In general, there are multiple technical barriers used in the oil and gas industry. These barriers form different protection layers. Among all the systems used, the Safety Instrumented System (SIS) is the most important and critical protection layer (Chang et al., 2015). Safety instrumented systems are also referred in the literature as safety-critical systems or electrical/electronic/programmable electronic (E/E/PE) safety-related systems, according to the technology used to develop them (Liu, 2020).

To determine if a SIS can be applied, it is necessary to show that it is operating correctly and will perform its function when a demand occurs. To do so, the producer needs to demonstrate, that each safety instrumented function (SIF) implemented by the SIS is within the desired reliability range, and that measures to detect, prevent and avoid hardware, software and systematic failures have been implemented (Lundteigen et al., 2009).

Functional safety and *safety integrity* are key concepts used to describe the desired performance of a SIS (Lundteigen et al., 2009). Functional safety is used to describe the overall safety of a system, whereas safety integrity describes the ability of a SIS to perform its safety functions. Safety integrity is therefore a measure of reliability of a given SIS/SIF and is expressed in a discrete scale comprising four levels. Each Safety Integrity Level (SIL), lies within a specified reliability range, where SIL1 represents the lowest level and SIL4 the highest. In order to attest a system's integrity, the industry follows standards and guidelines.

Some of the most important standards used on the Norwegian Continental Shelf (NCS) are the standards IEC61508 (2010) and IEC61511 (2016) , and the guideline NOG-070 (2020) which provide input on how to design, install, operate and maintain safety-related systems. These standards use the safety lifecycle as an approach to structure the SIL requirements. However, for designing a SIS or SIS sub-systems, the IEC61508 (2010) is the most relevant standard. From a producer of safety-related systems point of view, the standards lack orientation from a product development perspective, as the general requirements focus more on the overall system. Furthermore, there are other concerns that must be explored in order to deliver a final product that meets the customer requirements and ensure failure-free performance of the system.

1.2. Objectives

The main objective of the thesis is to present a holistic approach to improve the management of SIL requirements during the design of safety instrumented systems. This is going to be done by achieving the following sub-objectives:

i) Present the requirements necessary to design safety instrumented systems (SIS)

ii) Clarify the role of risk assessments in SIS/SIS element design and its limitations

iii) Propose a framework that allow a SIS producer to meet the safety requirements and cover the limitations presented

1.3. Structure of the thesis

This work is organized in 5 Chapters. The first chapter provides an overview of the project, initial considerations, and objectives of this study. Chapter 2 presents a review of concepts and the requirements necessary to the design phase of safety-related systems used in the Norwegian oil and gas industry. Chapter 3 discuss the role of risk assessment in SIS design. In Chapter 4, a framework incorporating the perspectives presented in the previous chapter is presented. Chapter 5 is dedicated to presenting conclusions as well as recommendations for future work.

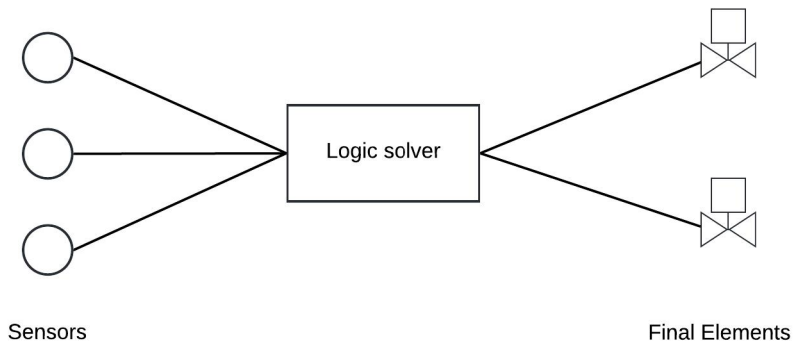
2. REVIEW OF BASIC CONCEPTS AND REQUIREMENTS FOR DESGINING SAFETY RELATED-SYSTEMS

2.1. Safety instrumented system (SIS)

Safety systems are extensively applied in the oil and gas industry to prevent hazardous events from occurring. A variety of systems can be applied for different purposes, usually they are designed to act differently and independently from each other, forming multiple protection layers (Chang et al., 2015). In general, safety systems are designed to perform one or more safety instrumented functions (SIF). The system under the protection of a SIS is denominated equipment under control (EUC) (Rausand, 2014).

A SIS generally consists of sensors, a logic solver and final element, presented in Figure 1. The sensors are responsible for monitoring and detecting any changes in the environment. Further, the information gathered by the sensors is sent to the logic solver in the form of signals. The logic solver interprets the signals and decides if an action against a specific condition, that is, if a SIF, must be performed. Lastly, the final element performs the preventive action, and is composed by one or more actuators (e.g., valves, circuit breakers, motors) (Gabriel et al., 2018).

Figure 1: SIS subsystems



Retrieved from: Murthy et al. 2008.

2.2. Safety instrumented function (SIF)

A SIF is intended to act against a *specific* hazardous event. In short, the main objective of a SIF is to bring the process or the EUC to a safe state. In addition, the parts of a SIS responsible of performing the SIF are called *safety loop* (Rausand, 2014).

2.3. Safety integrity level (SIL)

Safety integrity level (SIL) is a measurement of performance for a given SIF. This measurement has been introduced by the IEC61508-4 (2010) as the probability of a SIS to satisfactorily perform the specified SIFs under all the stated conditions within a specific time interval. It is important to highlight that a SIL is always related to a SIF and not to the SIS itself. SILs are presented in a discrete scale ranging from SIL1 to SIL4. These requirements reflect the reliability of a given SIF, with SIL1 presenting the highest probability of failure and SIL4 the smallest (Rausand, 2014).

2.4. Failure and failure classification

A *failure* can be defined as the "termination of the ability of a component to perform its required function" (Rausand, 2014). In other words, if the component is no longer able to perform its designed functions, than a failure has occurred, and the component is *failed*. Moreover, the component will stay in the *failed state* unless it is repaired. In this case, the component has a *fault* (Rausand, 2014). A fault is a state and can be due to random hardware failures or due to systematic failures.

There are several criteria used to classify failures. The NOG-070 (2020), categorizes failures in relation to consequence and detectability. In relation consequence, a failure can be classified as:

- *Dangerous failure* is a failure that impedes or disables a safety function.
- *Safe failure* is a failure that does not have the potential to put the system in a hazardous state.

In relation to detectability, a failure can be:

- *Detected* is a failure that is revealed by self-diagnostic (or automatic diagnostic) testing.
- *Undetected* is a failure which is not detected by self-diagnostic testing.

Furthermore, the hardware failures can be categorized as (Rausand, 2014) :

- *Dangerous undetected (DU)* failures are revealed only when a demand occurs. DUs are used to calculate the system reliability and are the main contributor to SIF unavailability.
- *Dangerous detected (DD)* failures are detected by self-diagnostic testing. The average time from when the DD occurs until the function is restored is called mean time to restoration (MTTR). That is, the average period where the SIF is unavailable due to DD.
- *Safe undetected (SU)* are non-dangerous failures which are not detected by self-diagnostic testing.
- *Safe detected (SD)* are non-dangerous failures that are detected by self-diagnostic testing.

2.5. Common cause failures (CCF)

Common cause failures (CCF) result from one or more events and are serious threat to SIS reliability. In a broader way, CCFs can cause redundant elements and safety barriers to fail simultaneously. This type of failures may result from the design or from operation. Failures caused by design are due to inadequate selection of hardware components and lack of understanding of failure mechanisms. CCFs resulting from operation may be caused by improper testing, maintenance, human errors and environmental stresses. (Lundteigen & Rausand, 2007).

The standards and guidelines used in the Norwegian oil and gas industry, focus more in CCFs caused by design and require their effect to be considered in reliability calculations.

2.6. Probability of failure on demand (PFD)

The probability of failure on demand (PFD) is a measure of reliability of a system. In the NOG-070 (2020) "*PFD is defined as the average probability that a safety system is unable to perform its safety function upon a demand"* and can be denoted as PFD_{avg}. For a single component, the PFD_{avg} is a function of the dangerous undetected failures (λ_{DU}) and the average period of time the component is unavailable during the proof test interval ($\tau/2$) as per Eq. (1) below:

$$PFD_{avg} \approx \lambda_{DU} \cdot \tau/2\,, \qquad \text{Eq. (1)}$$

where τ is the average duration of proof test interval.

For redundant systems, e.g. *koon* voted systems, the PFD_{avg} becomes a function of the independent failures and common cause failures. One of the most common ways of quantifying the PFD_{avg} is by the PDS method, which is a framework established to determine the unavailability of a SIS. This method accounts for all major factors affecting reliability during system operation, such as (SINTEF, 2017):

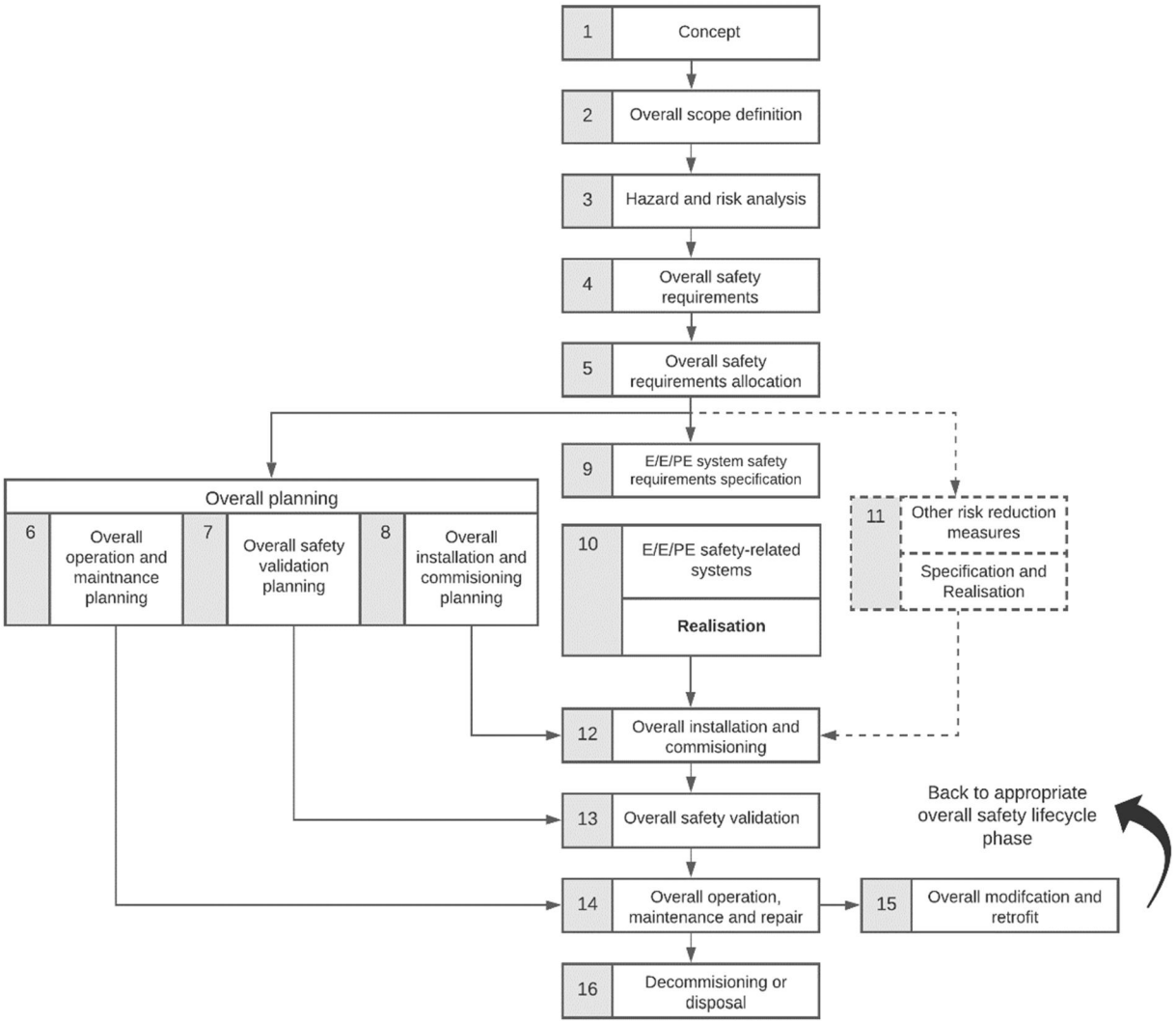
- All major failure categories/causes
- Common cause failures
- Automatic self-tests
- Functional (manual) testing
- Systematic failures
- Complete safety function
- Redundancies and voting logic

For the remainder of this work the PFD_{avg} will be referred only as PFD.

2.7. Safety lifecycle

The safety lifecycle is a concept introduced by both IEC 61508 (2010) and IEC 61511 (2016). The intention is to bring a framework to assess the requirements related to each identified phase of a SIS/SIF. The purpose of the safety lifecycle is to promote a consistent approach to identify and reduce risk in accordance to tolerability limits. The framework covers requirements from specification, design, operation, maintenance, to modification and decommissioning (NOG-070, 2020). The safety lifecycle is presented in Figure 2.

Figure 2: Safety lifecycle



Retrieved from: IEC 61508, 2010

Essentially, the safety lifecycle can be divided into three phases: analysis, realisation and operation. Analysis comprises the all the activities dedicated to identifying hazards, determining target SIL and to defining safety requirements specification (SRS). Realisation comprises the SIS design, SIL verification and SIS installation and test. Whereas the operation encompasses SIS operation and maintenance as well as decommissioning (Chang et al., 2015).

2.8. Hardware safety integrity

The requirements for hardware safety integrity analyse two different aspects of a SIS reliability: the quantitative requirements related to the average probability of a SIS to fail to perform its design function on demand and qualitative requirements related to architectural constraints.

2.8.1. Quantitative requirements

The estimation of hardware failures will depend on the demand mode that the SIS is intended to operate. A *demand* is a process deviation that requires the SIF to be activated. According to the frequency of demand, the probability of failure will be estimated based on the probability of failure on demand (PFD) or the probability of failure per hour (PFH). For systems operating in low-demand mode, the demand frequencies are typically less than once per year and the PFD is then used. While for systems operating in high or continuous demand mode, the demands are expected to occur several times in a year, or as a part of normal operation. In this case, the PFH is taken into account (Rausand, 2014).

Table 1: Safety integrity levels for safety functions operating on demand or in a continuous/high demand mode

Safety Integrity Level	**Demand Mode of Operation** (average probability of failure to perform its design function on demand - PFD)	**Continuous / High Demand Mode of Operation** (probability of a dangerous failure per hour - PFH)
4	$\geq 10^{-5}$ to $< 10^{-4}$	$\geq 10^{-9}$ to $< 10^{-8}$
3	$\geq 10^{-4}$ to $< 10^{-3}$	$\geq 10^{-8}$ to $< 10^{-7}$
2	$\geq 10^{-3}$ to $< 10^{-2}$	$\geq 10^{-7}$ to $< 10^{-6}$
1	$\geq 10^{-2}$ to $< 10^{-1}$	$\geq 10^{-6}$ to $< 10^{-5}$

Retrieved from: IEC61508, 2010.

2.8.2. Architectural constraints

The architectural constraints requirements take into account the robustness of the structure of the SIS by considering its subsystems (Lundteigen & Rausand, 2006). The necessary measures to determine the architectural constraints are expressed by Hardware Fault Tolerance (HFT), the Safe Failure Fraction (SFF) and depend also on the type of the subsystems that compose the SIS. Furthermore, the regulations provide two possible routes to fulfil the requirements of architectural constraints. According to NOG-070 (2020):

- Route 1_H based on the hardware fault tolerance (HFT) and safe failure fraction (SFF) concepts.
- Route 2_H based on reliability data from field feedback and similar devices, increased confidence levels of HFT for specified safety integrity levels.

The first route is recommended for development of new technology where no field experience is available, while the second requires the equipment to be developed in compliance with IEC 61508 or to be documented prior in use (NOG-070, 2020).

2.8.2.1. Hardware fault tolerance (HFT)

The HFT expresses the ability of a SIS subsystem to continue to perform a function in the presence of errors or faults. For a *koon* voted group the HFT is $n - k$ meaning that the group is functioning if at least k of its n channels are still functioning (Lundteigen & Rausand, 2006). For example, for a 2oo3 voting system, HFT = 1.

2.8.2.2. Safe failure fraction (SFF)

The SFF provides the fraction of overall hardware failure rate of device considered as safe or detected dangerous failure (Catelani et al., 2010). A SFF can be calculated as:

$$SFF = \frac{\Sigma\lambda_S + \Sigma\lambda_{DD}}{\Sigma\lambda_T} \qquad \text{Eq. (2)}$$

where $\Sigma\lambda_S$ is the sum of all safe failure rates, $\Sigma\lambda_{DD}$ is the sum of all detected dangerous failure rates and $\Sigma\lambda_T$ is the sum of all possible failure rates (safe and dangerous).

2.8.2.3. Type of the SIS subsystems

The SIS subsystem can be classified as being Type A or Type B. According to Rausand (2014), for a Type A the subsystem

- The failure modes of all components are well defined.
- The behaviour of the subsystem under fault condition can be completely determined.
- The exists failure data showing that the claimed rates of dangerous undetected and dangerous detected failures are met.

Whereas for a Type B subsystem at least one of the following statements are true:

- The failure mode of at least one component of the subsystem is not well defined.
- The behaviour of the subsystem under fault condition cannot be completely defined.
- The is not sufficient data to show the claimed rates of dangerous undetected and dangerous detected failures are met.

Furthermore, if the Route 1_H is used, the IEC 61508 (2010), defines the architectural requirements for different safety integrity levels in accordance with the type of the subsystems and corresponding SFF as presented in

Table 2: Architectural constraints for Type A and Type B systems

Safe failure fraction (SFF)	Hardware fault tolerance					
	Type A			Type B		
	0	1	2	0	1	2
< 60 %	SIL1	SIL2	SIL3	SIL1	SIL2	SIL3
60 % - 90 %	SIL2	SIL3	SIL4	SIL2	SIL3	SIL4
90 % - 99 %	SIL3	SIL4	SIL4	SIL3	SIL4	SIL4
> 99 %	SIL3	SIL4	SIL4	SIL3	SIL4	SIL4

Retrieved from: IEC61508, 2010.

2.9. Systematic safety integrity

Systematic safety integrity is referred as part of “Management of functional safety” in the context of NOG-070 (2020) and includes all activities required to ensure that all measures to identify, design and maintain SIL requirements during the entire lifecycle of the systems. The systematic safety integrity is specified by qualitative requirements and demands extended examination of the design. The IEC61508-2 (2010) provides the recommendations for techniques to control failures caused by hardware design,

environmental stress or influence, and to control failures during operation. Additionally, the NOG-070 (2020) presents some general examples of improvement areas for avoidance and control of systematic failures such as: avoidance of unnecessary complexity, ensure that the equipment is fit for its intended use to name a few. Furthermore, systematic failures are defined in the IEC61508-4 (2010) as a "failure, related in a deterministic way to a certain cause, which can only be eliminated by a modification of the design or of the manufacturing process, operational procedures, documentation or other relevant factors"

2.10. Systematic capability

According to the IEC61508-2 (2010), to achieve the required SIL the elements of the SIS must comply with systematic safety integrity requirements. In other words, systematic capability (SC) is a measure of confidence that the elements' systematic safety integrity meets the requirements of a specified safety integrity level (Creech, 2014). Furthermore, three routes that can be used to achieve SC are defined:

- Route 1S: This route covers the requirement for elements and components designed in accordance with IEC61508 (2010).
- Route 2S: Covers proven-in-use components.
- Route 3S. Covers pre-existing software elements.

3. REVIEW AND DISCUSSION OF METHODS USED TO ALLOCATE AND VERIFY SIL REQUIREMENTS

The SIL requirements are a risk-based approach to determine if a SIF can provide the preestablished level of risk reduction. The requirements are based on industry needs and derived from a long history of operation of safety-related systems. They serve as a reference to dimensioning systems and determining the desired performance/safety requirements of the system. The standard approach suggested by the IEC61508 (2010) uses risk analysis and risk acceptance criteria to establish the required properties of the system. Both qualitative and quantitative assessments are used to establish a risk picture and are further translated into tolerability limits and necessary levels of risk reduction. Whether qualitative or quantitative assessments are used, risk acceptance criteria in the context of IEC61508 (2010) state what are the tolerable limits with respect to the frequency (or probability) of the hazardous event and its specific consequences (Dean, 1999). In this sense, quantitative risk assessment plays a critical role in the design process and is used to understanding the risk a process, activity or system is subjected to. Hence, being an essential tool to allocating and managing necessary risk reduction measures.

Even though risk assessments are a powerful tool, they have limitations. These limitations need to be acknowledged, as they impact directly on the extent the analysis should be used in the decision-making process. To address some of issues that may arise when using the standard approach to allocate and verify performance requirements, Chapter 3 starts by elucidating the importance of risk assessments in safety integrity, how risk is represented in this context and what are the principles used to justify the implementation of safety measures. The Chapter also includes a brief explanation of the methods used and which stage of design they are applied, as well as what are their limitations and their impact on the overall performance, if they are not accounted for.

3.1. The role of risk assessment in safety integrity

Risk assessment is an essential part of the safety lifecycle and, therefore, extremely important for designing safety-related systems. It concentrates on defining the system risk scenarios. There are two distinct phases during the design that risk assessment methods are used:

- During the conceptual phase of the project, before detailed design commences, hazard identification and risk analysis are used establish the desired performance of the system and set parameters to achieve it.
- Throughout detailed design to identify potential failure modes, their impacts on the system and to perform reliability checks. This process is iterative and can be performed many times depending on the complexity of the system considered and how well the design is judged to meet the performance requirements.

For designing purposes the IEC61508 (2010) is the dominant standard, and presents a generic approach for systems intended to perform safety functions that are comprised of E/E/PEs (Gabriel et al., 2018). This standard follows a risk-based approach, which makes necessary to define some risk acceptance criteria that will be used to guide decision making through the safety lifecycle. The steps are:

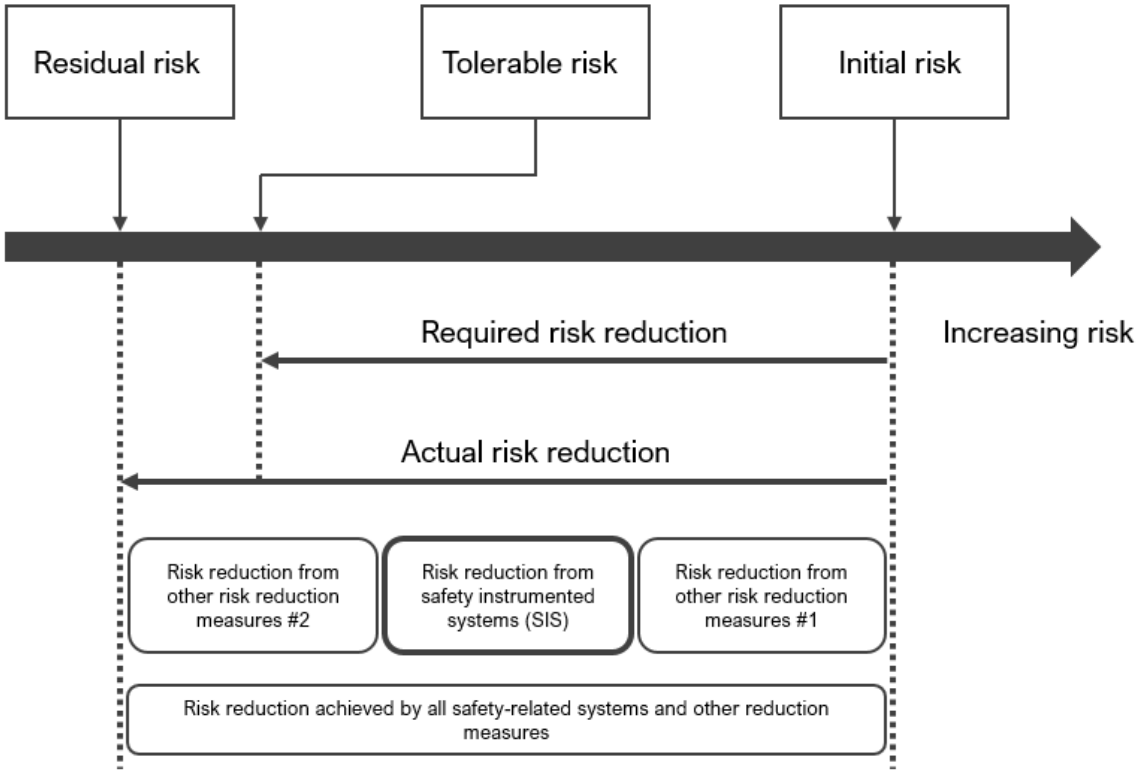
1. To understand the initial risk
2. To define the risk acceptance criteria
3. To identify the *necessary risk reduction* to meet the criteria
4. To define the necessary SIFs
5. To allocate SIL to each safety function
6. To verify if the designed product/system achieves the desired SIL

3.1.1. The concept of necessary risk reduction

The necessary risk reduction is the reduction in risk that must be achieved to meet the tolerable risk for a specific situation (IEC61508-5, 2010). The overall safety is achieved by having a combination of risk reducing measures, SIS and other barriers. All technical, operational and organisational elements must be accounted to determine the actual risk reduction, as indicated in the Figure 3.

The assessment can be conducted using any risk related standard, such as the NORSOK-Z013 (2010) , and will involve the following tasks: *hazard identification, risk analysis* and *risk evaluation* (PSA, 2018). The risk assessment will give the required safety integrity of the system for risk to be acceptable.

Figure 3: Necessary risk reduction



Retrieved from: NOG-070, 2020.

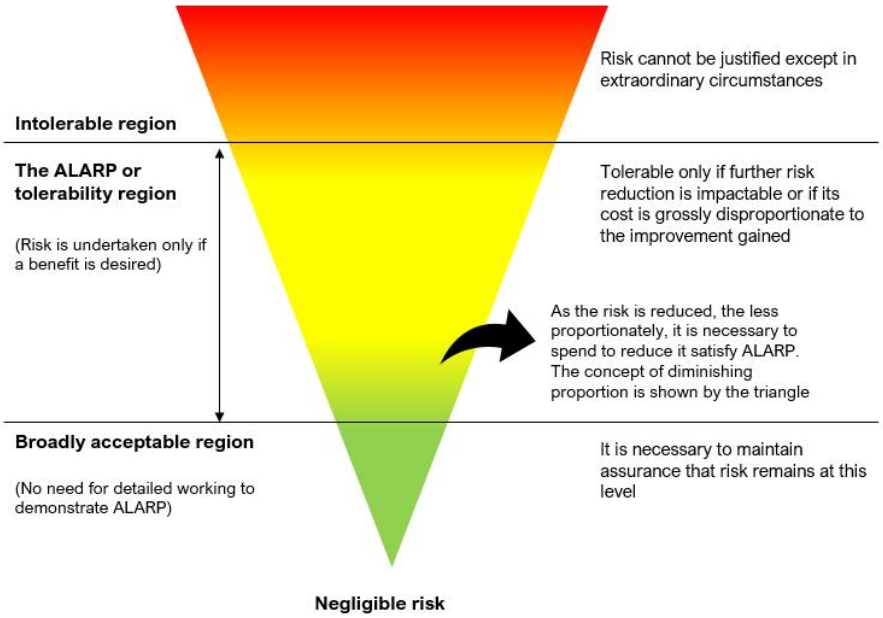
3.1.2. The ALARP principle

The ALARP principle states that risk should be kept in a level that is "as low as reasonably practicable", i.e. a safety measure shall be implemented unless it can be demonstrated that the costs are in gross disproportion to the benefits gained (Aven, 2014).

This principle is broadly used in the oil and gas industry in the NCS to guide decision-making in safety management (Abrahamsen et al., 2018). In general, cost-benefit analyses and the grossly disproportionate criterion are used to support decision-making. That is, to determine what is practicable cost-benefit analyses are used as a basis for the decision on whether a risk-reducing measure should be implemented (Bai & Jin, 2016).

The ALARP principle is illustrated in Figure 4, according to the IEC61508-5 (2010). In the ALARP region, risk is only tolerable if the costs of implementing the measure are disproportionate to the benefits, or if risk reduction is impracticable. This region is delimited by an "upper tolerable limit" and a "lower tolerable limit". If the risk is above the upper limit, risks are usually not undertaken, unless in extraordinary circumstances (Bai & Jin, 2016). ALARP may be used to determine the SIL requirements for safety functions (IEC61508-5, 2010). The approach proposed by the standard is to match a consequence of an event to a tolerable frequency. Depending on the frequency and the severity of the consequences, risk can be sorted into classes. Then, the SIL requirements for a specified safety function are increased until the frequency of occurrence is reduced to fall into the ALARP or broadly acceptable region (IEC61508-5, 2010).

Figure 4: ALARP principle



Retrieved from: IEC61508-5, 2010.

3.2. Hazard and risk analysis

To determine the level of risk reduction to bring the EUC to a safe state, it is necessary to describe the inherent risk associated with the hazards and to identify the requirements for risk reduction (NOG-070, 2020). The risk arising from the process can be determined by qualitative and quantitative techniques. Some commonly used techniques are:

- Hazard identification (HAZID)
- Hazard and operability studies (HAZOP)
- Process Hazard Analysis (PHA)
- Review of past data and experience
- Expert judgements
- Information in databases and data handbooks

Hazard and risk analysis should cover aspects such as (IEC61508-1, 2010 p. 28):

- the consequences and likelihood of the event sequences with which each hazardous event is associated;
- the tolerable risk for each hazardous event;
- the measures taken to reduce or remove hazards and risks;
- the assumptions made during the analysis of the risks, including the estimated demand rates and equipment failure rates; any credit taken for operational constraints or human intervention shall be detailed.

3.3. Defining safety integrity requirements – SIL Determination

The information from the Hazard and Risk Analysis serves as an input to determine all SIFs necessary to protect against the identified risks and to allocate the associated SIL requirements (Gabriel et al., 2018). In other words, risk analysis in this phase provides answers for the following questions (IEC61508-5, 2010):

- "What safety function has to be performed?" – the safety function requirements
- "What degree of certainty is necessary that the safety function will be carried out?" – the safety integrity requirements.

The safety integrity requirements are defined using a risk-based approach for determining the SIL and setting a numerical target for failures related to the SIL, which is the maximum PFD (or PFH). The determination of SIL requirements in the oil and gas industry is done by using the risk-graph method, the Layers of Protection Analysis (LOPA) or using the Minimum SIL requirements table proposed by the NOG-070 (2020). The operator or the system integrator will therefore use the determined SIL values to select the components/elements that are part of the SIS. For example, for the most common functions, the Minimum SIL requirements present the required SIL as shown in Table 3.

Table 3: Minimum SIL requirements for a "standard" function

SIF	SIL	Functional boundaries
Isolation of production/injection bore in one topside well from the production/injection manifold/flowline	SIL 3	The function starts at the unit where the demand is initiated (unit not included) and ends with the valves shutting in the well. The following equipment is needed: • ESD logic (wellhead control panel) incl. I/O • PWV OR PMV OR Down hole safety valve (DHSV), incl. solenoid(s) and actuator

Retrieved from: NOG-070, 2020

3.4. Verifying the provided level of risk reduction – SIL Verification

After all requirements haven been stablished, design takes place and the hardware and systematic safety integrity (as referred in Chapter 2) must be verified for each safety function. This is a crucial step in the safety lifecycle, once it is investigated if the designed SIS meets the required failure measure (Gabriel et al., 2018).

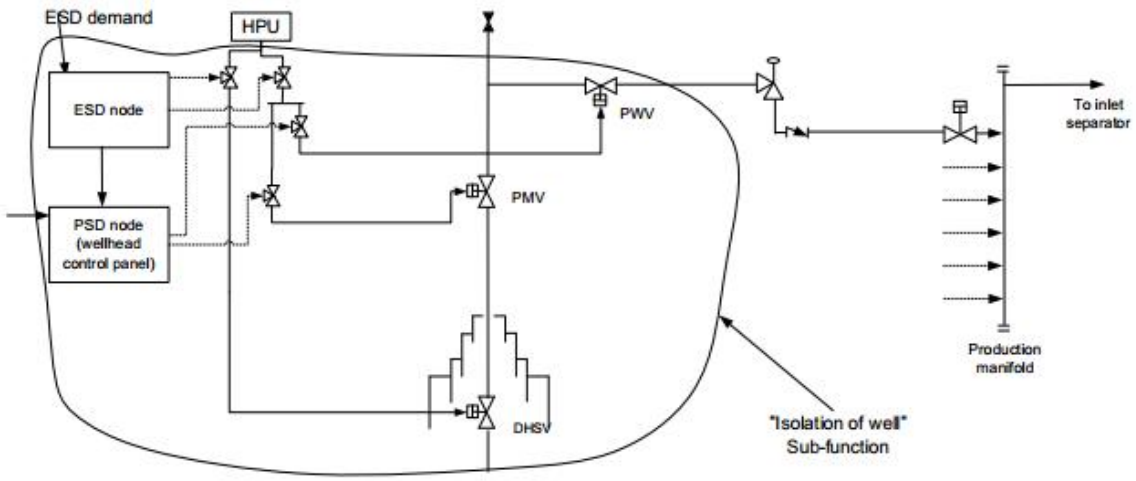
There are many methods that can be used to SIL verification. The process can be carried out by using Reliability Block Diagrams (RBD), Fault Tree Analysis (FTA) and Failure Modes and Effects Analysis (FMEA). Typically, a Failure Modes and Effects and Diagnostic Analysis (FMEDA), which is a variant of the FMEA, is used to provide input to the reliability calculations for each component of the safety-loop and determining the architectural constraints.

The FMEDA is realised by each component or sub-system producer and provides a way of classifying failure rates for safety systems. The analysis is conducted by a team of experts from different backgrounds that have experience in analysing and quantifying the relationship between the effects, risks and improvement actions (Huang et al., 2020). Each component or subsystem is analysed separately, and the experts list all the failure modes, effects, causes and detection strategies according to their knowledge and working experience. This information allows reliability improvements to be performed (Huang et al.,

2020). Further, the designers to verify the PFD value for the overall system using probabilistic modelling methods for the architecture of the SIS taking into account the criteria presented in Table 1 for the selected SIL.

For example, for a “standard” function, such as the emergency shut down (ESD) “Isolation of production/injection bore in one topside well” illustrated by Figure 5, the NOG-070 (2020) presents the PFD requirements for all the elements in the safety-loop.

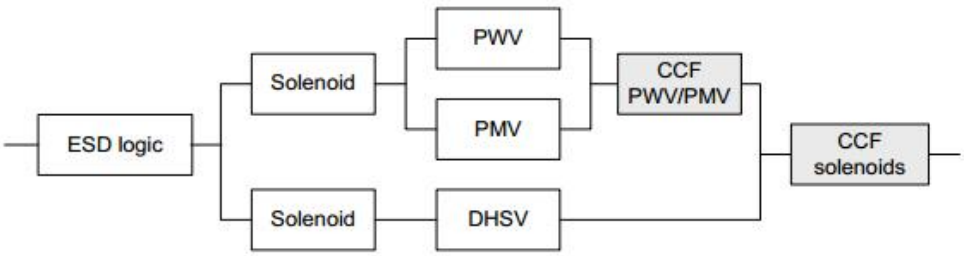
Figure 5: Isolation of production/injection bore in one topside well



Retrieved from: NOG-070, 2020

The reliability block diagram for this function represented in the Figure 6 and the PDF for each component is given in Table 4.

Figure 6: RBD for the ESD function



Retrieved from: NOG-070, 2020.

Table 4: PFD results for the ESD function

Component	Voting	PFD per component	PFD	
			CCF	Indep.
ESD Logic (redundant I/O and redundant CPU)	1oo1	$1.9 \cdot 10^{-4}$		$1.9 \cdot 10^{-4}$
Upper branch:				
PMV/PWV Solenoid	1oo1	$2.6 \cdot 10^{-3}$		
PMV/PWV	1oo2	$4.4 \cdot 10^{-3}$	$4.4 \cdot 10^{-4}$	
Total upper branch (indep.)		$\mathbf{7.0 \cdot 10^{-3}}$	$\mathbf{4.4 \cdot 10^{-4}}$	$9.5 \cdot 10^{-5}$
Lower branch:				
DHSV Solenoid	1oo1	$2.6 \cdot 10^{-3}$		
DHSV	1oo1	$7.0 \cdot 10^{-3}$		
Total lower branch (indep.)		$\mathbf{9.6 \cdot 10^{-3}}$		
CCF solenoids	**1oo2**		$\mathbf{2.6 \cdot 10^{-4}}$	$2.6 \cdot 10^{-4}$
Total for the function				$\mathbf{5.5 \cdot 10^{-4}}$

Retrieved from: NOG-070, 2020.

Based on the RBD and the subsystems PFDs the SIL for the function can be determined. For this case, the safety function satisfies SIL 3, as the PDF lies within the range $\geq 10^{-4}$ to $< 10^{-3}$. As presented previously, there are many methods that can be used to determine the PFD. For example, the NOG-070 (2020) suggests the PDS method, which uses reliability theory and field historical observations. These observations are found in databases such as the OREDA database[1]. The discussion related to specific methods used to calculate the PFD will not be elaborated in this report as the remarks presented in the following section may be applied to all calculation methods.

[1] SINTEF Technology and Society, Norges teknisk-naturvitenskapelige Universitet, & DNV GL *(2015)*. *OREDA: Offshore and onshore reliability data handbook*

3.5. Limitations

The standard approach suggested by the IEC61508 (2010) uses risk analysis and risk acceptance criteria to establish the required properties of the system. QRAs are used to establish a risk picture and are further translated into tolerability limits and necessary levels of risk reduction. Once the RAC has been established and the need for risk reduction identified, QRAs are used to define a set of performance requirements that need to be delivered by the system. Usually, the performance requirements are seen in relation to the probability of a safety function being performed, which, in turn, form the basis for the design.

During the design of safety systems, these requirements are used and referred on a regular basis and new QRAs are conducted to determine if the predefined limits are met. Even though, QRAs are a valuable tool, there are limitations that need to be acknowledged. These limitations impact directly on the extent the analysis should be used in the decision-making process, especially when considering new systems.

For designing systems, the use of QRAs raise several issues related to uncertainty and knowledge. According to the new risk analysis perspectives, risk analysis should not be limited to presenting probabilities and consequences of an event (Aven & Ylönen, 2016), as proposed by the IEC61508 (2010). Uncertainty and strength of knowledge should also be addressed. These uncertainties are caused by randomness due to systems inherent variability (aleatory uncertainties), or due to imprecision resulting from lack of knowledge (epistemic uncertainties). Furthermore, epistemic uncertainty can be reduced if the knowledge increases (Chang et al., 2015).

Probabilities are a natural choice to represent uncertainty. However, the probability numbers are conditional to a background knowledge and represent the degree of belief of the analyst in relation to an event. This knowledge can be more or less strong depending on the amount of information available. Thus, the strength of knowledge should be considered, since there might be underlying factors that are not reflected by the assessment (Aven, 2014). In this section, it is evaluated to what extent the use of risk acceptance criteria and the tools used to verify if the criteria are met can be justified. The section end with discussion related to the limitations of the methods, allowing one to understand and decide how much weight the QRA should be given.

3.5.1. The use of risk acceptance criteria

The first issue that may result from the use of such standards, as discussed above, is related to the specification of quantitative thresholds for risk acceptance. First, defining thresholds for acceptable risk requires a detailed understanding of "dose-response relationship" (Coglianese et al., 2003). That is, to fully understand all the links between the multitude of causal agents and specific observed effects. This becomes particularly difficult when one tries to predict the system's performance just by looking at how the individual parts or components work (PSA, 2020). Thereby determining optimal thresholds for risk acceptance can be a demanding task, as predicting the interaction between the components may be hard (Coglianese, 2003 and PSA, 2020). For quantitative estimations of risk, the quality of data, as well as its relevance will also play a significant role on defining the RAC. Second, one must consider that uncertainty can be also related to the models used, as these represent simplifications of real-world phenomena (Aven, 2014). Since it is virtually impossible to model every possible scenario, predictions are limited to the scope of the analysis. The scope, in turn, is defined based on assumptions related to these potential scenarios (Coglianese et al., 2003). Judgements about the likelihood of a scenario happening or not, can lead to disregarding an event with potential for disaster if the probabilities are considered low. Hence, the risk can be very different from what it has been presented by the risk picture, and surprises can occur. As this information is crucial for developing the requirements, estimates can lead to weak formulation of limits if the risk picture is misrepresented. On the other hand, the opposite can also occur. If risk is overestimated, the requirements will lead to too stringent limits, meaning more resources will have to be used to design a system that provides higher level of risk reduction. Although, overestimation of RAC is less likely to happen as companies are prone to set limits that are easier to achieve (Abrahamsen & Aven, 2012). Yet, strict limits can lead to increased complexity and over expenditure, and this possibility needs to be addressed (Abrahamsen & Abrahamsen, 2015).

When considering the oil and gas industry, it can be argued that it is possible to list all sorts of adverse events that can happen, but how they materialise is not as straight forward (PSA, 2020). Which leads to most systems being described as complex. However, when taking a closer look to accidents and near-misses, one may find that despite the large quantity of components, there is a good understanding of most systems and activities, and how these components interact. Still, surprises happen, and when they happen, are mainly caused due

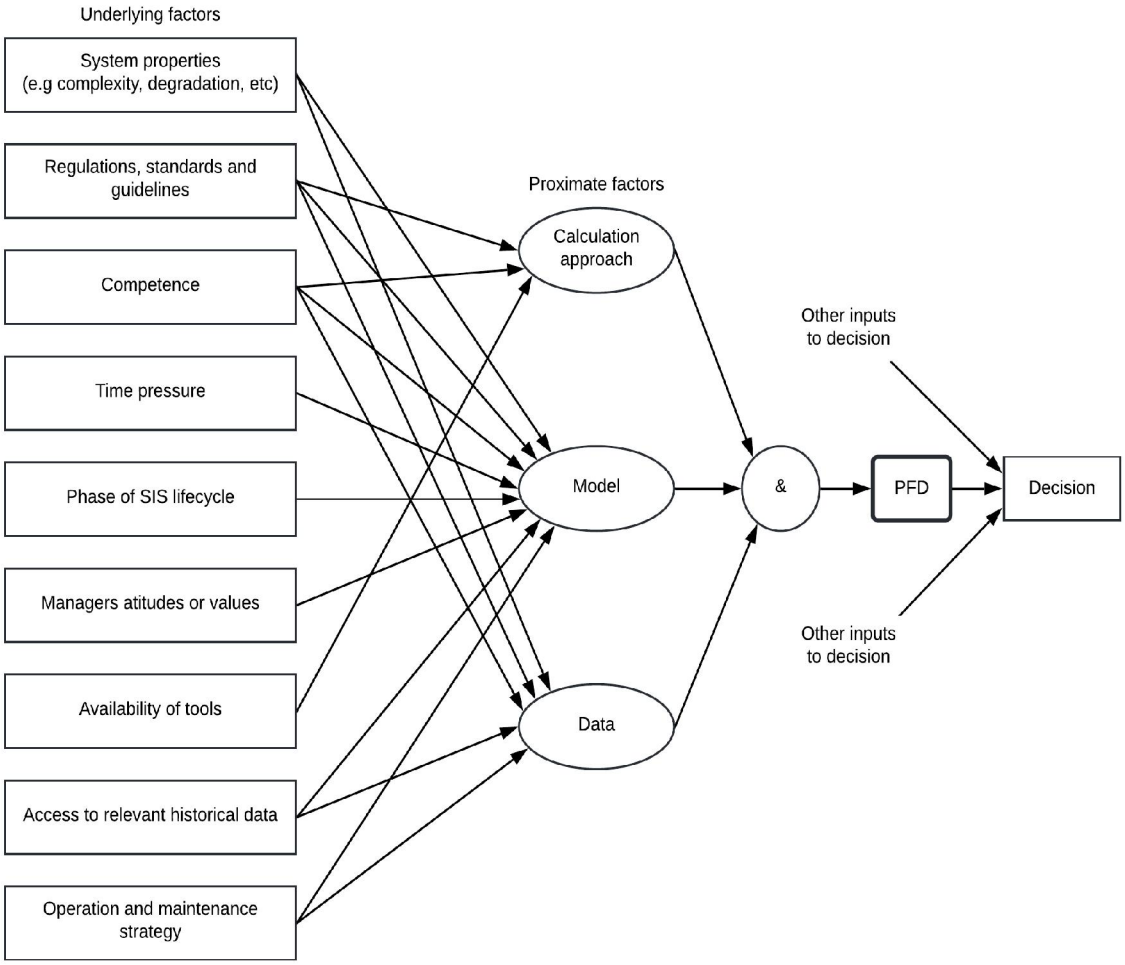
to incorrect or deficient usage of knowledge but not due to lack of understanding about the process and governing phenomena. To meet the challenges related to potential surprises, robustness/resilience must be emphasised, so the systems are able to withstand and correct problems efficiently (PSA, 2020).

Another problem highlighted by Aven and Ylönen (2016) is the used of RAC to determine minimum performance criteria, which seems to contradict the ALARP principle. It is important to reflect if risk is indeed reduced to the lowest level when limits are based on minimum performance requirements. Without paying the proper attention to the level of uncertainty and knowledge ALARP can be used in a way that does not contribute to actual risk-reduction (PSA, 2018). For instance, by making information that contributes to risk less accessible to decision-makers. The decision will depend only on the assumptions and choice of risk assessment method, while factors can be hidden away in the calculations. Underlying knowledge is, therefore, a key element to establishing RAC and balancing concerns that are not captured by probabilities and expected values. Furthermore, there is only one category where it is possible to use quantitative risk acceptance criteria in the form of upper limits: in situations characterized by limited uncertainties and where one possess a strong knowledge base so that in practice the potential surprises can be disregarded (PSA, 2020).

3.5.2. The use of QRAs

The second issue is related to the tools used to verify if the criteria are met. Safety and reliability assessments are used to this purpose and provide useful information to SIS designers and manufacturers, as well as to end users (Lundteigen, 2008). In this context, the methods used to perform SIL verification are conditional to underlying factors. Uncertainty in this setting is often related to (i) the model (ii) the data and (iii) the calculation approach. These aspects will influence the level of uncertainty in PFD calculations, and are related to what extent the model, the data and the calculation approach are representative of the system in question. Furthermore, there are underlying factors that may influence uncertainty levels in PFD calculations as indicated in the Figure 7.

Figure 7: Factors that influence the PFD calculation



Retrieved from: Lundteigen (2008)

The model is a simplification, as already mentioned. In relation to SIS, it expresses the degree of belief regarding the system's functions, components, configuration, modes of operation, other operational assumptions, etc. The way the model is constructed can be influenced by various factors, as shown in Figure 7. For instance, different industries and applications will use different modelling strategies, where various regulations standards and guidelines are recommended.

The data is used as input parameters for the reliability model. It comprises failure rates, test intervals, and so on. This data is subjected to uncertainty associated with its quality, relevance, and quantity. Uncertainty related to the amount of data available is often referred as statistical uncertainty and it is usually represented by probability density functions, cumulative distributions, or confidence intervals for the parameter values. Furthermore, data

will also influence model selection, as it is not advantageous to select a model to which data is not available. However, in situations where the technology is new, the reliability data is derived from similar systems or based on expert judgements.

The calculation approach is related to how the PFD is determined. It can be calculated using approximations formulas or exact expressions. It is also important to stress that usually, the results produced by these approaches do not present major discrepancies. Hence, the selection of approach is down to the analyst's preference. One contributor to uncertainty in this setting, is whether to use a time dependent PFD or the average value (Lundteigen et al., 2009). The IEC61508 (2010) recommends the average PFD, which can raise questions regarding the use of expected values, as they do not give much information about variation of the estimated quantity. More information on how underlying factors may affect reliability determination is given in Table 5.

Table 5: Influence of underlying factors on PFD calculations

Underlying factors	Data	Model	Calculation approach
System properties	The properties of a given system may deviate from the ones in historical data	Life distribution of SIS components. For example: time to failure is exponentially distributed	-
Regulations and standards	Different guidelines on how to handle uncertainty. For example, IEC61508 requires a 70% confidence interval for failure modes	Recommend different modelling strategies	Recommend different calculation approaches.
Competence	Use of expert judgements to estimate input parameters when data is not available	The experience of analysts may influence their choice of model. Experienced analysts might prefer more complex models, for example.	May influence the choice of calculation approach
Time pressure	-	Time constraints may influence model selection	-
Phase of SIS lifecycle	-	In early design phases simpler model might be used, while in later phases these models are update by more complex ones	-
Managers attitudes or values	-	If the focus is only to comply with regulations and standards the scope of reliability analysis might be limited	-
Availability of tools	-	Different tools may offer different options of models	Different tools may offer different options of calculation approach
Access to relevant historical data	Historical data is usually based on a number of installations/systems having components with different technologies. To make proper selection of input parameters, access to underlying information is required.	May offer guidance on dominating features that should be considered, such as critical/relevant failure modes	-
Operation and maintenance strategy	These strategies can give an idea about the distribution of down times. mean time to repair, etc.	Whether to consider planned or unplanned down times	-

Adapted from Lundteigen (2008)

The sources of uncertainty are somewhat addressed by the standards. Additionally, it is important to consider that the assessments are based on a series of assumptions and subjected to variation. These assumptions are specially related to the system properties and under which conditions it will operate. Uncertainty needs to be accounted for, especially in relation to changes in these scenarios. If the potential for variation of assumptions and operating conditions are not accounted for, the final system may not be suitable for its intended use. The result can be a product that is either too complex or too simple to provide the necessary risk reduction (Lundteigen, 2008). Of equal importance is to consider the knowledge dimension and seek opportunities to increasing it throughout all phases of the system.

4. A HOLISTIC APPROACH TO DEVELOPING SAFETY-RELATED SYSTEMS IN COMPLIANCE TO THE IEC61508

So far in this work, it has been discussed some challenges related to design of safety-related systems, where some limitations related to methods used to determine and verify the requirements where presented in the previous section. For systems that are designed for high-reliability applications, such as ones applied in the oil and gas industry, compliance with a series of requirements must be shown in order to "assure that the systems have the required reliability and quality before they are put into operation" (Rahimi & Rausand, 2015). Furthermore, to build safety into design one must adopt practices that produce the required safety attributes (Drogoul et al., 2007). In this sense, reliability (or SIL) is determined by the technical decisions made during the design, development, and production phases of the product life cycle (Murthy et al., 2008). Nevertheless, when new technological solutions are presented (i.e., products containing new materials and/or unproven[2] components), producers must show that the final product is *fit for purpose* (Rahimi & Rausand, 2015). In these situations, the producer of the of SIS components or entire SIS applications must ensure that the risk related to the technology was reduced to an acceptable level (Zikrullah et al., 2019). However, this task is not as straight forward as it might seem, as some risks only become apparent once the technology is employed (van de Poel & Robaey, 2017). It is, therefore, the producer's job to build confidence regarding the performance of the presented solution. To achieve this task, it may be fruitful to have a well-defined process to follow. This is done by finding ways to overcoming the challenges presented, such as the uncertainty related to the product performance.

In this chapter a framework that allows a SIS producer to meet the safety requirements and cover the limitations presented. The framework is divided into six phases, as presented in Figure 8. It starts from the definition of performance requirements necessary to achieve the intended SIL to installation, commissioning, and operation of the SIS. This chapter

[2]Unproven: Design or concepts that are have not been previously applied and do not count with operational experience and/or previous engineering documentation and analyses.

provide the reader a picture of the overall process used to design a safety-related system and demonstrates how the producer's responsibility changes throughout the safety-life cycle.

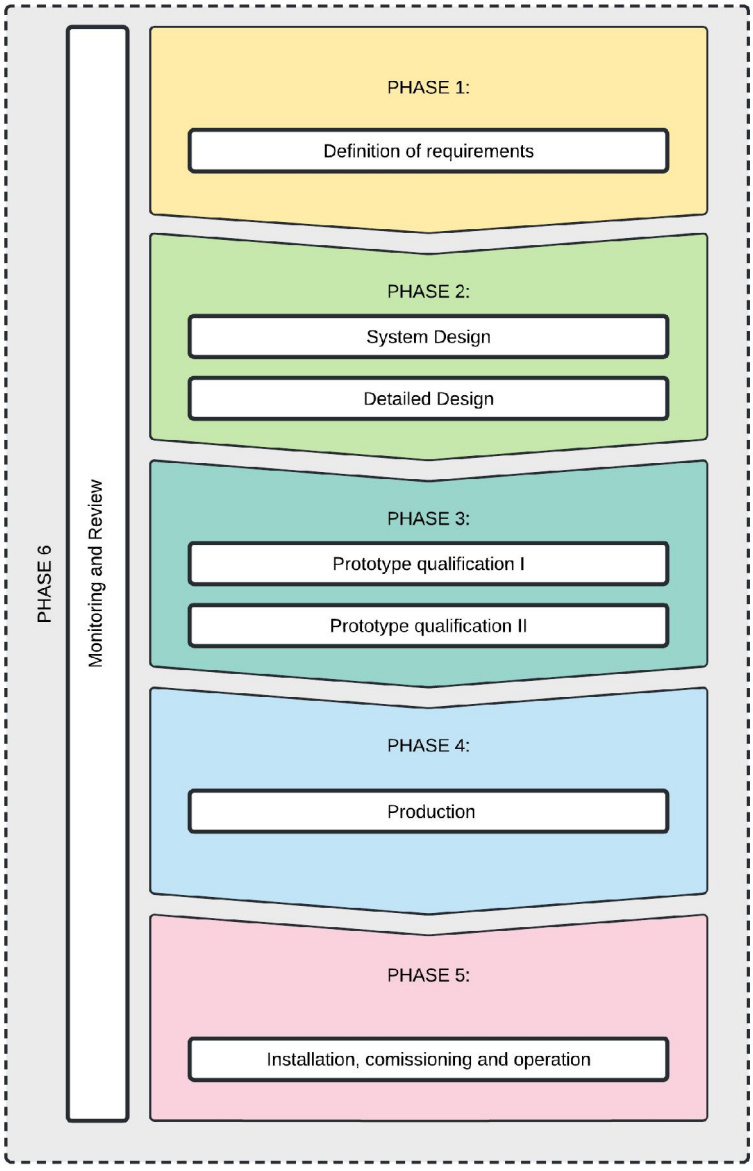
4.1. The framework

The design of a safety-related system follows the requirements of the IEC61508 (2010) and other sector specific standards. The IEC standards present a framework based on the safety lifecycle, which embraces the design, installation, operation, maintenance and decommissioning of a SIS. For the producer of such systems/components, however the most relevant phase in the safety lifecycle is the phase 10 (Realisation). The main problem is that the standards are not very detailed about the product development phases (Lundteigen et al., 2009). The realisation of new product can be based on many different models. Usually, companies have their own internal procedures detailing this process, which from an overall perspective is possible to identify the similarities between them. Perhaps, the most common feature between different processes is that the NPD starts with the idea of building a product to meet specific needs and ends with the product being launched (Osteras et al., 2006).

In the literature, the product development model of Murthy et al. (2008) is presented as a good alternative as it can be integrated the IEC 61508 safety life cycle, as presented by Lundteigen et al. (2009) and Rahimi and Rausand (2015). It comprises eight phases that are divided into a matrix containing three different stages (pre-development, development, and post-development) and levels (business, product, and component). In this work, however, this model is adapted into 6 phases, where the final phase, monitoring and review is presented in parallel to the other phases. Since the producer must provide evidence that the requirements of a standard are adequately met, monitoring and review activities must occur in parallel to the remainder phases of product development model. Experiences and information should be properly transferred and communicated to the interested parties throughout the process. The following phases are described in this section:

- Phase 1: Definition of requirements
- Phase 2: Design
- Phase 3: Prototype qualification
- Phase 4: Production
- Phase 5 Installation, commissioning and operation
- Phase 6: Monitoring and Review

Figure 8: Product development framework



4.1.1. Phase 1: Definition of requirements

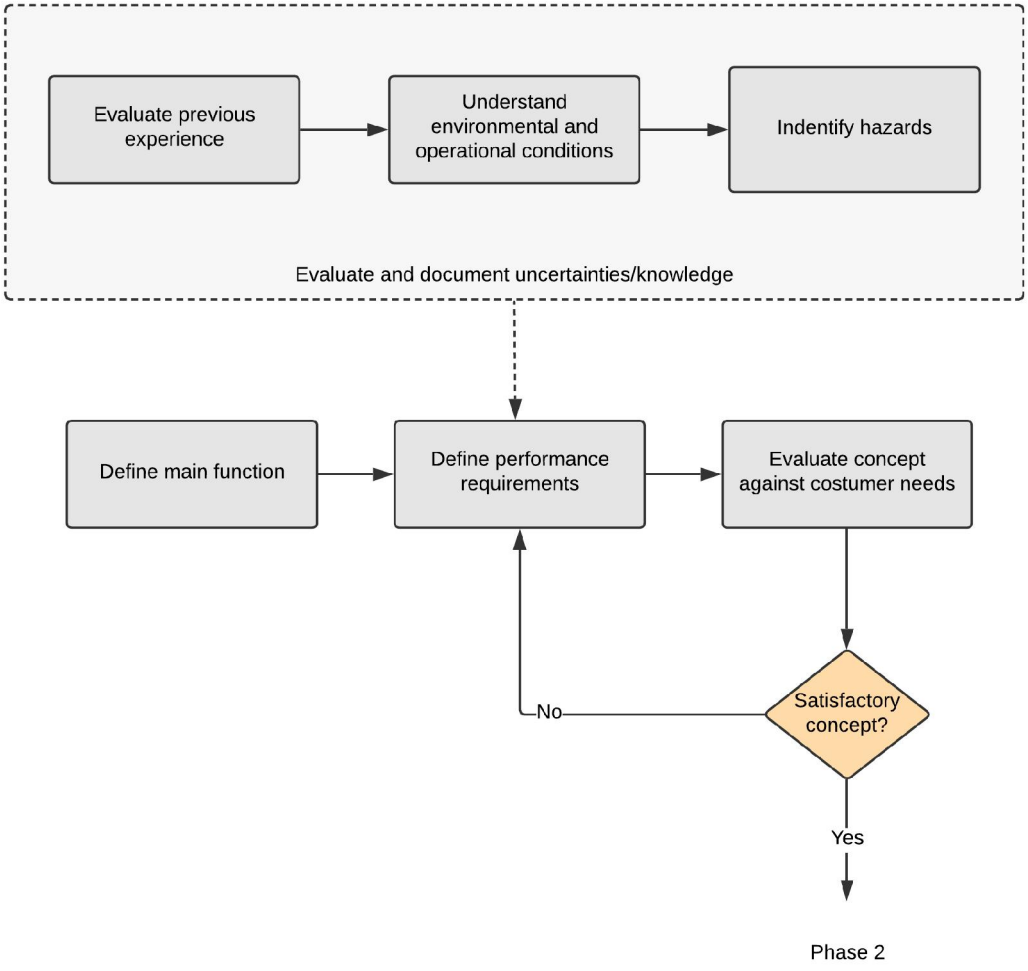
The first phase of the framework is related to identifying the need for a new product or modification of an existing product and aligning it to business objectives and strategies. In this phase the product only exists as a concept, and in most cases only the main functions are known. The producer must describe the product, its intended use and define the performance requirements. The latter is defined based on assumptions regarding the main features of the product. It is necessary to ensure that the concept reflects the needs of the end-user, as well as fulfil business goals.

The requirements should be defined prior to design of the product; as this allows to identify the resources needed throughout the development process (Rahimi & Rausand, 2015). With regards to safety, "the requirements are generated by the end-user and are appropriate for the intended application" (Drogoul et al., 2007). These requirements are further assessed and redefined once the design process commences. The producer must establish the overall strategies and frameworks that will be used throughout the process, define the functional safety management plan, develop procedures and guidelines for the product development process containing verification and validation activities (Lundteigen et al., 2009). During this phase, the producer must also identify which requirements are dependent on the SIL rating and which ones are generic, define roles and responsibilities for different phases of the project and set in a quality management system and other documents if they are not already available. The type of requirements is also dependent on the type of system being produced. For full SIS applications the work is more extensive than for SIS components. Where for the later the producer is mainly concerned with the systematic capability.

It is also beneficial to evaluate experiences from previous processes and products, such as failure rates and hazards. At this point using qualitative or semi-quantitative risk assessments can help identifying possible hazards that require further analysis once more detailed knowledge about the system becomes available (such architecture, components, properties, etc.). Uncertainties regarding the system must be documented and resolved throughout the design process. A list containing critical items can be elaborated (Lundteigen et al., 2009). Contacting potential customers can help identifying issues with similar products and provide inputs regarding the concept. Workshops can be organized to review the concept and identify costumer needs, as well as to promote understanding related to the operational

and environmental conditions that the product may endure. Furthermore, deviations from the intended application can be identified and measures to correct them can be planned. Phase 1 can be summarized as presented in Figure 9. Once the concept is approved the producer moves on to the next phase of the framework.

Figure 9: Definition of product requirements

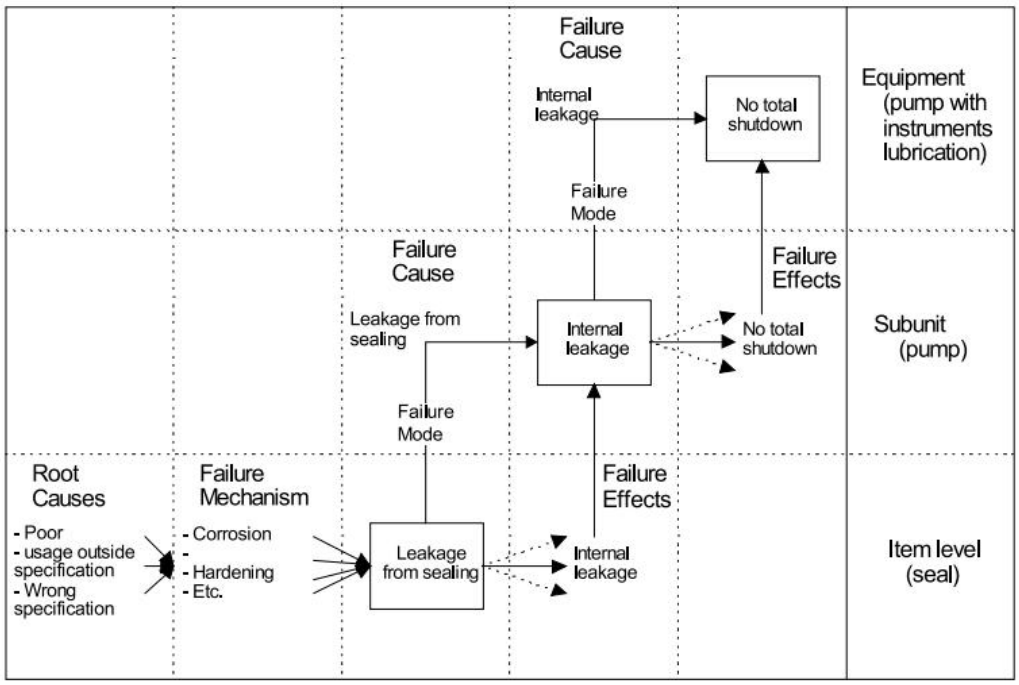


4.1.2. Phase 2: Design

4.1.2.1. System design

In this phase the preliminary design is elaborated. The main objective of Phase 2 can be understood as the effort to transform the requirements into the physical product. The main tasks in this step are related to ensuring that the system design is appropriate and potential hazards have been identified and corrected. This is achieved through the combination of design reviews and QRAs which are performed iteratively. As described previously, a common practice is to perform FMEA and/or FMEDA to identify the system's failure modes and use the results as an input for preliminary reliability calculations. Rahimi and Rausand (2015) also suggest performing HAZOP analysis for complex systems as an aid to identify deviations from the design intent and examine their causes and consequences. Furthermore, it is important to investigate the underlying factors that may cause the failures. In addition to that, it is suggested to evaluate the knowledge to which the failure modes are based on, especially in cases where real feedback data is not available. This can assist the design team prioritizing treatment strategies and reflect if more work needs to be done to investigate the possible mechanisms that lead to the specific failure mode (Brissaud et al., 2010). Hence, allowing more efficient risk management. A systematic approach for identification of possible failure modes and their related failure mechanisms is depicted in Figure 10. The idea is to investigate the relationship between a root cause failure mechanism in a hardware assembly and its consequences.

Figure 10: Relationship between failure, failure causes and effects



Retrieved from: DNVGL-RP-A203, 2017

Once the design team has concluded the FMEDA, identified and corrected all the possible failures by means of changing the system design, reliability analysis is performed to verify if the system is able to meet the performance requirements taking into account the allocated SIL. Performing uncertainty assessment in this step is highly recommended. A sensitivity analysis or importance measures can assist the design team identifying the most critical components. The first technique investigates how changes in one input parameter affects the output, whereas importance measures determine how changes in reliability of specific components will affect the overall system reliability. These analyses can be used to determine the appropriate reliability of the system's components and to propose system modifications and follow-up activities. The critical items list is then updated with the new information generated by the analyses.

In addition to focusing on random hardware failures, the producer must also enforce and ensure that working procedures to avoid systematic failures have been followed and all assumptions, changes and conflicts have been identified and documented, and that the hardware architecture meets the architectural constraints.

4.1.2.2. Detailed design

During this phase the producer will elaborate the design specification of the product. This document contains a detailed description of all the components and their properties, such as diagnostic capability and fault behavior to name a few (Lundteigen et al., 2009).

This can be done following the same approach as for the system design, by identifying each component failure modes and updating the reliability assessments with the information gathered in this step. The producer must also ensure that the interactions between the components are appropriate (via compatibility[3] and interface analysis[4]) (Rahimi & Rausand, 2015). Thereafter, a new design review is conducted to ensure the system is able to deliver the intended functions and no modifications are needed.

Another important task to be covered in this step, is performing quality assessment to evaluate components and parts purchased from subcontractors. This is done to assure items must provide the necessary performance as established in the design specification requirements. Furthermore, plans concerning the assembly of components should be elaborated and attention should be given to avoidance of hazards during the construction, installation, and testing. The results of the analysis are used to develop initial versions of instructions manual – installation, operation, and maintenance manuals. If the performance is judged adequate the producer moves to the next phase of the framework.

4.1.3. Phase 3: Development

Phase 3 refers to the physical product, where the development of product prototype takes place. This phase initiates the product qualification process, where main objective is to test the prototype to against the performance requirements. Testing the prototype may reveal failure modes that were not identified during the design phase. For new technology the qualification requires more effort than for proven technology. This is given the uncertainty associated with the technology, which will present higher margins to meet the functional specification. Furthermore, the development phase can be subdivided into two sub-categories: prototype qualification I, starting at a component level to prototype qualification II regarding the overall system.

---

[3] The ability of two or more components to perform their functions when sharing the same hardware or software

[4] Relation between the elements of a system

4.1.3.1. Prototype qualification I

This activity is related to component testing against the desired performance. Here, the components are tested in controlled environments according to their criticality and the uncertainty related to them. The term component in this phase can be used to designate individual parts of the system, or sub-systems depending on the complexity of the final product. When dealing with new technology, working procedures related to technology qualification must also be initiated. For instance, following the guidelines proposed by DNVGL-RP-A203 (2017).

- Verify if the requirements for the components are met
- Perform function testing for the components
- Document hazards and analyse if new hazards have been identified
- Update the reliability analyses with new information
- Decide if the performance is satisfactory before moving on to the next phase

The challenges associated with this activity may include: simulating realistic environments and determining the number of tests required to demonstrate reliability (Rahimi & Rausand, 2015).

4.1.3.2. Prototype qualification II

The main purpose of this activity is to investigate the effects of the interactions between components and ensure that they are combined properly. The prototype assembly is tested to verify if it has the required quality, which consists of operational testing. The information derived from testing the system in operational environment provides the producer insights regarding necessary design changes. Depending on the application of the system, it is not possible to test the prototype in a real environment. For instance, for subsea equipment, virtual testing is performed instead of operational testing. This consists of performing the tests simulate the operational conditions, for example keeping components submerged in a pool or under exposure of high temperatures and pressure (Lundteigen et al., 2009).

- Verify quality of the whole product
- Verify if the product achieves the required performance
- Identify if there are any problems between components
- Document hazards and analyse if new hazards have been identified
- Update the reliability analyses with new information
- Decide if the performance is satisfactory before moving on to the next phase

4.1.4. Phase 4: Production

The construction phase involves the production of the product. At this stage the producer is concerned in ensuring that the performance of the product matches the established criteria. It is also important to certify that the production process does not introduce new failures to the system. This is achieved by:

- Quality assurance of manufacturing process
- Quality assurance of the product
- Verification of performance requirements
- Factory and Site acceptance test

4.1.5. Phase 5: Installation, commissioning, and operation

In this phase the product is put into operation. The producer may or may not be involved in the installation, commissioning, and operation of the product. However, cross-company cooperation is highly recommended, as it is the producer's job to:

- Monitor the product performance
- Collect data regarding the product performance to check if the predicted performance is adequate
- Identify the need for changes or improvements
- Update controlling documents with information related to failures and new failure modes.
- Update the reliability estimation with actual field data

4.1.6. Phase 6: Monitoring, review and lessons learnt

This phase is executed from a business perspective. The process is evaluated with regards to costs, profits from sales, impacts of failures or inadequate performance on business reputation, etc. The producer must register the experience from this phase and use it to promote organizational learning for the next product generation (Lundteigen et al., 2009).

4.2. Challenges regarding safety demonstration of new technology

As stated previously, the use of RAC and associated requirements are only acceptable for situations involving a strong knowledge base and small uncertainties, and that holds for most technical systems in the industry where there is experience and are based on known technology. For situations involving new technology, however, such requirements might not be suitable. In these cases, new safety philosophies might be needed as the traditional practice may not offer support for demonstrating safety of such concepts (Zikrullah et al, 2019 and DNV-GL, 2021). For instance, the IEC61508 (2010) allow the users to establish their own criteria and technical solutions necessary to meet the criteria. This enables creativity in design, as the standard only prescribes the safe attributes of the system. For example, one attribute is based on independency between safety functions. The user is then allowed to propose their own solutions, given that it meets this principle. However, in practice, fully independent systems are difficult to achieve and the level of independence between them is not entirely known. Some degree of dependence resulting from logical and physical interactions or due to increased digitalisation should be expected (DNV-GL, 2021). For example, Zikrullah et al. (2019) describe that the concept involving integration of process control and safety (IPC&S) systems[5] generate problems when formulating requirements for detailed architecture, as the current approaches focus on models that do not investigate the interaction between components/systems. Furthermore, the concept violates the requirements of independence between systems. The authors also present two solutions (i) development of a new approach that considers dependency between all protection layers and (ii) propose a method to quantify the dependency between both systems. In addition, the requirements should focus on providing safe implementation of systems that present dependencies and should clarify which situations independency can be relaxed and when it should be met (DNV-GL, 2021). This would aid SIS designers as the current approaches for safety demonstration might be cumbersome, or even limiting the amount of innovation proposed.

[5] IPC&S system: a concept where the process control system (PCS) and process shutdown (PSD) are performed by the same hardware, with functional separation of software.

5. CONCLUSIONS AND FUTURE WORK

This thesis has presented a review of the safety requirements used during the design of safety-related systems as well as a framework to aid this process. The most important standards used in the NCS are the IEC61508 (2010), IEC61511 (2016) and the NOG-070 (2020). Those standards use the safety lifecycle framework to structure the requirements for designing, installing, operating, and maintaining safety-related systems. However, the most important standard used to design these systems is the IEC61508 (2010), which relies on a risk-based approach. Even though the standard provides a general approach to demonstrate safety, it was argued in this work that the approach relies in an incomplete representation of risk, as it does not encompass uncertainty and knowledge as part of the risk description. This can lead to misrepresentation of risk, which might affect the quality of the final product.

After reviewing the requirements, some challenges regarding the traditional methods to allocate and verify SIL were highlighted, such as the way RAC and QRAs are used. Risk acceptance criteria and QRAs serve as the basis to establishing the desired performance of the system, and further in the design process new QRAs are conducted to assure the product meets the desired reliability targets. It was presented that in the Norwegian Oil and Gas industry, however, there is an extensive knowledge regarding most processes. Hence, the use of RAC to establish the performance requirements can be justified. The problem seems to arise when QRAs are used to predict the performance of the systems, as the traditional approach does not give much weight to uncertainty and knowledge as suggested by the PSA in their management regulations. In fact, uncertainty assessments are very limited or non-existent. Generally, the focus lies on the statistical uncertainty regarding the models and data used to predict the PFD. Using this approach to measure the performance of a system is particularly challenging when it involves new technology, as there is not much data and information related to the actual performance of the system. Nevertheless, some failure modes only become apparent after the system is put into operation. Thus, the focus should be in designing robust/resilient systems, so they can withstand hazards accordingly. Furthermore, this approach can lead to safety being demonstrated in a mechanical way, as the focus would be in meeting numerical targets instead of focusing on improvements.

It is, therefore, suggested that due diligence on the background knowledge supporting the analysis is performed, and that the results are documented and properly transferred. Risk is more than probabilities and expected values, and aspects such as uncertainty and knowledge must be evaluated. Technical and non-technical aspects need to be addressed in order to design systems appropriately and ensure they will operate correctly.

Another problem with the safety lifecycle is that it is not as detailed from the product development perspective. In view of the limitations presented earlier in the text, a framework was elaborated to help SIS producers with the task of demonstrating safety according to the IEC61508 (2010), while giving more attention to uncertainty. The framework covers the whole product development process and suggests ways of increasing knowledge throughout the design such as investigating the underlying factors that may cause the failures. It is proposed to conduct a root causes of the failure modes, so the design team can prioritize treatment strategies. Furthermore, it allows designers to reflect if more work needs to be done to investigate the possible mechanisms that lead to a specific failure mode. This results in more efficient risk management throughout the process. Additionally, the design team must focus on revealing weaknesses earlier in the design, so corrective actions can be taken as soon possible. A deeper understanding of the operational conditions and previous experiences may provide guidance on both technical and non-technical aspects that must be considered throughout the design and offer a more concreate starting point.

Finally, it is important to stress the challenges related to safety demonstration of new technology not only considering the uncertainty related to predicting the performance due to lack of data and experience, but also due to the fact that new concepts might violate some established design principles, such as independence between safety functions. As technology is evolving in a fast pace, new safety philosophies might be needed to predict and verify the performance of concepts that are not covered by the standard approach. For instance, due to increasing digitalisation some level of dependence may be present, and new approaches to quantify it are made necessary. Therefore, new methods that consider dependency between systems are highly recommended, as mentioned by Zikrullah et al. (2019). Further work should focus on creating mechanisms to quantify and deal these challenges as well as elucidating which cases independency must be met and which cases it can be relaxed.

REFERENCES

Abrahamsen, E. B., Abrahamsen, H. B., Milazzo, M. F., & Selvik, J. T. (2018). Using the ALARP principle for safety management in the energy production sector of chemical industry. *Reliability Engineering & System Safety, 169*, 160-165. https://doi.org/https://doi.org/10.1016/j.ress.2017.08.014

Abrahamsen, E. B., & Aven, T. (2012). Why risk acceptance criteria need to be defined by the authorities and not the industry? *Reliability Engineering & System Safety, 105*, 47-50. https://doi.org/https://doi.org/10.1016/j.ress.2011.11.004

Abrahamsen, H., & Abrahamsen, E. (2015). On the appropriateness of using the ALARP principle in safety management.

Aven, T. (2013). Practical implications of the new risk perspectives. *Reliability Engineering & System Safety, 115*, 136-145. https://doi.org/https://doi.org/10.1016/j.ress.2013.02.020

Aven, T. (2014). *Risk, surprises and black swans: Fundamental ideas and concepts in risk assessment and risk management*. Routledge.

Aven, T., & Ylönen, M. (2016). Safety regulations: Implications of the new risk perspectives. *Reliability Engineering & System Safety, 149*, 164-171. https://doi.org/https://doi.org/10.1016/j.ress.2016.01.007

Bai, Y., & Jin, W.-L. (2016). Chapter 38 - Risk Assessment Methodology. In Y. Bai & W.-L. Jin (Eds.), *Marine Structural Design (Second Edition)* (pp. 709-723). Butterworth-Heinemann. https://doi.org/https://doi.org/10.1016/B978-0-08-099997-5.00038-1

Brissaud, F., Charpentier, D., Fouladirad, M., Barros, A., & Bérenguer, C. (2010). Failure rate evaluation with influencing factors. *Journal of Loss Prevention in the Process Industries, 23*(2), 187-193. https://doi.org/https://doi.org/10.1016/j.jlp.2009.07.013

Catelani, M., Ciani, L., & Luongo, V. (2010). The FMEDA approach to improve the safety assessment according to the IEC61508. *Microelectronics Reliability, 50*(9), 1230-1235. https://doi.org/https://doi.org/10.1016/j.microrel.2010.07.121

Chang, K., Kim, S., Chang, D., Ahn, J., & Zio, E. (2015). Uncertainty analysis for target SIL determination in the offshore industry. *Journal of Loss Prevention in the Process Industries, 34*, 151-162. https://doi.org/https://doi.org/10.1016/j.jlp.2015.01.030

Coglianese, C., Nash, J., & Olmstead, T. (2003). Performance-based regulation: Prospects and limitations in health, safety, and environmental protection. *Admin. L. Rev., 55*, 705.

Creech, G. (2014). IEC 61508 Systematic Capability. *Measurement and control, 47*(4), 125-128. https://doi.org/10.1177/0020294014528895

Dean, S. (1999). IEC 61508 - Understanding Functional Safety Assessment. *Measurement and control, 32*(7), 201-204.

DNV-GL. (2021). *Safety 4.0 - Demonstrating safety of novel subsea technologies*. Retrieved 03/05/2021 from https://www.dnv.com/research/oil-gas/safety40/project-description.html

DNVGL-RP-A203. (2017). *Technology qualification*.

Drogoul, F., Kinnersly, S., Roelen, A., & Kirwan, B. (2007). Safety in design – Can one industry learn from another? *Safety Science, 45*(1), 129-153. https://doi.org/https://doi.org/10.1016/j.ssci.2006.08.004

Gabriel, A., Ozansoy, C., & Shi, J. (2018). Developments in SIL determination and calculation. *Reliability Engineering & System Safety, 177*, 148-161. https://doi.org/https://doi.org/10.1016/j.ress.2018.04.028

Huang, G., Xiao, L., Zhang, W., Li, J., Zhang, G., & Ran, Y. (2020). An improving approach for failure mode and effect analysis under uncertainty environment: A case study of critical function component. *Quality and Reliability Engineering International, 36*(6), 2119-2145. https://doi.org/https://doi.org/10.1002/qre.2686

IEC61508. (2010). *Functional safety of electrical/electronic/programmable electronic safety-related systems. Part 1-7*.

IEC61511. (2016). *Functional safety - Safety instrumented systems for the process industry sector*.

Johansen, I. L., & Rausand, M. (2015). Barrier management in the offshore oil and gas industry. *Journal of Loss Prevention in the Process Industries, 34*, 49-55. https://doi.org/https://doi.org/10.1016/j.jlp.2015.01.023

Liu, Y. (2020). Safety barriers: Research advances and new thoughts on theory, engineering and management. *Journal of Loss Prevention in the Process Industries, 67*, 104260. https://doi.org/https://doi.org/10.1016/j.jlp.2020.104260

Lundteigen, M. (2008). *Safety instrumented systems in the oil and gas industry: Concepts and methods for safety and reliability assessments in design and operation* [PhD thesis,

Lundteigen, M., & Rausand, M. (2006). *Assessment of Hardware Safety Integrity Requirements*.

Lundteigen, M. A., & Rausand, M. (2007). Common cause failures in safety instrumented systems on oil and gas installations: Implementing defense measures through function testing. *Journal of Loss Prevention in the Process Industries, 20*(3), 218-229. https://doi.org/https://doi.org/10.1016/j.jlp.2007.03.007

Lundteigen, M. A., Rausand, M., & Utne, I. B. (2009). Integrating RAMS engineering and management with the safety life cycle of IEC 61508. *Reliability Engineering & System Safety, 94*(12), 1894-1903. https://doi.org/https://doi.org/10.1016/j.ress.2009.06.005

Murthy, D. P., Rausand, M., & Østerås, T. (2008). *Product reliability: specification and performance*. Springer Science & Business Media.

NOG-070. (2020). Norwegian oil and gas application of IEC61508 and IEC61511 in the Norwegian petroleum industry (Recommended SIL).

NORSOK-Z013. (2010). *Risk and emergency preparedness assessment*

Osteras, T., Murthy, D. N. P., & Rausand, M. (2006). Product performance and specification in new product development. *Journal of Engineering Design, 17*(2), 177-192. https://doi.org/10.1080/09544820500275735

PSA. (2011). *The framework regulations - § 11 Risk reduction principles*. https://www.ptil.no/en/regulations/all-acts/the-framework-regulations3/II/11/?expandGuideline=true&hideParagraph=true

PSA. (2017). *Black swans*. https://www.norskoljeoggass.no/contentassets/d3183372438841a180e14938177f6ec7/black-swans.pdf

PSA. (2018). Integrated and unified risk management in the petroleum industry. https://www.ptil.no/contentassets/8d93722526cb4c57a5068e680be90a7b/risikostyring-2018-engelsk.pdf

PSA. (2020). Bruk av risikoakseptkriterier – en evaluering. https://www.ptil.no/contentassets/4deea346d8cb4008a2eef488f85313ae/bruk-av-risikoakseptkriterier---en-evaluering.pdf

Rahimi, M., & Rausand, M. (2015). Technology Qualification Program Integrated with Product Development Process. *Int J Performability Eng, 11*(1), 3-1. https://doi.org/10.23940/ijpe.15.1.p3.mag

Rausand, M. (2014). *Reliability of Safety-Critical Systems: Theory and Applications* (Vol. 9781118112724). Somerset: John Wiley & Sons, Incorporated. https://doi.org/10.1002/9781118776353

SINTEF. (2017). *What is the PDS method?* Retrieved 10/06/2021 from https://www.sintef.no/projectweb/pds-main-page/the-pds-method/what-is-the-pds-method/

van de Poel, I., & Robaey, Z. (2017). Safe-by-Design: from Safety to Responsibility. *NanoEthics, 11*(3), 297-306. https://doi.org/10.1007/s11569-017-0301-x

Zikrullah, N., Kim, H., Lundteigen, M., & van der Meulen, M. (2019). *Clarifying Implementation of Safe Design Principles in IEC 61508: Challenges of Novel Subsea Technology Development*. https://doi.org/10.3850/978-981-11-2724-3_0112-cd

APPENDIX

The new risk perspectives used in the Norwegian Continental Shelf

New risk perspectives focus on the importance of highlighting uncertainty rather than adopting a probabilistic-based perspective which is generally seen in risk assessment and risk management approaches applied in industrial activities.

In Norway, the PSA has adopted these perspectives and defines risk as *"the consequences of the activities, with associated uncertainty"* (PSA, 2011). Where uncertainty is related to the potential consequences of the activities – which are not known. The PSA (2011) also recognises that the risk associated with the activities is dependent of the context that is being evaluated and the information base considered. Therefore, when using risk assessments to support decision-making, dependencies of the background knowledge shall be contemplated. This is, once more, consistent with current thinking among risk science community, where knowledge (or lack of) is seen as crucial element when describing risk and its characterization can provide useful insights about potential for surprises/black swans (Aven, 2013).

In contrast, in the Norwegian Continental Shelf (NCS), risk-based thinking is still dominant, and it is used as the basis for elaborating regulations, standards and requirements. Where emphasis is given to the use of risk assessments, risk acceptance criteria and tolerability limits (Aven & Ylönen, 2016). The focus lies on risk reduction, and it is translated into risk acceptance criteria (RAC).

Essentially, the risk acceptance criteria are used in the petroleum sector to support decision-making. These criteria highlight features such as keeping the risk to a level that is As Low as Reasonably Practicable (ALARP principle), in accordance to safety regulations. If the criteria are not met, risk reducing measures need to be implemented (Abrahamsen & Aven, 2012). To define if the RAC are achieved, various quantitative risk assessments

(QRAs) are performed. Hence, the results of QRAs are then used to determine if additional measures are needed.

For practical execution of activities, pre-determined criteria simplify the decision-making process. In this context, having concrete reference values is very useful to judge if the risk is acceptable or not at any phase of a project. However, this can lead the process to be conducted in a mechanistic way, resulting in the need for a change in the risk perspective to which the criteria are used.

One reason is that RAC used in the industry are expressed as minimum safety requirements, which may seem contradictory considering the ALARP principle highlighted by safety regulations. The contradiction is caused by the fact that attention is often placed on achieving the criteria and not on following adequate risk reduction processes. Which, in turn, is related to the industry having autonomy to decide themselves what are the acceptance limits. Such practices might result in weak formulation of the limits that risk can be considered acceptable. In fact, weak limits are only justifiable if it can be demonstrated that the benefits are sufficiently large. Otherwise, more stringent criteria need to be adopted.

Another reason is related to the limitations of the risk assessment tools used to determine the criteria and its achievement. Probabilistic approaches are not enough to describe risk. The background knowledge is equally important and needs to be considered (Aven, 2014).

To summarize, risk-based perspectives do not capture the overall essence of risk. The two main contributing factors are related to:

1. The use of risk acceptance criteria
2. The tools used to determine if the criteria are met.

Regarding the RAC, it is acknowledged that it may give the wrong focus, where companies are too concerned in meeting the criteria instead of finding the best solutions. As for the risk assessment, the tools used (here being understood as the QRAs), to large extent, were developed in the 70s and 80s and did not experienced any major changes since they were firstly introduced. These tools lack the level of precision necessary for using of mechanical criteria (Aven & Ylönen, 2016). Risk analysis has limitations and does not capture all aspects of risk. Thus, there is a need for an approach that reflects uncertainty and knowledge.(Aven, 2014). According to Aven and Ylönen (2016), the analysis must cover not only the probabilistic assessment but also the following items:

- The knowledge on which the probabilities are based
- The strength of knowledge
- Assumptions and risks related to the deviations of these
- Surprises relative to one's knowledge and beliefs

The impact of the new risk perspectives on the way risk acceptance criteria are used in the Norwegian continental shelf

As mentioned in the previous section, the PSA (2017) has adopted the new risk perspectives, and suggests that adjustments are to be made regarding the risk acceptance criteria. Risk acceptance shall be reviewed in relation to knowledge and uncertainty. In this manner, assessment of background knowledge supporting the probabilistic assessment needs to be reflected. According to Aven (2014) and the PSA (2017), the following items can be used to deliberate if risk is, in fact, acceptable:

1. If the strength of knowledge is not weak, and risk is found to be acceptable in relation to probabilities with large margins, the risk is judged acceptable.
2. If the strength of knowledge is strong, and risk is found to be acceptable in relation to probabilities, risk is considered acceptable.
3. If the strength of knowledge is not strong, and risk is found to be acceptable in relation to probabilities with small or moderate margins, risk is unacceptable, and measures are required to reduce risk
4. If risk is found to be unacceptable in relation to probabilities, the risk is unacceptable, and measures are required to reduce risk.

The above points are elaborated further by the PSA (2020) in the report "The use of risk acceptance criteria – an evaluation" *("Bruk av risikoakseptkriterier – en evaluering").*

According to this report, for situations where the knowledge base is very strong, for instance, cases involving known technology, ideas and problems, risk acceptance criteria can be replaced by more specific requirements, such as solutions tailored for such situations. Additionally, qualitative risk analyses can be used to identify potential surprises. In this case, the desired level of safety can be achieved while simplifying risk management. However, for situations that are not "standard", the uncertainties are higher and "broader risk assessments" are required, as the knowledge is relatively weak. The use of quantitative RAC alone do not provide good decision support as only one aspect of risk is reflected in the risk picture. Its usage is, therefore, weakly justified. Lastly, when the problems are new, or for situations with little experience or for new solutions, the uncertainties are large, it is necessary to strengthen the knowledge base, while giving emphasis to robustness/resilience. For situations where uncertainties are large, quantitative risk descriptions and associated RAC may not provide useful insights to managing risk in a good way.