University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study programme: Societal safety and security	Spring semester, 2021 Open
Author: Marita Pytte	
Programme coordinator: Morten Sommer Supervisor: Morten Sommer	
Title of master's thesis: Learning potential of accident investigation reports from Norway and the United States' national investigation boards.	
Credits: 30	
Keywords: Accident investigation; Learning from accidents; Investigation expertise; Investigation methods; Accident models; Accident methods; National investigation boards; Cross-continental comparison	Number of pages: 64 + supplemental material/other: 85 Stavanger, 23.06.2021

Acknowledgements

This master's thesis is the end product after two years of studying Societal Safety at the University of Stavanger. The choice of topic stood between a variety of interesting and engaging issues, but I am very happy I chose to deep-dive further into accident investigations. The interest awoke after taking the subject Investigation Methods (Granskingsmetodikk) in the third semester, and it sparked enthusiasm in me. Particularly learning from accidents is something that appears to still need more research and something we have the potential to improve on to save more lives through mitigating measures.

I would like to take the opportunity to thank everyone that has helped me throughout the process of writing this thesis. First, I want to thank my supervisor, Morten Sommer, for encouragement and constructive feedback. You were always patient, as well as being quick with responses and feedback. I appreciate all the time you took to read through my different drafts and helping me make the final product better. I also want to thank you for sparking my interest in the first place, the way you shared your enthusiasm for accident investigations through the Investigation Methods subject.

Secondly, I would like to thank Gunnar Haugland from Safetec for an interesting discussion early on in the project. It helped me adjust the aim of my thesis and gave me ideas for interesting dilemmas and talking points to discuss. So early on I felt a little "lost" and our conversations made me get back on track with new energy.

I also want to thank my friends Maria and Benjamin in particular, for keeping me company, motivating me, and cheering me on every single day, and having faith in me when I doubted myself. Last but not least, I want to thank my family, for always being there for me and loving me unconditionally, as well as providing me with a warm and safe space to come home to whenever I need it.

You have all been invaluable in this process, this thesis would not have been the same without every single one of you.

Marita Pytte

Randaberg, June 2020

Summary

Accident investigations are an important tool for learning from accidents, by giving insight into what went wrong, and how to prevent similar accidents from happening in the future. However, there is a lack of standard guidance regarding best practices for how to execute such investigations. There are a plethora of different aspects to these investigations, as well as different alternatives and approaches for each aspect. The purpose of this thesis is to investigate how a few of these aspects might influence the outcome of the investigations, and how this in turn can influence learning potential.

To answer this, the objective of this study was to find out what influence standard procedures, investigation methods, expertise, and accident models have on conclusions and safety recommendations. The study has focused on two accident boards from two different countries: the Norwegian Safety Investigation Authority (NSIA) from Norway and the National Transportation Safety Board (NTSB) from the United States. They were chosen to highlight potential cross-continental differences that could give valuable insight that both sides can learn from.

The research in this thesis was done mainly through comparative document analyses, both a quantitative and qualitative approach. The qualitative analysis was used to gain insight into the fundamental differences between the two boards and entailed official documents from the two boards, as well as relevant laws and regulations regarding their mandate and responsibilities. This was in turn used as potential explanations for differences found through the quantitative analyses. The quantitative analyses included railway accident reports from the last decade and investigated the effect the different aspects had on the report's focus, safety recommendations, socio-technical levels, and conclusions.

The data collection, hypotheses and conclusions were all done in light of the theoretical context of the thesis. Throughout the thesis, I have relied on Hollnagel's three major groups of accident models consisting of sequential, epidemiological, and systemic models; Rasmussen's socio-technical levels; and Argyris and Schön's distinction between double-loop and single-loop learning. Furthermore, I chose to focus on the three accident models AcciMap, MTO and STEP. This theoretical context must be considered when reviewing the results and conclusions made in this thesis.

The results from the analyses implied that different standard procedures, expertise, and accident models all have a significant effect on the outcome of investigations and may influence learning potential. A larger cross-sectoral investigation board with more diverse expertise appears to give a more even distribution of safety recommendations between the different socio-technical levels, and more focus to higher levels. This could provide learning to more of the levels, as well as increasing double-loop learning by including higher levels.

The results also indicate that epidemiological accident models give an increased focus on human causal factors, while systemic accident models lead to an increased focus on organizational factors. Lastly, investigation methods did not have any apparent effect, possibly due to the two accident boards having similar approaches and overlap. Both the NSIA and the NTSB have different aspects that can influence their learning potential in positive and negative ways, and it would be beneficial for both of them to learn from each other's approaches to increase their learning from accidents.

Sammendrag

Ulykkesgransking er et viktig verktøy for å lære av ulykker, ved å gi innsikt i hva som gikk galt og hvordan man kan forhindre liknende ulykker fra å skje igjen. Det er derimot en mangel på standard prosedyrer og beste praksis når det gjelder hvordan ulykkesgranskinger bør utføres. Det er mange forskjellige aspekt ved slike granskinger, i tillegg til at det er ulike alternativer og tilnærminger til hvert av aspektene. Formålet med denne oppgaven er å undersøke hvordan noen av disse aspektene kan påvirke utfallet granskingene har, og hvordan dette igjen kan påvirke læringspotensialet.

For å svare på dette har det vært et mål å finne ut hvordan standard prosedyrer, granskingsmetodikk, ekspertise og ulykkesmodeller kan påvirke konklusjoner og sikkerhetstilrådninger. Studien har fokusert på to granskingskommisjoner fra to forskjellige land: Statens Havarikommisjon fra Norge (SHK) og the National Transportation Safety Board (NTSB) fra USA. De ble valgt for å sette lys på mulige forskjeller på tvers av kontinent, og gi innsikt som begge sider kan lære fra.

Undersøkelsene i denne oppgaven er gjort hovedsakelig gjennom dokumentanalyser, både kvantitative og kvalitative. Den kvalitative analysen ble brukt for å få innsikt i fundamentale forskjeller mellom de to granskingskommisjonene, og besto av offisielle dokumenter fra de to kommisjonene i tillegg til relevante lover og regler som omhandlet deres mandat og ansvar. Resultatene av dette ble igjen brukt som mulige forklaringer på forskjeller funnet i de kvantitative analysene. De kvantitative analysene inkluderte rapporter etter jernbane-ulykker fra det siste tiåret, og undersøkte effekten de forskjellige aspektene hadde på rapportenes fokus, sikkerhetstilrådninger, sosio-tekniske nivå og konklusjoner.

Datainnsamlingen, hypotesene og konklusjonene ble utformet i lys av det teoretiske rammeverket brukt i oppgaven. Gjennom oppgaven har jeg fokusert på Hollnagels tre hovedgrupper for ulykkesmodeller som involverte sekvensielle, epidemiologiske og systemiske modeller; Rasmussens sosio-tekniske nivå; og Argyris og Schöns definisjon på dobbelkretslæring og enkelkretslæring. I tillegg valgte jeg å fokusere på de tre kjente ulykkesmodellene AcciMap, MTO og STEP. Dette teoretiske rammeverket må tas i betraktning med tanke på resultatene og konklusjonene trukket i denne oppgaven.

Resultatene fra analysene indikerer at forskjellige standard prosedyrer, ekspertise, og ulykkesmodeller alle har en signifikant effekt på utfallet av ulykkesgranskinger, som igjen kan påvirke læringspotensialet. En større granskingskommisjon som inkluderer flere sektorer og ulik ekspertise ser ut til å bidra til en mer jevn distribusjon av sikkerhetstilnærminger på de ulike sosio-tekniske nivåene. I tillegg ga det også mer fokus på de øverste nivåene. Dette kan bidra til læring på flere av nivåene, i tillegg til økt dobbelkretslæring ved å inkludere høyere nivå.

Resultatene tydet også på at epidemiologiske ulykkesmodeller gir økt fokus på menneskelige årsaksfaktorer, mens systemiske ulykkesmodeller fører til høyere fokus på organisatoriske faktorer. Granskingsmetodikk så derimot ikke ut til å ha noen signifikant effekt, muligens grunnet at de to granskingskommisjonene benyttet overlappende granskingsmetoder uten store nok forskjeller. Både SHK og NTSB viste seg å ha forskjellige aspekter som kan påvirke læringspotensialet i både positive og negative retninger, og de kan begge ha nytte av å lære av hverandres tilnærminger på ulike områder.

Content

List of figures

List of tables

1. Introduction

1.1 Background

It has become a common worldwide practice to execute accident investigations in all types of fields. One can investigate criminal acts, natural disasters, health care, fraud, transportation accidents and much more. An accident investigation can be defined as “the collection and examination of facts related to an occurred specific event” (Harms-Ringdahl, 2004, p. 14). Accident investigations differ from criminal investigations, as they are not out to find evidence to convict a perpetrator. Instead, they intend to promote learning from accidents and prevent future accidents from happening or lessening the negative consequences. While there are differentiations between accidents, incidents, disasters, events and so on, ‘accidents’ is defined in this thesis as “an unexpected, unwanted chain of events, with consequences on health, safety and environment or equipment damages” (Dechy, et al. 2012, p. 1382).

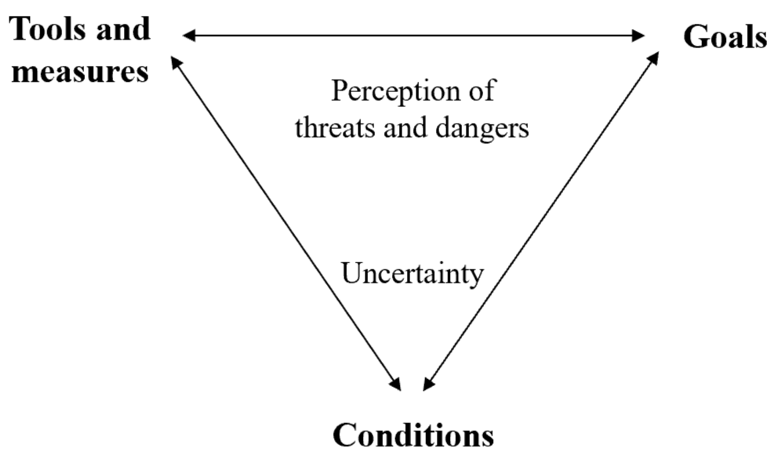
1.1.1 Accident investigations and safety management

During an accident investigation, it is relevant to examine the company’s safety management to gain insight into possible causal factors. Safety management can be defined as “all measures implemented to achieve, maintain, and further develop a security level in line with defined goals” (Njå, et al., 2020, p. 65). Despite not giving any indication of what level of safety one should aim for, the definition illustrates an important relationship between safety measures and goals. Njå et al. (2020) also point out how both the goals and safety measures are influenced by a framework of conditions, such as available resources, environmental factors, and social and cultural influences. Challenging conditions, as well as any lack of sufficient goals or safety measures, can all be contributing factors in an accident trajectory.

Figure 1 shows the relationship between the three factors at play in safety management, in addition to the central role of the actors’ risk perception and the uncertainty factor. Different tools and safety measures are used to reach goals and visions regarding safety and security. Furthermore, both the goals and the safety measures are influenced by the relevant conditions in play. These conditions are usually something the involved actors do not have much power to influence, at least not in a reasonable amount of time (Njå et al., 2020).

Figure 1

Safety management model. Adapted from Njå et al. (2020)



Accident investigations do not only investigate safety management but are an important part of the safety management itself. Knowledge regarding how accidents happen can increase efficient safety management and can be considered a measure that helps both achieve and maintain a sufficient security level. Hovden et al. (2004) explain that accident investigations contribute to more efficient safety management by monitoring the condition and trends in priorities, resources, and security measures, as well as contributing to modelling an understanding of causal factors and accident mechanisms. Accident investigations also create motivation, draw attention and interest towards risk mitigation, and influence attitudes towards a positive safety culture.

Safety measures are an important aspect of safety management, and the outcome of accident investigations often consists of making safety recommendations regarding implementing new measures or fixing existing ones that are not working sufficiently. This is one of the main ways to find out whether implemented measures work, whether they indeed do prevent accidents from happening and whether current safety levels and goals are sufficient.

1.1.2 Approaches to accident investigations

Despite widespread use, there is no agreed-upon “best way” to execute accident investigations, even within the same sector. This has resulted in investigations differing in many aspects, such as accident theory, methods, and expertise. Sklet (2004) listed 14 main categories of methods, but there are many more, especially when considering that many companies have developed their own methods (Roed-Larsen & Stoop, 2012).

Earlier studies have examined and compared different accident methodologies. In 1985, Benner rated and evaluated 14 accident models and 17 accident investigation methodologies used by 17 different government agencies. He made a rating system based on how well the different models and methods fulfilled different criteria. After ranking the methods and models, they were used to reinvestigate earlier accidents to see if differently ranked models would produce different findings. The results were overwhelmingly in favour of the higher-ranking methods and models, despite two years having passed since the accident. In fact, the original investigations using lower-ranking methods had excluded important data from the accident.

Sklet (2004) used a similar approach while evaluating his 14 main categories of investigation methods. He evaluated whether each method fulfilled different criteria but did not rank the results. Instead, he presented a table with information on what each method focused on and included. He concluded that each method has different areas of application and unique positive and negative sides, useful during different parts of the investigation process. Filho et al. (2019) support this idea of combining methods. They investigated the same accident with two different analysis approaches: STAMP and AcciMap. Their results similarly showed that each approach was better at capturing different aspects of the accident.

A survey looking into European organizations' accident investigation procedures showed a lack of a standard method as well as a lack of adequate investigation knowledge and training for the participants (Roed-Larsen & Stoop, 2012). This lack of standard guidance could influence learning potential, especially if the importance of knowledge and training of the participants is underestimated.

1.1.3 Learning from accidents

Learning from accidents entails how well underlying risk factors, as well as risk-mitigating measures, are identified (Hovden et al., 2004). To increase learning potential from accident investigations there should also be solutions for how to implement such risk-mitigating measures. Hovden et al. (2004) suggest that to increase learning potential, accident investigations should give insight into underlying contributing factors, as this reveals more efficient measures to prevent future accidents.

Cedergren and Petersen (2011) argued that whether causal factors are determined at the micro level (physical processes, actor activities and equipment), meso level (organizational factors) or macro level (conditions related to regulators, associations, and government) will determine which of these levels will achieve most learning. After comparing railroad accident reports from three Scandinavian countries, they revealed that a large majority of attributed causes were determined at the micro level (68%-78%). Between 19% and 27% were at the meso level, and only 3 to 5% were at the macro level. This indicates a lack of learning potential for the higher levels. Furthermore, when Klaveness (2012) investigated the Norwegian petroleum industry's internal accident reports, only 3% of safety recommendations were aimed at the meso level, and 0% at the macro level, further supporting this lack of learning at higher levels.

Brath (2020) also investigated the Norwegian petroleum industry and found that there is a lack of evaluation of safety interventions after they have been implemented. This supports earlier studies with similar findings in other sectors and organizations (Cedergren, 2013; Drupsteen et al., 2013). Furthermore, the safety recommendations might never be implemented at all, due to resistance for a variety of reasons (Lundberg, et al. 2012). Stormo (2011) emphasises a few main contributing factors to this resistance: a lack of procedures for follow-up of safety measures, conflicts between different levels of government, and a lack of economically realistic recommendations.

To prevent such resistance Lundberg and his co-workers (2012) suggested 18 strategies, based on which safety culture they are dealing with. All 18 strategies use either power, trade-offs, duty, or orientation (increasing knowledge) to minimize resistance. Trade-off strategies are largely used for resource-weak cultures, as well as being the only culture where duty is emphasized. Power and orientation are used for cultures that simply have low safety standards, while rational cultures are more easily won over by using orienting tactics. Cultures that do not prioritize safety need a combination of power, trade-offs, and orienting.

There are clear difficulties with learning from accidents, despite accident investigations being common. Similar accidents keep occurring, without there being a clear way to prevent them. Investigation methods, accident models, socio-technical levels and recommendation strategies all seem to be aspects that can help improve learning potential and thus reducing either accident volumes or negative consequences, even if it is unlikely to prevent all accidents.

1.2 Purpose and research problem

Despite good research into accident methods, models, and learning, there is a lack of research into cross-continental differences in accident investigation reports. Many of the comparative studies are either between different organizations within the same country or neighbouring countries such as the Scandinavian countries. This could prevent long-distance learning, where differences might be bigger. In addition, focusing research on theoretical research methods rather than real reports might not show the real picture of how accident investigations are done in practice. Indeed, Karanikas et al.'s (2015) study showed that modern safety thinking was not present in newer Dutch accident reports and that their approach had not changed from 1999 to 2013, even though theories have evolved. There is also a lack of quantitative research on how different aspects of investigations influence the outcomes of the reports in practice.

The purpose of this thesis is to gain insight into how different aspects of modern accident investigations may influence outcomes, and further improve learning from accidents with this insight. A cross-continental comparison is done to highlight potential cross-continental differences, the effect of these, and the potential for the investigation boards to learn from each other. I have chosen to compare railroad accident reports from the last decade from the Norwegian Safety Investigation Authority (NSIA; Norwegian: Statens Havarikommisjon) and the United States' National Transportation Safety Board (NTSB). A quantitative analysis of different aspects in their reports could reveal weaknesses and strengths in their accident investigation practices, and how these aspects influence learning potential.

To examine the impact of different aspects of accident reports, and the difference between the two boards, I chose the following major research problem:

How do different aspects of accident investigations by the NSIA and the NTSB influence learning potential?

The relevant theory was used to develop further research questions to help answer the main research problem. These are presented in section 2.7.

1.3 Investigation boards

The accident investigation boards compared in this thesis are the NSIA and the NTSB, due to them both being national government agencies in charge of investigating transportation-related accidents. This entails aviation, railroad, marine and road/highway accidents, while the NTSB also investigates pipeline accidents and hazardous materials, and the NSIA includes the defence sector. Neither of these investigation boards appoints blame or liability under criminal or civil law, and their findings cannot be used in criminal investigations. They also do not have the authority to enforce their safety recommendations. The purpose of their investigations is to learn and prevent future accidents, and they describe their missions similarly. The Ministry of

Transportation explains “The purpose of the NSIA’s investigations is to elucidate matters deemed to be important to the prevention of accidents and serious incidents” (Samferdsels-departementet, 2020, p. 3) while the chairman of NTSB explains that their mission is “to learn from the accidents we investigate to keep them from happening again” (NTSB, 2017, p. 1).

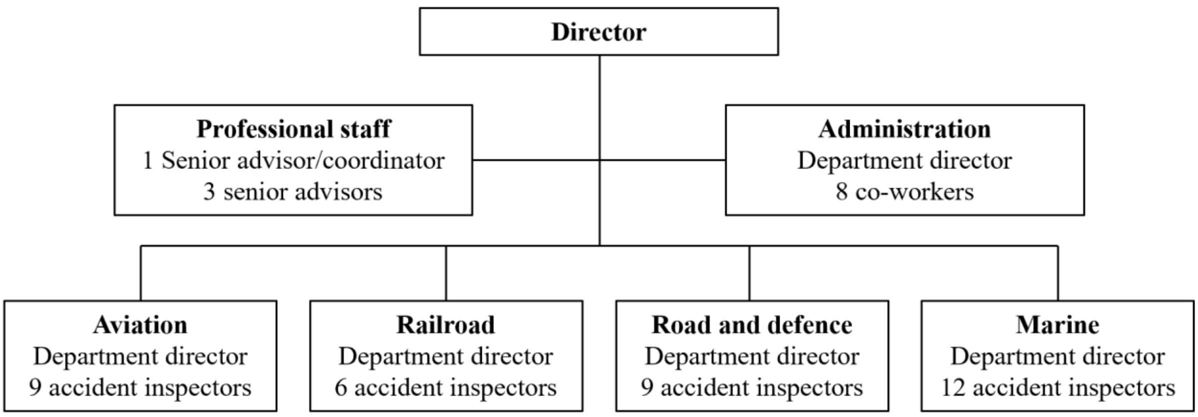
1.3.1 The Norwegian Safety Investigation Authority

The NSIA, previously named Accident Investigation Board Norway (AIBN), is under the Ministry of Transportation in Norway and was established in 1989. Before this, accident investigations were conducted by ad-hoc temporary committees. In the beginning, they only investigated aviation accidents and expanded to include railroad accidents in 2002, road accidents in 2005, marine accidents in 2008, and defence accidents in 2020 (NSIA, n.d.).

The NSIA is organized per Figure 2: lead by a director, with a staff of four advisors and an administration underneath the director, and lastly the different transportation sectors with their respective directors and inspectors. In total, they have 54 employees as of December 2020 (Statens Havarikommisjon, 2021).

Figure 2

NSIA‘s organization chart. Reproduced from Statens Havarikommisjon (2021)



The investigation board is mandated to investigate both accidents and serious incidents in their assigned transportation sectors (Samferdselsdepartementet, 2020). The road, aviation and railroad sector are under the Ministry of Transportation’s responsibility, the marine sector is under the Ministry of Trade, Industry and Fisheries, and the defence sector is under The Ministry of Defence. Their instructions are as follows (Samferdselsdepartementet, 2020, p. 4):

1. Investigate accidents and serious incidents in the specified sectors.

2. Write reports that include a statement from the NSIA about causal factors and potential safety recommendations, without underlying concrete solutions.
3. Perform specific tasks with a security-related purpose that the Ministry of Transportation may impose on the agency.
4. Represent the Ministry of Transportation and/or the Ministry of Trade, Industry and Fisheries and/or the Ministry of Defence when needed, or participate in meetings with these ministries.
5. Make statements regarding cases related to the mentioned ministries, and help with case processing when asked.
6. Collaborate with other businesses when deemed beneficial.

Specifically for the railroad sector, the NSIA is mandated to investigate railroad accidents and serious incidents in line with the Railway Investigation Act (Jernbaneundersøkelsesloven, 2005). This law entails what information needs to be provided to the investigators, how they can obtain information, rights of affected parties, confidentiality, and more.

1.3.2 The National Transportation Security Board

The NTSB was established as an independent agency as early as 1967, 22 years before their Norwegian counterpart. Already from the start, they were investigating accidents within aviation, highway, marine, pipeline, railroad, public transportation, and transportation of hazardous materials (NTSB, 2017), giving them more experience than the NSIA. They were originally under the U.S. Department of Transportation, but in 1974 they became a separate entity, independent from any other government agency (NTSB, 2017).

The organization of the NTSB is a little different from the NSIA. The official board consists of five Board Members, nominated by the President before being confirmed by the Senate to serve a 5-year term (NTSB, 2020). With their approximately 400 employees, the NTSB is also significantly larger than NSIA. The organizational chart for NTSB is shown in Figure 3.

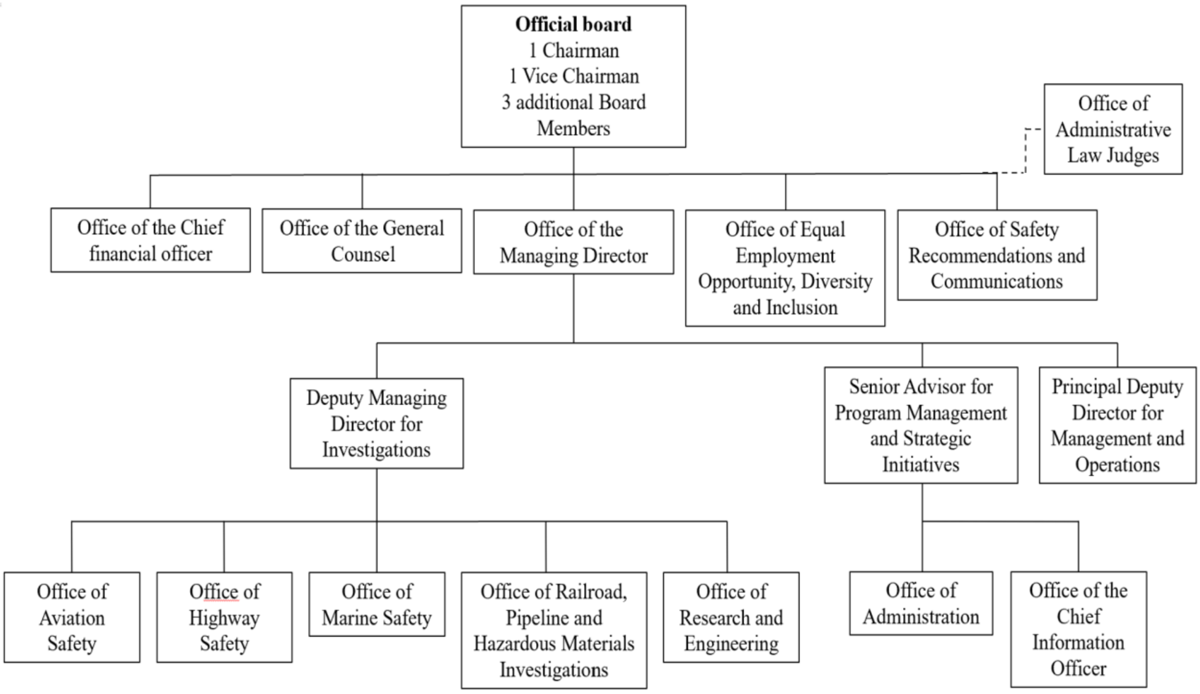
The NTSB’s legislative mandate entails (NTSB, 2017, p. 5):

1. Maintaining their congressionally mandated independence and objectivity.
2. Conducting objective, precise investigations and safety studies.
3. Performing fair and objective airman and mariner certification appeals.
4. Advocating and promoting safety recommendations.

5. Assisting victims of transportation accidents and their families.

Figure 3

NTSB Organizational chart. Reproduced from NTSB (2020)



U.S. Code Title 49, chapter 11 (§§ 1101-1155) contains statutes regarding the organization and administration of the NTSB, their authority, and enforcement and penalties. Furthermore, the Code of Federal Regulations Title 49 describes their responsibility within the railroad sector: "The NTSB is responsible for the investigation of railroad accidents, collisions, crashes, derailments, explosions, incidents, and releases in which there is a fatality, substantial property damage, or which involve a passenger train" (49 CFR § 831.40).

1.4 Further structure

The purpose of this study is to investigate how different aspects of accident investigations may influence the outcome, and thus also the learning potential. Chapter 2 consists of an introduction to the theoretical framework that will be used to identify and analyse these different aspects. The methods, research design and executions of all analyses are explained in chapter 3, while the results follow in chapter 4. These results are discussed in chapter 5, considering the theoretical framework presented earlier. At the end of the thesis, chapter 6, I present some concluding remarks, reflections, and summaries regarding the findings and their possible indications, in addition to possible future research.

2. Theory

2.1 Causality

One purpose of accident investigations is to find out both how and why accidents happen (Hollnagel, 2004). This leads us to the concept of causality. To gain knowledge about how accidents happen, and thus how to prevent them, accident investigations often seek out the *cause* of the incident. Merriam-Webster's Collegiate Dictionary (2021) defines a cause as "something that brings about an effect or a result". This might sound simple, but there are disagreements about what a cause is.

David Hume, a philosopher from the 1700s, is known for his philosophy regarding causality. He argued that for something to be the true cause, it needs to proceed the effect in time, there needs to be a certain connection between the cause and effect, and the same cause always has the same effect. If it ever fails to produce the same effect it cannot be the sole cause but must be assisted by some other circumstance (Hume, 2009).

A strong cause and effect relationship is a common way to view causality, similarly to the sequential relationship in laws of physics (Hollnagel, 2004). Leveson (2004) refers to this as a direct, linear cause. This is related to the principle behind Root Cause Analyses (Cojazzi & Pinola, 1994), which focuses on the possibility of finding one root or origin of an event. Often, the root cause is seen as close in space and time, and what Woods et al. (2010) named 'sharp-end' factors in 1994. However, the true cause can also be a factor removed in space and time, a so-called 'blunt-end' factor (Woods et al., 2010). Leveson (2004) explains that the most important factor in the occurrence of accidents seems to be related to management commitment to safety culture, which is at the blunt end of the scale.

A concept of such a 'one true cause' is considered naïve by some, regardless of it being at the blunt or sharp end of the scale. Hollnagel (2004) claims that accidents might not even have a cause, but rather explanations. Accidents can happen because of several factors coming together at a specific time, likely a combination of both sharp and blunt end factors. None of these factors is necessarily the cause of the accident. The cause, if any, is the simple coincidence that these factors occurred at the same time, not the factors themselves. A more effective strategy for accident mitigation would therefore be to find and control the conditions that lead to accidents, rather than finding and destroying the cause.

This illustrates how the view of accident causality may have large implications for the accident investigation itself, and the importance of gaining insight into this effect. Hollnagel (2004) refers to the stereotypical ways of thinking about how an accident occurs as ‘accident models’, although they are also known as ‘accident causation models’ (Katsakiori et al., 2009). These accident models are frames of references or a common frame of understanding. Leveson (2004) states that accident models affect both the data collected and the factors considered as possible causes in accident investigations. As a result, she says “they may either act as a filter and bias toward considering only certain events and conditions or they may expand activities by forcing consideration of factors that are often omitted” (p. 237).

2.2 Accident models

Heinrich (1941) is considered to have developed the very first accident causation model: The Domino model. His model is in line with what Leveson (2004) describes as a direct and linear cause. The model implies that events happen one-by-one, until they eventually lead to the accident, just like dominoes falling over in a chain reaction. As only one event happens at a time, it can be considered a one-dimensional sequence of events (Katsakiori et al., 2009).

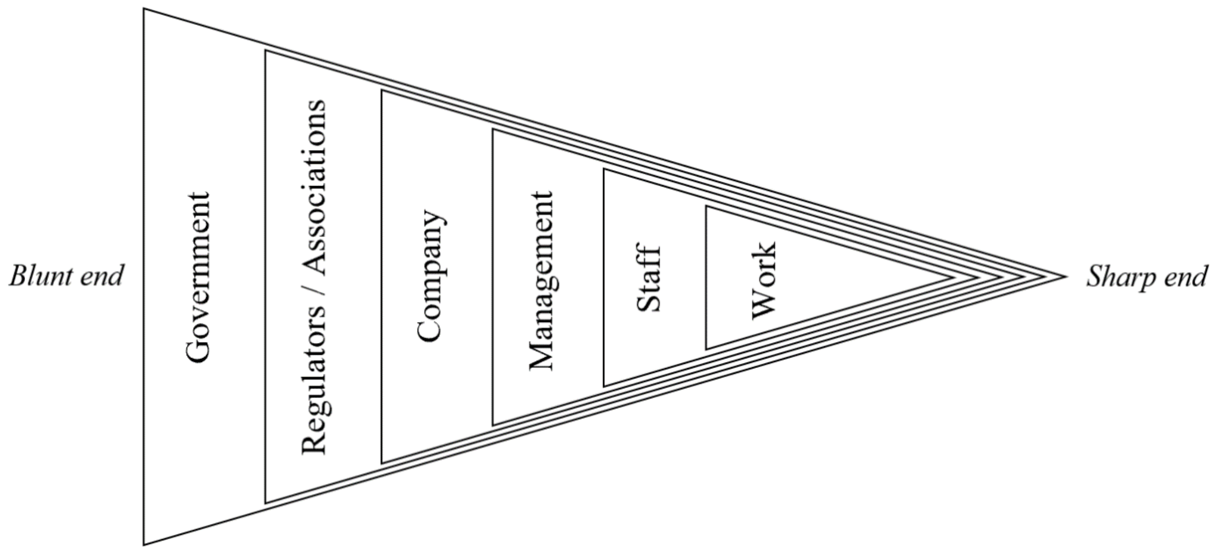
In the late 1980s, James Reason (1997) introduced the idea of active failures interacting with latent conditions. Active failures are the more immediate and obvious acts leading accidents to happen, similar to sharp-end factors. Latent conditions, on the other hand, go beyond individual acts and are related to factors in the system, which can be hidden for years before being exposed through the active failures. These latent conditions have similarities with blunt-end factors, but they are considered more dormant than simply being removed in space and time. Reason’s model, called the Swiss cheese model, demonstrates these latent conditions as the holes in cheese slices (symbolizing barriers), which an accident trajectory must go through before potentially causing an accident.

A lot of models throughout the years have focused on the contribution of human activity to accidents (Katsakiori et al., 2009). In fact, Hollnagel (2004) has it as part of his definition of accidents, as he says that accidents “must directly or indirectly be the result of human activity” (p. 5). Human error can essentially be seen in two ways: The ‘new view’ and the ‘old view’. The old view considers human error the cause of accidents, while the new view suggests that human error is the symptom of latent conditions deeper in the system (Dekker, 2014). Hale and Glendon (1987) proposed a model where human action controls a danger in the workplace that is always present. In this way, danger can both be created and prevented through people’s actions on several levels in the system.

This brings about socio-technical approaches and the consideration of different socio-technical levels. A socio-technical system refers to modern systems consisting of both human and technological interactions, in addition to being a part of a bigger social structure with multiple levels (Qureshi, 2007). Rasmussen (1997) presented a model which included six different levels in the socio-technical system involved in risk management. These are (1) government, (2) regulators and associations, (3) company, (4) management, (5) staff, and (6) work. Rassmussen's model was intended to give an overview of risk management, but this is also closely connected to accident causation. If all these levels are involved in the management of risk, they are also all involved in preventing – or failing to prevent – an accident. His levels can further be combined with the idea of sharp- and blunt-end factors, and in Figure 4 the levels are arranged in terms of 'sharpness'.

Figure 4

Rasmussen's (1997) socio-technical levels as sharp- and blunt-end factors



Nancy Leveson (2004) argued that many accident models were simply too subjective to properly investigate accidents and that modern socio-technical systems are too complex to be explained through simple cause-effect relationships. She introduced the so-called Systems-Theoretic Accident Model and Process (STAMP), developed with Rasmussen's socio-technical levels in mind. According to STAMP, accidents are not caused by events, but rather a result of a lack of 'constraints', or safety controls. These constraints must be enforced on each of Rasmussen's socio-technical levels. Systems need to constantly adapt and change as a part of a feedback loop between information and control, resulting in an interrelated and dynamic design (Leveson, 2004).

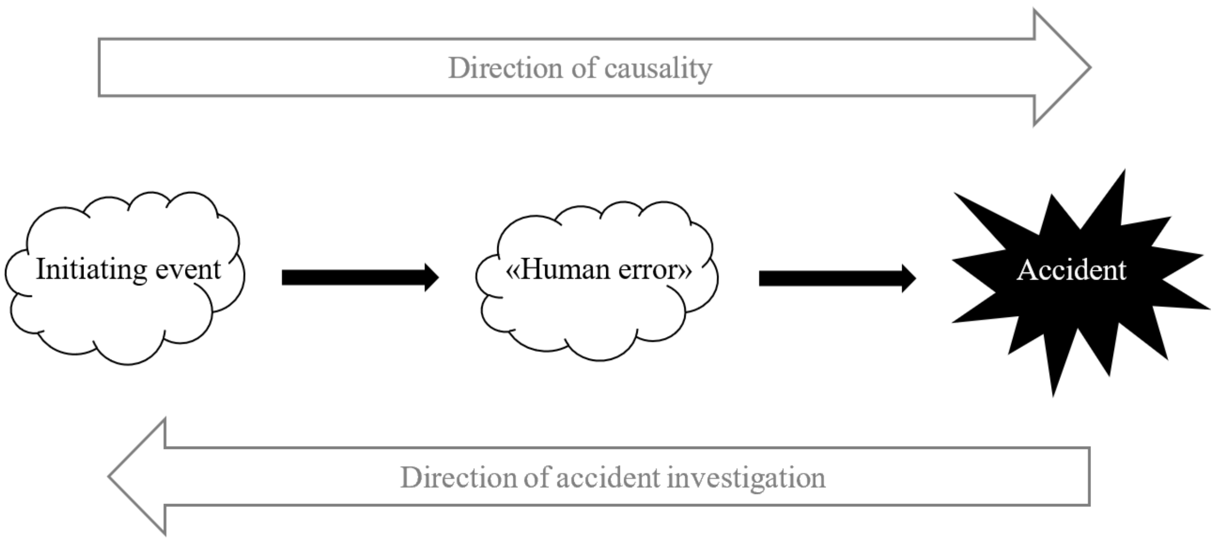
As many more accident models have emerged over the years, Hollnagel (2004) suggested using three categories in which most of these models fit into sequential accident models, epidemiological accident models, and systemic accident models. I will briefly introduce all three categories, and how each type can influence accident investigations.

2.2.1 Sequential accident models

As the name indicates, sequential accident models refer to accidents as the result of a sequence of events. These events occur one-by-one in a specific order, where the last event is the accident (Hollnagel, 2004). The triggering unexpected event is often assumed to be an unsafe act, or 'human error', according to Hollnagel (2004), even though it can just as likely be something else. Heinrich's (1941) Domino model is an example of a sequential accident model, where each domino represents one event in the chain. These types of models assume a clear cause-effect link, which coincides with the principle of a root cause. This simplistic relationship is illustrated in Figure 5.

Figure 5

Sequential accident models. Adapted from Hollnagel (2004)



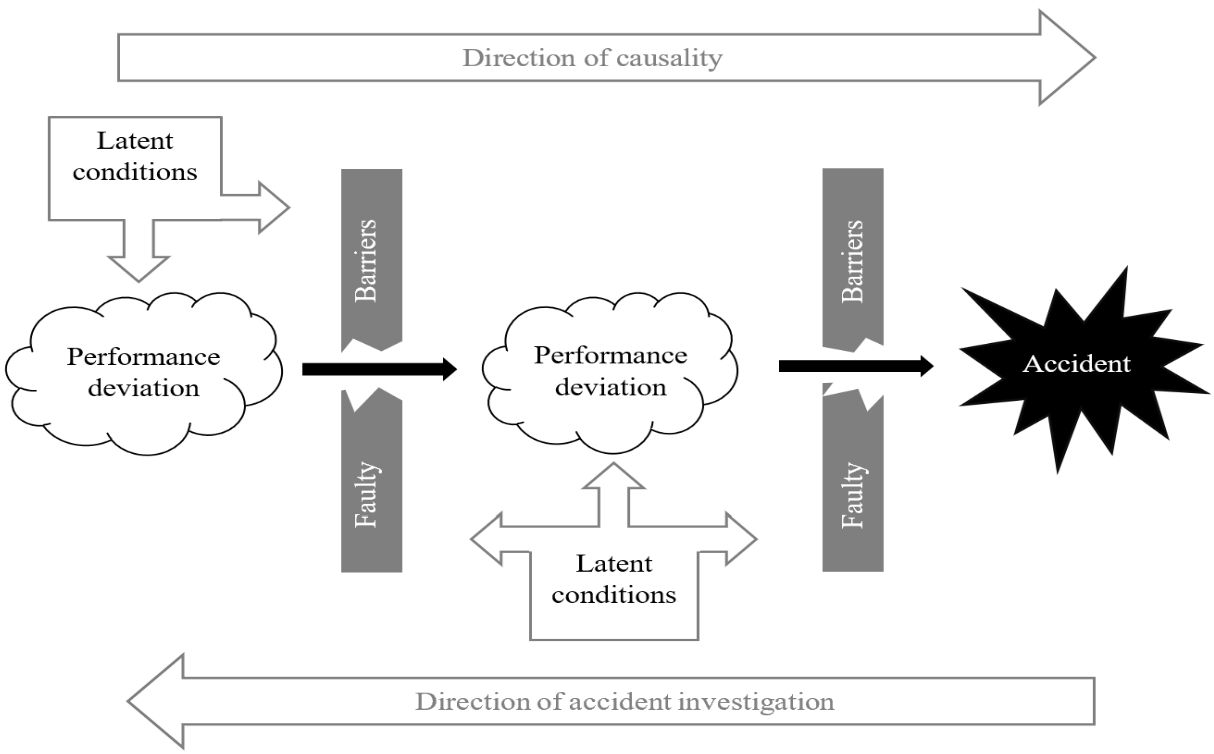
Accident investigations using sequential accident models would search for, and eliminate, the initiating event to prevent similar accidents from happening again. This search for the cause would typically start at the accident and work its way back in the chain of events until the root is found (Hollnagel, 2004).

2.2.2 Epidemiological accident models

While epidemiological accident models also work with a sequential understanding of accident causation, they are not one-dimensional. Instead, they compare accident causation with diseases, and how environmental factors can play an important role in the offset of an accident (Hollnagel, 2004). They include a combination of both latent conditions and active failures, in line with Reason's (1997) thinking, rather than one triggering unexpected event. This is an essential change of focus, as latent conditions can be present a long time before an accident occurs. Additionally, 'human error' is replaced by 'performance deviations', as humans are not the only cause of such events, as well as 'deviation' being less loaded than 'error' (Hollnagel, 2004). In line with this change of wording comes the consideration of environmental factors that could be the cause of the performance deviations, instead of the deviations being to blame. This is in line with the new view of thinking of human error, while sequential models use the old view to a bigger extent.

Figure 6

Epidemiological accident models. Inspired by Hollnagel (2004)



In contrast with sequential accident models, epidemiological models include barrier thinking, which is different factors that could prevent an accident from happening. Reason's (1997) Swiss cheese model illustrates these barriers through the cheese slices that the accident trajectory goes through. Figure 6 shows the relationship between latent conditions, barriers and performance deviations commonly found in epidemiological models.

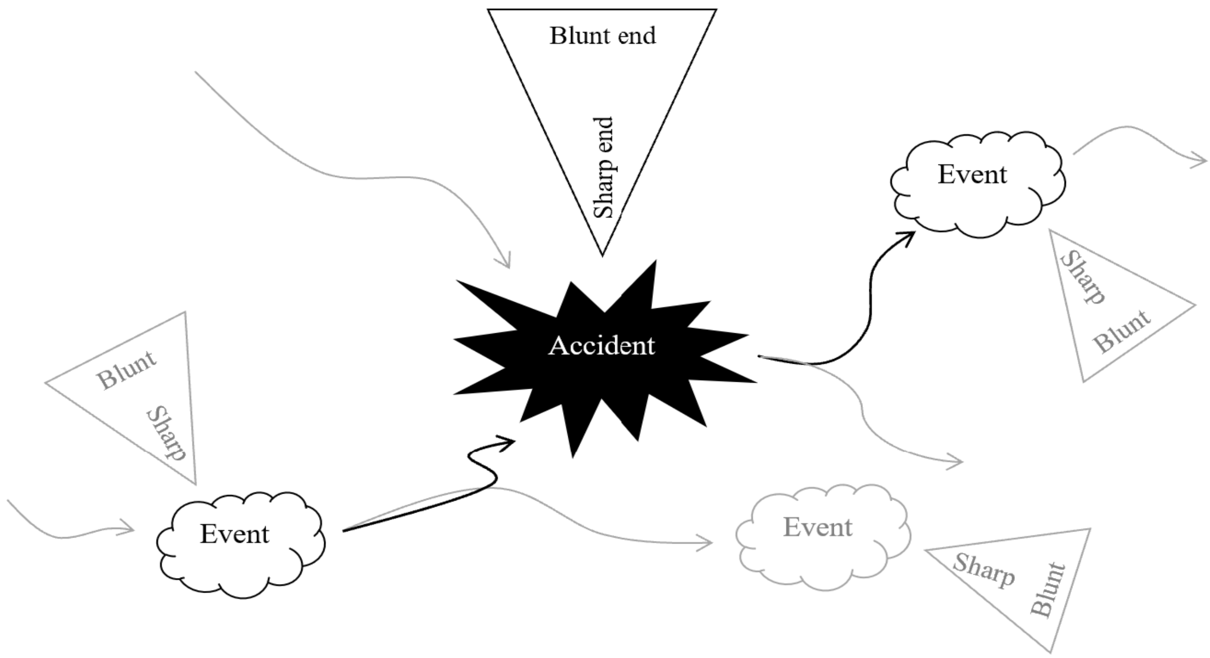
As a result of these differences, accident investigations also change their focus. Instead of seeking out and destroying the assumed one true cause, the investigators search for contributing latent conditions, and failed or missing barriers. Attempts to prevent future accidents would be done by improving or implementing barriers and strengthening defences (Hollnagel, 2004).

2.2.3 Systemic accident models

Taking a step away from linear thinking, systemic accident models consider the whole system, and the interplay between human, technological and organizational factors. Accidents are naturally still developing in line with time, but systemic accident models consider everything that happens along the way as part of one whole system, rather than distinct events (Hollnagel, 2004). Every accident is both preceded and followed by events, each of which has contributing sharp-end and blunt-end factors. Figure 7 shows an example of this.

Figure 7

Systemic accident models. Reproduced from Hollnagel (2004)



Hollnagel (2004) points out the difference between accidents as ‘resultant’ and ‘emergent’. Something resultant is predictable based on the contributing factors, while something emergent is not. These models consider systems to have such complex interactions that all the possible ways they may interact are simply not possible to predict. Accidents thus may seem random and unavoidable. This, largely because even small and seemingly insignificant events can set off very large consequences (Hollnagel, 2004).

Rasmussen’s (1997) socio-technical levels is an example of a systemic accident model. It shows how different levels of a system – and the interactions between them – have influences on accident occurrences. Leveson’s (2004) STAMP model also fits into this category, with a focus on the dynamic big picture, and safety control rather than unexpected events.

Systemic accident models also have consequences for the focus in an accident investigation. Rather than seeking out isolated causes, one would try to analyse the system’s performance as a whole and variability in this performance. Some dependencies and interactions might be correlated with the emergence of accidents, and these patterns can be used to prevent accidents before they happen rather than just as an ad-hoc response. It is also necessary to understand the difference between positive growth-related variability in the system, and possible negative variation (Hollnagel, 2004).

2.2.4 Alternative categories

While this paper focuses on Hollnagel’s three categories, it is not the only way to categorize accident models. For example, Katsakiori et al. (2009) review five different ways to categorise accident models, before deciding to replace epidemiological accident models with human information processing accident models. This gives more focus to cognitive psychology, how the human brain works, and the processes that influence human behaviour during unexpected events. On the other hand, it makes less of a distinction between the largely different models that all see accidents as happening in a sequence of events.

Katsakiori et al.’s (2009) human information processing models are inspired by one of Lehto and Salvendy’s (1991) classifications. They discovered that almost all accident models at the time explicitly considered human factors, and they called these ‘models of human error and unsafe behaviour’. Lehto and Salvendy identified different types of models of human error and unsafe behaviour, and human information processing models was one of these categories. In addition, there were behavioural models, focusing on human traits as causes of behaviour, which would largely be covered by Hollnagel's sequential models, as they do not consider outside influences. In contrast, another of the categories points more attention to situational reasons for human error, which Hollnagel would identify as epidemiological models.

Fu et al. (2020) use a more comprehensive classification, where accident models are first divided as linear or nonlinear. While linear models focus on a chain of events and the interaction between different contributing factors, nonlinear models choose to focus on a few factors of an accident. After that, the nonlinear accident models are split into four new categories: human-based, statistics-based, energy-based, and system-based.

Despite various other categories, Hollnagel's alternative is both acknowledged and suitable for traditional accident models and more modern socio-technical models (Fukuoka & Furusho, 2017). Other alternative categories are largely covered by Hollnagel's categories while putting the main focus on different aspects. Some alternatives are also essentially the same categories, but with different names. One example is Toft et al.'s (2012) three historical phases of accident models. Choice of categories may also depend on which context they are to be used in, to purposefully highlight different aspects.

2.3 Perspectives on accidents

Despite the amount of literature on accident causation, theories are claiming that it might not be possible to uncover the cause of an accident at all. Rather, it has been suggested that causes are something that we *construct*, based on which accident models the investigators use. Lundberg et al. (2009) call this phenomenon 'What-You-Look-For-Is-What-You-Find', or WYLFIWYF. These accident models tend to be grounded in perspectives on major accidents, which entails accident theory in a broader sense. While accident models concern *how* an accident happens (causes), these perspectives also include theory on *why* accidents happen (explanations), and how to prevent them. The accident models can thus fit into different ones of these perspectives.

In this section, I will briefly present a few of the most acknowledged perspectives on major accidents; the energy-barrier perspective, man-made disasters, normal accidents theory, high reliability organizations (HRO) theory, resilience engineering, and the conflicting objectives perspective. These perspectives have all had a major impact on the field of safety science, in addition to influencing practical safety management.

2.3.1 The energy-barrier perspective

The energy-barrier perspective, developed by Haddon (1970), explains accidents through harmful energy that reaches vulnerable targets. This can happen if there is a lack of effective barriers to stop the harmful energy before it reaches the target, and the perspective has been highly influential on the idea of safety in design (Hovden, 2010; Rosness et al. 2004). Haddon (1970) identified ten different strategies to reduce losses from accidents, that he later separated into three categories: (1) reduction or modification of the energy source, (2) separating the energy and target, and (3) resilience and rehabilitation of the target (Haddon, 1980).

Sklet (2004, p. 31) defines safety barriers as “any means used to control, prevent or impede the hazard from reaching the target”. He categorizes them into physical barriers and management barriers, but they can also be grouped into proactive barriers (frequency-reducing) and reactive (consequence-reducing) barriers (Hovden et al., 2010). Accident investigations influenced by this perspective would have an increased focus on these safety barriers, and how they influenced the accident (Sklet, 2004). Reason’s (1997) Swiss cheese model is largely compatible with the energy-barrier perspective, as his cheese slices illustrate barriers, and the accident trajectory fits with the idea of harmful energy passing through.

2.3.2 Normal accidents theory

Charles Perrow (1984) argued that systems with both tight coupling and interactive complexity are simply built in a way that makes accidents inevitable, or normal. Tight coupling entails interactions that are close in space and time, to ensure speedy production. This also makes the system more difficult to stop if an accident were to happen, which further allows negative consequences to propagate through the system (Dekker, 2014). Interactive complexity, on the other hand, describes the interactions between different system components. Instead of the interactions being simple and linear, they are connected in ways that are so complicated that it is impossible to foresee how they might influence each other (Perrow, 1984). Systems with this combination should simply not be allowed to exist, according to Perrow, as accidents are bound to happen.

While the normal accidents theory gained massive attention, it is also considered ‘debunked’, as critics have shown it cannot be applied to any accidents – not even the accident Perrow himself used to develop his theory (Hopkins, 2001). Nevertheless, the theory still had a profound influence on how major accidents are understood. The normal accidents theory was largely developed to lead the blame away from individuals and onto dangerous characteristics of certain technologies and control structures (Perrow, 1984). The systemic accident models are useful in combination with the normal accidents perspective for this reason. The shift of focus has also inspired other perspectives, such as Turner's man-made disasters, and Rasmussen's conflicting objectives perspective. Furthermore, HRO originated as a counter-response to the normal accident theory.

2.3.3 High reliability organizations and resilience engineering

As a response to normal accidents, HRO theory is based on studies on organizations that handle the interactive complexity and tight coupling without major accidents (La Porte & Consolini, 1991; Rochlin et al., 1987). According to the HRO perspective, accidents are avoidable if the system is organized and controlled in a good way (Dekker, 2014). Kongsvik et al. (2018) point out three characteristics of an HRO: use of redundancy, spontaneous reconfiguring, and 'mindfulness'. Redundancy refers to having more than the bare necessities and overlapping work tasks and competencies. Spontaneous reconfiguring describes the organizations' ability to change between centralised and decentralised management, as they are needed at different times. For example, when unexpected events occur it is necessary to use a decentralised line of command, as the workers 'on the ground' have more hands-on experience. Lastly, mindfulness refers to their continuous risk awareness and attention to potential failures.

This perspective is not so much about how accidents happen, but how to prevent them from happening. However, if an organization is lacking any of the needed characteristics, this could be considered the explanation for why an accident occurred. The lack of focus on the causal link does on the other hand 1leads to a lack of fitting accident causation models.

Resilience engineering is a perspective that can be associated with the HRO perspective, as they both focus on resilience and how to make organizations capable of maintaining security – even during pressing and difficult situations that could potentially cause harm (Kongsvik et al., 2018). HRO is used more frequently, as it includes more aspects than resilience, in addition to being the first of the two to be developed. Resilience engineering does on the other hand encourage increased focus on what causes things to go right, rather than only focusing on accidents and what causes them. The perspective also warns against trusting rules and routine too much, as this will prevent adaptability in new and challenging situations (Kongsvik et al., 2018).

2.3.4 Man-made disasters

Turner and Pidgeon (1997) proposed the MMD perspective, suggesting that accidents are not sudden phenomena. Instead, they develop over time during the so-called ‘disaster incubation period’. During the incubation period, several latent errors can build up because of norms and beliefs about hazards that do not add up with reality. This leads to the system getting increasingly more vulnerable as new events occur without properly being understood or noticed. Pidgeon and O’Leary (2000, p. 16) further explain that “accidents arise from an interaction between human and organizational arrangements of the socio-technical systems set up to manage complex and ill-structured risk problems” rather than from chance or ‘Acts of God’.

The MMD perspective describes six stages of an accident: (1) normal operations, (2) incubation period with misperceptions and lack of information flow, (3) precipitating event, (4) onset, (5) rescue, dealing with immediate problems, and (6) full cultural readjustment (Turner & Pidgeon, 1997). Accident investigations using this perspective could look for these steps, to gather information about how the accident developed over time.

Kongsvik et al. (2018) highlight a few factors that can contribute to the build-up of latent errors: lack of sufficient flow of information and keeping employees involved and informed, that information is simply overlooked or misinterpreted or that information is not seen in the right context. There is almost always someone with information about possible incubating accidents, but no action is taken due to this lack of information flow (Turner & Pidgeon, 1997). A combination of perception, interpretation, and flow of information across socio-technical levels is thus highly important and resembles aspects of systemic accident models. Additionally, the focus on latent conditions building up over time resembles the idea behind epidemiological accident models, so these accident models could be used with a man-made disasters perspective.

2.3.5 The conflicting objectives perspective

Rasmussen (1997) developed the conflicting objectives perspective building on his socio-technical levels. The perspective highlights conflicting interests between actors at different levels, and how these interests could compromise safety in unexpected ways. These interests are often related to efficiency, economic costs, and workload. It is necessary to find a balance, to avoid bankruptcy and overworked employees, but also to keep a solid security level. While investigating accidents, these conflicting interests would be relevant to look for as potential explanations. This puts the focus on the system as a whole, and as Rasmussen developed both the conflicting objectives perspective and the socio-technical levels, systemic accident models can be a good fit for this perspective. It is also one of the only perspectives that have a scope reaching outside of the organization, including government and regulatory levels as possible explanations.

2.3.6 Combining perspectives

While these perspectives of major accidents are vastly different in ways, they are not necessarily mutually exclusive. It has been proposed that the use of one single perspective alone is not sufficient to explain accidents, but that their combined use gives a better overall understanding (Kim & Haugen, 2015). Combining perspectives may broaden what accident investigations look for and consider as possible explanations and contributing factors, including all socio-technical levels and both sharp-end and blunt-end factors. This combination might also contribute to a lesser WYLFIWYF-effect.

2.4 Investigation methods

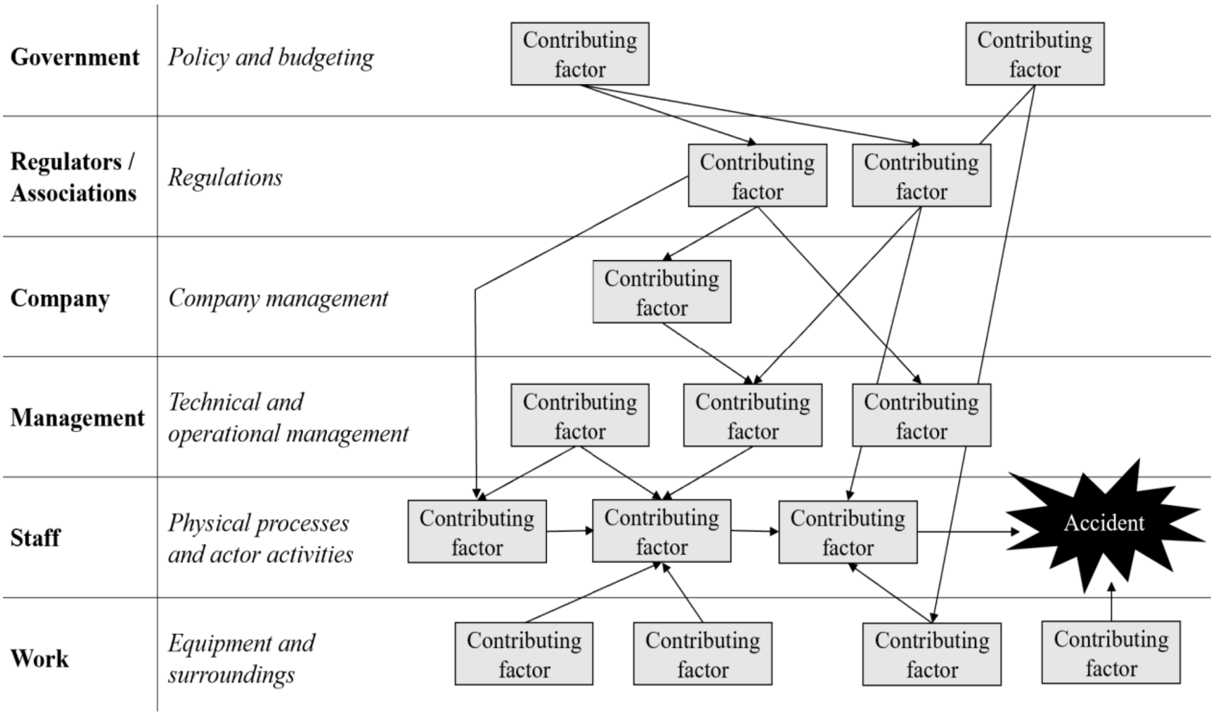
Despite being influenced by differing accident models and perspectives, the process of investigating accidents appears to have some common traits. Lundberg et al. (2009) identified four steps after analysing eight investigation manuals in Sweden: plan and initiate, collect data, analyse, and recommend. The U.S. Department of Energy (DOE, 2000) goes more in-depth in their investigation workbook with seven sections. However, they fit these steps into three phases that have high similarity to Lundberg et al.'s steps: a collection of evidence and facts, analysis, and developing judgements and report. The DOE emphasizes that these phases do not occur strictly chronologically but have a high level of overlap between them.

Despite these similarities, there are different analytical tools developed to make the investigation process easier. They are also referred to as different investigation methods, and they use graphical displays to give a better understanding of how an accident unfolds. Some of the most used investigation methods are AcciMap; Man, Technology, and Organization (MTO) analysis; and Sequentially Timed Event Plotting (STEP).

AcciMap is developed by Rasmussen (1997) and shows a visual presentation of decisions and actions at different socio-technical levels. The model emphasizes how these decisions and actions relate to each other and how they can influence the risk of accidents (Kjellén & Albrechtsen, 2017). An example of this is shown in Figure 8.

Figure 8

Principal illustration of AcciMap. Adapted from Goode et al. (2019)

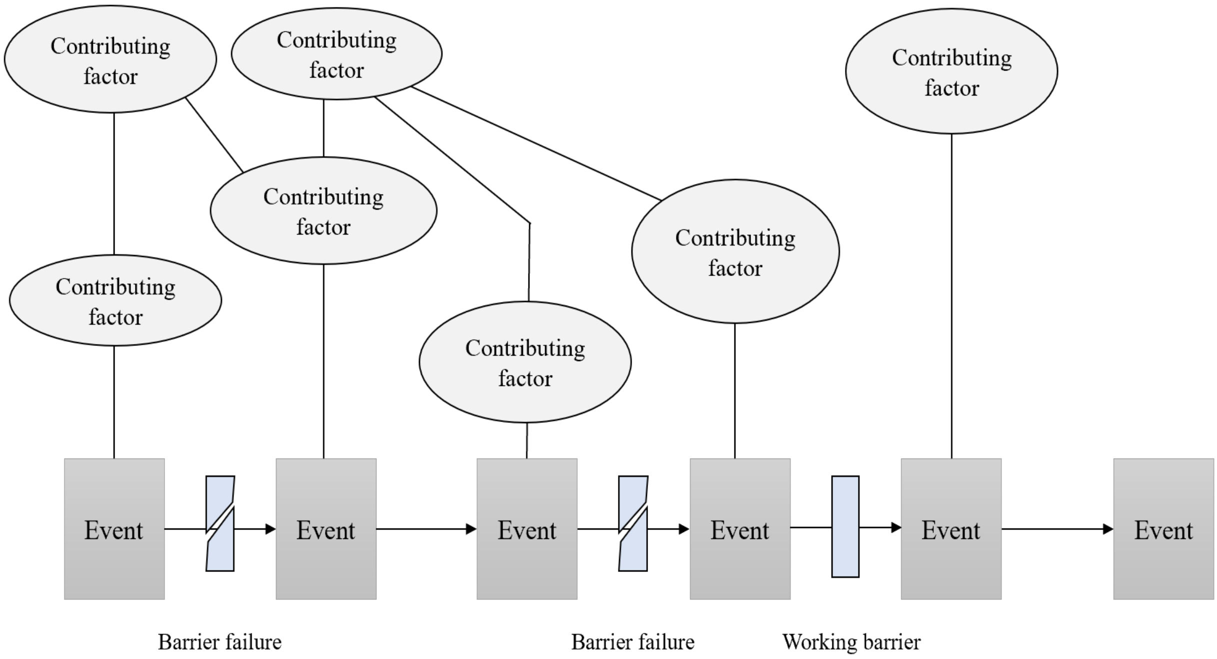


There are apparent connections between AcciMap and systemic accident models, as well as the conflicting objectives perspective. Rasmussen developed his systemic accident model, the conflicting objectives perspective, and AcciMap all based on his socio-technical levels. This clear connection is a good illustration of how the underlying understanding of accidents and causation influenced both perspectives on accidents and the method of choice. Since the accident models focus on such vastly different causal factors, it is essential to have a method that helps investigate exactly these factors.

In contrast with AcciMap, MTO uses events and causal factors charting, which means a linear model of interlinked events and causal factors (Kjellén & Albrechtsen, 2017). Each event's causal factors are analysed, and the MTO method investigates factors linked to man, technology, and organization. MTO includes a focus on barriers and performance deviations, illustrated in Figure 9.

Figure 9

Principal illustration of an MTO analysis. Adapted from Tinnmannsvik & Kjellen (2018)



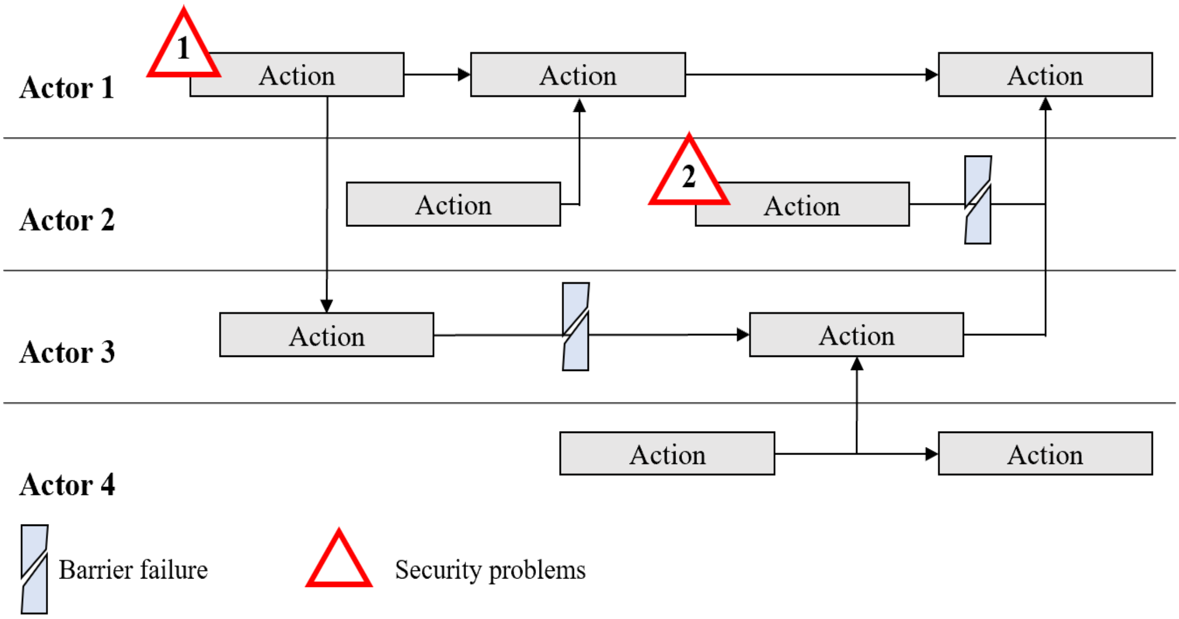
This inclusion of barriers, performance deviations, and latent contributing factors makes this method a good match for epidemiological accident models. It does not simply investigate one cause and effect, making it more advanced than what would be necessary with a sequential accident model. Furthermore, a systemic accident model requires more focus on interactions between different actors, actor levels, and the system as a unit. This analysis does not separate actors or socio-technical levels, making it more complimenting for epidemiological models.

While STEP uses a linear model similar to the MTO method, it is multi-linear, and more suitable for complex accidents involving many interacting actors (Kjellén & Albrechtsen, 2017). The NTSB is considered to have introduced multi-linear events sequencing concepts as early as the 1970s and are still frequently seen using such diagrams in their reports today (Sothivanan & Siddiqui, 2015).

STEP is a matrix-based version developed based on these concepts. It illustrates all relevant actors involved in an accident, both on the sharp end and blunt end, and their actions over time. These actions are the focus of the diagram, rather than events, exemplified in Figure 10. This results in multiple interacting sequences of events, rather than just one, and highlights that several activities take place at the same time. This is more in line with a systemic accident model, but instead of looking at actor levels, it identifies individuals. Although it is more complimentary of a systemic accident model than the MTO analysis, the STEP analysis is also mostly in line with epidemiological accident models because of this. However, it lacks the barrier analysis of the MTO method, which is an important focus in epidemiological models.

Figure 10

Principal illustration of STEP. Adapted from Tinnmannsvik & Kjellen (2018)



As all these methods have different strengths and weaknesses, the NSIA constructed their own method based on all three. They include the multi-linear sequence of the STEP model, the barrier-focus and causal analysis of each event from MTO, and the analysis of interacting socio-technical levels from AcciMap.

2.5 Learning from accidents

One of the most important goals of accident investigations is to learn from previous mistakes and prevent them from happening again, thus reducing accident frequency or negative consequences. Organizational learning involves, according to Argyris and Schön (1978), both detecting and correcting errors, making accident investigations ideal for this purpose. However, Fiol and Lyles (1985) point out that organizational change does not necessarily imply learning, and that there is a difference between learning and adapting. Learning implies cognitive development while adapting relates to behavioural development. One can change behaviour without it being grounded in knowledge, and gain knowledge without changing behaviour.

Argyris and Schön (1978) differentiate between 'single-loop' learning and 'double-loop' learning, which Fiol and Lyles (1985) later refer to as lower- and higher-level learning:

> When the error detected and corrected permits the organization to carry on its present policies or achieve its present objectives, then that error-detection-and-correction-process is single-loop learning. […] Double-loop learning occurs when the error is detected and corrected in ways that involve the modification of an organization's underlying norms, policies, and objectives. (Argyris & Schön, 1978, p. 3)

While single-loop learning focuses on specific actions and behaviours, double-loop learning focuses on overall rules and norms (Fiol & Lyles, 1985). Double-loop learning can be argued to have a more long-lasting effect because of this, while also having an impact on the whole organization rather than a specific part. Double-loop learning may lead to new cognitive frameworks that can guide decisions and prevent future accidents. However, not all double-loop learning is automatically positive. It can contribute to an organization with dysfunctional norms and superstitions that lead to avoiding problems rather than handling them.

Models and perspectives that focus on blunt-end and systemic factors, like the STAMP model and the conflicting objectives perspective, show more signs of double-loop learning. This is because they investigate the organization as a unit, and its underlying norms, policies, and objectives. Models and perspectives focusing on sharp-end factors, like the domino model and the energy-barrier perspective, show more signs of single-loop learning. Systemic and epidemiological accident models thus might show a larger learning potential than sequential accident models. The same thing applies to accident methods and their focus on higher or lower socio-technical levels. The levels of the safety recommendations can as a result indicate whether they will achieve double-loop or single-loop learning.

2.6 Investigator expertise

Based on previous studies, investigators' area of expertise is theorized to influence the investigations. Cedergren and Petersen (2011) found that investigation reports reflect the investigators' knowledge and concluded that the investigators tend to focus on areas of their expertise. As a result, investigators with purely technical and operational expertise might limit the scope to micro-level factors (actor activities and equipment). Micro-level factors usually are not enough to explain an accident on their own but could be symptoms of higher-level problems. Focusing solely on these factors is similar to sequential accident models, and an idea of a simple cause and effect. Such a scope might not be sufficient to gain a deeper understanding of the factors leading to accidents (Cedergren & Petersen, 2011).

To minimize this effect, it has been suggested to use larger cross-sectoral investigation boards with more diverse competencies (Cedergren & Petersen, 2011; Stemn et al., 2020; Svenson et al., 1999). Stemn et al. (2020) argue that investigating with different professional perspectives will increase the learning potential. This is because it can give a wider insight into the accident and interactions between different system components, which again can improve understanding of a complex system. In this way including more diverse expertise might promote systemic accident model thinking.

Le Coze (2013) points out how the use of accident models also depends on the user. He argues that models such as Reason's Swiss Cheese model are not very specific, and leaves investigators with their expertise to determine what the 'holes' in the slices are. Since these holes, or latent conditions, can be related to different parts of a big and complex system, many different qualifications are required to identify and learn from them.

2.7 Research questions considering the theory

These theories have been used to develop research questions to guide the analyses and answer the main research problem: *How do different aspects of accident investigations by the NSIA and the NTSB influence learning potential?* They relate to the influence of accident models and views, expertise and background, investigation process and methods, and socio-technical levels:

1. How can investigation procedures and methods affect the outcome of investigations?

While the investigation process and method utilized is argued to have implications for the whole investigation, there is a lack of details on what exactly the effect is, and to which parts. Knowledge of how much (if any) impact this aspect has on the outcome of investigations can give insight into the learning potential following different investigation methods.

2. How can the investigation boards' expertise influence the investigation?

The investigators' expertise is another aspect of accident investigations that has been theorized to influence the investigations' scope, outcomes, and learning potential. It is thus a relevant aspect to examine to answer the problem definition.

3. How do different accident models influence conclusions and safety recommendations?

Accident models are considered a central aspect of accident investigations, and previous research indicates that accident models and their following views on causality have implications for the investigation process and outcomes. It is relevant to find out whether this is true in practice for the selected boards, particularly because the conclusions and safety recommendations are important tools for learning from accidents.

Based on previous research and findings, I hypothesized that different accident models would show significant differences in learning potential, in the way that more system-aware models increase safety recommendations on higher socio-technical levels. I further hypothesized that the investigation boards would show differences in fatalities (and thus severity), accident models, and percentage of specific recommendations, as well as expertise and methods. The combination of these differences was hypothesized to influence the percentage of recommendations at each socio-technical level. Systemic models were expected to have most recommendations at the top levels, with epidemiological models focusing more on the middle levels and sequential focusing on lower levels.

3. Methods

The objective of this thesis was to investigate whether different aspects of accident investigations influence learning potential. In this section, I will explain the research design and the method I used to answer my main research problem and research questions, and why I chose this procedure. A research design refers to a plan for getting from one place to another, or "the process that connects the research questions, empirical data and research conclusions" (Blaikie, 2000, p. 39). The research strategy, data collection and reduction, and analyses are also described in this section.

3.1 Research strategy

Blaikie (2000) emphasises the importance of choosing a research strategy, as our approach influences the whole study. He describes four main research strategies: inductive, deductive, abductive and retroductive. A researcher's assumptions will influence their work, making it important to acknowledge them and show transparency. The research in this thesis largely aims to investigate already existing theories on how different aspects of accident investigations influence the outcome. According to Blaikie's model, this closely resembles the aim of a deductive strategy, which goes from theory to data, before re-evaluating the theory considering the new data.

Danermark et al. (2002) use the same four categories but refer to them as modes of inference rather than research strategies. Inference in this context is understood as "various procedures, ways of reasoning and arguing applied when we in science relate the particular to the general" (Danermark et al., 2002, p. 75). The deductive strategy is in this view based on making logically valid arguments, and that knowledge of individual phenomena can be found through universal laws.

While this view is in line with Blaikie's, they have different views when it comes to abduction. Blaikie describes an abductive approach as developing a theory based on lay accounts. Danermark et al. (2002) describe it instead as the researcher interpreting data in the light of a specific context, or conceptual framework, and that this is not the only possible way to interpret it. This makes the researcher's conclusions and interpretations less rigorous and acknowledges that the results are influenced by the choice of theories and conceptual context. Although there are positive sides to this, it also puts the validity of the conclusions into question.

These two research strategies are quite different, but the adaption of both in this study can be of interest, as the limitations and positives of the two strategies complement each other. Danermark et al. (2002) consider the different strategies (or modes of inference) as complimentary when doing research, and not that the researcher must adapt just one. It will be useful to follow the guidelines of logical conclusions and theory testing of the deductive approach, while still being aware that our interpretation will not necessarily be the one and final answer, but a logical conclusion in the chosen theoretical context.

There are differences between the ontological and epistemological assumptions connected to the four research strategies. That is, the assumptions about reality, and how to gain knowledge about it (Danermark, 2002). The chosen deductive strategy relies on a realist ontology. This entails a view of the social phenomena existing regardless of the people involved – that there is an existing reality that determines the social behaviour (Blaikie, 2000). The epistemological view of deduction acknowledges that it is impossible to know whether a theory is true or not. Therefore, the conclusions drawn are not final but are involved in a search for the most reasonable and causal explanation based on rigorous testing. Since the chosen research strategy is closest fitting a deductive strategy, these are also the ontological and epistemological assumptions adapted in this study.

3.2 Data collection

There are many types of quantitative and qualitative methods available for data collection. Quantitative methods quantify data and tend to focus on a larger data sample than qualitative methods. Qualitative methods, on the other hand, focus on smaller samples and go more in-depth into the details (Blaikie, 2000). Each has its positive sides, as quantitative methods have more potential for generalization, while qualitative methods can give insight into details that simply is not possible with large sample sizes and quantified data.

Blaikie (2000) presents different types of commonly used data collection techniques, showcased in Table 1. Some techniques can be seen in both quantitative and qualitative versions, such as interviews and content analyses of documents. The main difference in these cases is how the data is coded – usually either in the shape of words or numbers – as well as the sample size and the depth of the topic focus. For example, a quantitative content analysis of documents could code the data into categories that are assigned numbers and counted. A qualitative content analysis could on the other hand identify phenomena and patterns in relationships in the text (Blaikie, 2000).

Table 1

Commonly used data collection techniques. Adapted from Blaikie (2000)

Quantitative	Qualitative
Observation: Structured	Participant observation
Questionnaire	Observation: semi-structured and unstructured
Interview: Structured	Interview: Focused or In-depth
Content analysis of documents	Content analysis of documents
	Focus groups/Group interviews
	Oral/Life histories

This thesis has a multi-modal approach by including both qualitative and quantitative analyses. The qualitative analysis was used to gain more insight and in-depth knowledge of the different investigation boards themselves. According to Neuman (2014), this is a descriptive research type, which can be used to provide a detailed picture and inform about the background and context of a situation. This knowledge supplements the quantitative data and provides more potential understanding as to why the two investigation boards' results might differ. The quantitative analyses were however helpful to give insight into modern investigation practices, and the consequences of different aspects of each report. Quantitative analyses were needed to see the overarching trend and avoid simply anecdotal results. Research that tests a theory's predictions, as well as determining which of several explanations are best, is a type of explanatory research (Neuman, 2014).

I chose a content analysis of documents for both the qualitative and quantitative analyses. Due to much data already existing available to the public about the two investigation boards, this was an accessible way to gather details about their different aspects. Participant observation, life histories and focus groups were not relevant to the topic. Interviews were also a less accessible option, without necessarily being able to provide more info about the organizations than document analyses could. Interviews with individual people from the two investigation boards would also provide their personal perspectives and be coloured by their roles in the investigations. A few questions regarding investigator expertise were still sent per email to each board, to supplement and confirm information from job listings.

Accident reports are documents by nature, so looking into these documents and their content was a natural way to gain insight into different aspects of the accident investigations. Even though the process of the investigation is not necessarily detailed in the reports, it gives valuable insight into the safety recommendations, which aspects their analyses focused on, and what their conclusions were regarding causal factors. Analysing trends in the reports might also be a more reliable way to determine accident models than relying on self-reports. Additionally, going through each report gives specific information about each investigation, rather than overarching statements about usual practices.

3.2.1 Qualitative analyses

All documents analysed are published by the two investigation boards, to ensure accurate information supported by the boards themselves. NTSB's Major Investigations Manual (2002), and NSIA's framework of their own NSIA-method (2008) provided the information needed regarding their investigation procedures. To further gain an understanding of their investigators' qualifications, I used the two boards' official web pages, statements, and job listings that mention their investigators' expertise and education. For job listings, I looked for listings for jobs that specifically entailed investigating railroad accidents, not their other sectors. NTSB had information on their website that could confirm information from their job listings, while NSIA replied to questions per email that gave further information regarding what expertise they seek.

3.2.2 Quantitative analyses

The investigation reports (N = 73) consisted of railroad accident reports from the last decade (2010-2021), to focus on contemporary accident investigations. The reports were found published on the boards' respective web pages. Only reports regarding derailment, collisions, personal injuries and fatalities, and smoke and fire accidents were included in the analyses, to ensure that both boards had investigations regarding the same types of accidents.

The railroad sector was chosen because none of the investigation boards put their main focus into this sector. The NTSB has the most expertise and experience with the aviation sector, and NSIA also originally started only in the aviation sector. Furthermore, NSIA has the least number of investigators hired for the railroad sector. Not choosing the sector they give the top priority could show potential weaknesses that would not otherwise be apparent. Furthermore, the railroad sector gives the boards more opportunity to aim recommendations at employees and management in comparison with highway accidents, due to highway accidents largely involving random individuals who are not connected to any relevant organisation. Thus, looking into railroad accident reports has a bigger potential to uncover a focus on human error.

Limited reports, preliminary reports, combined reports, and accident briefs were excluded. This resulted in 28 reports from the NTSB (38%) and 45 reports from the NSIA (62%). A complete list of all included reports can be found in Appendix A.

3.3 Data reduction and analyses

3.3.1 Data reduction

Both NTSB and NSIA have detailed, and complicated procedures outlined in their manuals. The different steps in their investigations were identified and categorized in main phases, rather than including all actors' responsibilities. The NSIA has largely identified and outlined these main steps themselves, so the NTSB manual's steps were categorized on a similar overarching level to achieve the best potential for comparison. Excessive information and details intended to guide involved actors were excluded, to highlight the underlying process.

The investigation reports went through a content analyses, identifying and quantifying key aspects. Reports were coded according to which board they belonged to, how many safety recommendations they resulted in, how many of these recommendations were specific, and how many recommendations were at different socio-technical levels. The number of fatalities was also included, as an indication of accident severity.

The determined probable causes reported were coded according to whether they related to human (1), technological (2), or organizational (3) factors. When the report determined that they could not find any probable cause, it was coded as "none" (0). Whether the reports investigated human, technological, and organizational factors were documented, and whether their safety recommendations were grounded in facts and observations.

Data was also collected on whether sequential, epidemiological, or systemic accident models were the most prevalent throughout the reports, determined by using the information in Table 2. The determined probable cause was specifically not used to determine the accident model, to see if the rest of the trends in the reports influenced the probable cause.

Table 2

Categorization of investigation reports' accident models

Model type	Investigation focus	Recommendation focus
Sequential	Simple and specific cause-effect links.	Eliminate and contain isolated causes.
Epidemiological	Barriers, and interactions between active failures and latent conditions.	Make defences and barriers stronger.
Systemic	System interactions rather than specific factors.	Monitor and control performance variability.

An example of a report determined to use a systemic model is NSIA's report number 2015/08. The accident entailed a tram colliding with a bus in an intersection. The bus driver had driven into the intersection despite a red stoplight, to give free passing to an emergency vehicle. Throughout the report, NSIA focused on how the combination of different events happening simultaneously created a situation that was simply too complex. Neither the single events nor the involved actors' behaviour was the centre of attention in the investigation. They also decided not to give any safety recommendations, as nothing was deemed able to reduce the complexity of the situation.

Had the investigators, in this case, used an epidemiological model, the report would show signs of a barrier focus and contributing factors, such as communication between emergency vehicles and other drivers, the stoplights, the road and the intersection's design. Furthermore, a sequential model would possibly focus on the bus driver's actions, without considering further context and contributing factors.

A report from NTSB (RAR-17-02) regarding an Amtrak train striking a backhoe with a worker inside explains that "the lack of consistent knowledge and vision for safety across Amtrak's management created a culture that facilitated and enabled unsafe work practices by employees". They continued to explain that the unsafe actions occurred because of inconsistent views of safety and safety management in the corporate structure, collaboration issues with unions, and safety not being a priority. Despite being quite different from NSIA's report, this is also determined to be a systemic accident model, as it focuses on system variability rather than specific contributing factors or human error. Instead of simply focusing on the worker's actions, they investigated systemic factors that influenced the employees' behaviour.

In contrast, NTSB was determined to use an epidemiological model while investigating a passenger train derailment (RAR-16-02). Their focus throughout the report was on different contributing factors, both active and latent failures. For example, they argue that the engineer accelerated the train too much, but that this was caused by distraction due to an emergency with another train, and an insufficient train control system. Their safety recommendations show a barrier focus, such as increased occupant protection.

To estimate the socio-technical level of the safety recommendations, Rasmussen's levels were adjusted and specified for the railroad sector. The resulting levels are presented in Table 3, in addition to their relation to the macro-, meso- and micro levels of analysis.

Table 3

Specified socio-technical levels used in analyses

Level of analysis	Number	Rasmussen's original levels	Adjusted levels	Specifications
Macro	1	Government	Government	Legislative regulations and rules
	2	Regulators / Associations	Industry	Non-legislative standards, certifications, and recommendations.
Meso	3	Company	Company	Company procedures and practice.
	4	Management	Management	Local management, training, and supervision.
Micro	5	Staff	Actors	Employee and passenger behaviour and attitude/awareness.
	6	Work	Surroundings	Equipment, technical design, railroad conditions, control systems.

Every safety recommendation was classified based on the nature of the content, rather than who it was addressed to. When the same recommendation was made several times to different parties, it was only counted once. To prevent overlap, reiterated safety recommendations were excluded.

To compare safety recommendation distribution between groups despite large differences in recommendation volume, the number of recommendations at each level was coded as the percentage of the total amount of recommendations in the report.

3.3.2 Quantitative analyses

To investigate the difference between the investigation board's procedures and methods, I used a Brown-Forsythe one-way ANOVA with investigation board as the group factor, and number of fatalities, number of safety recommendations, percentage of specific recommendations, and accident model as the dependent variables. Brown-Forsythe was used to correct for unequal sample sizes. Another ANOVA was run with investigation board as group factor, and the six socio-technical levels (government, industry, company, management, actors, surroundings), as well as probable cause as dependent factors. This analysis was intended to gain insight into the effect of the different boards' procedures on the outcomes of their investigations.

By using a Brown-Forsythe one-way ANOVA with accident model (sequential, epidemiological, systemic) and then with probable cause (human, technological, organization) as group factors, I examined how accident models and views on causality can influence conclusions and safety recommendations. The dependent variables were the six socio-technical levels of safety recommendations and number of recommendations in both cases, as well as probable cause when accident model was the group factor.

Pearson's correlation (two-tailed) was used to examine the correlation between each of the six socio-technical levels. This was done to investigate the relationship between the recommendations at each level, and whether a high percentage of recommendations on one level influences the percentage on any of the other levels.

The main analyses in this study were ANOVAs, due to the nature of the data collected. Most of the variables were nominal, meaning that they were categorical with numbers representing different categories. The number values thus had no actual numerical meaning other than representing groups, and analyses relying on scale variables would not be of use. This includes regression analyses and further correlations. ANOVAs are useful for showing differences between groups, which was the main focus of this thesis. Not only differences in the two investigation boards, but also differences in results between accident models. A correlation test was however also relevant, as it is good at showing the relationship between different scale variables. This was ideal to further understand the relationship between the socio-technical levels, and whether two levels had a lot of overlap, or if focus on one level meant less focus on another.

All analyses were done in IBM SPSS Statistics, version 26.0.0.0 (IBM Corp, 2019). The results were determined significant at a value of $p < .05$ for all analyses.

3.4 Research ethics

The research in this thesis is conducted on the content of published and public documents, without pointing out individuals or personal information. Because the study does not involve humans, healthcare information or/and human biological tissue, it was not necessary to get the research approved by a regional research ethics committee. I have done my best to maintain transparency, quality, and accountability throughout this thesis.

3.5 Generalization, validity, and reliability

Two important indicators of quality related to research are validity and reliability. Validity involves measuring what you intend to measure, while reliability entails that the same method should be able to reproduce similar results if done repeatedly (Golafshani, 2003).

A common way to test for reliability is the test-retest method (Neuman, 2014). Doing your test, content-coding, questionnaire, or experiment at two different times gives you the possibility to see if the results remain stable. To verify that my quantitative data were reliable, I went through the investigation reports and coded the data based on my criteria twice, one month apart. This was to make sure that my results were consistent and that my criteria were specific enough for me to get the same results both times. Neuman (2014) further suggest four ways to improve reliability: clearly conceptualizing constructs, using a precise level of measurement, multiple indicators, and a pilot test. Because of this, I consciously tried to make my measurements and categories as specific, distinct, and concrete as I could. I also used several indicators to determine investigator expertise and looked at multiple aspects of the investigation reports before determining accident models.

Validity can be split into internal and external validity. Internal validity is related to drawing correct conclusions about the research subject, while external validity concerns the representativeness of the results (Blaikie, 2000). Without inner validity, the results will not answer what it claims to, but without external validity, the results will not have any usefulness or meaning outside of the context of the study itself.

For qualitative research, it is suggested that using methods such as triangulation, respondent validation, and clear detailing of every step of the research process might help improve validity (Malterud, 2001). Every finding should be questioned instead of taken for granted, as well as considering the effect of both context and bias. I have carefully considered the context and bias related to the collected qualitative research, and reached out for interviews over email, getting a response from a department director at NSIA. Furthermore, I am not looking to find qualitative results that are representative outside of the two investigation boards in question.

Related to quantitative research, it is common to also distinguish between face validity, content validity, criterion validity, and construct validity (Neuman, 2014). Face validity entails how much a measurement 'makes sense' as a measure for that specific construct, at face value. Content validity, on the other hand, concerns whether the measure manages to represent the whole content of the concept, or if it only represents a part of it. Criterion validity refers to using an already existing measure that is deemed valid as a criterion or reference point. Lastly, construct validity refers to a measure with multiple indicators (Neuman, 2014). If all indicators operate consistently, the construct validity is estimated to be higher.

Neuman (2014) points out how it is not possible to have absolute confidence about validity, but that some measures are more valid than others. I have tried to achieve higher validity by basing my measurements, indicators, and categorisations on acknowledged theories, increasing both face validity and criterion validity. I have also researched different sides of learning potential with different measurements, to be able to account for more of the content of the learning potential concept. I am however not claiming to account for the entirety of the learning potential concept. Furthermore, I have used several indicators to measure different aspects, such as expertise. By checking for consistency across interviews, different sources and documents regarding expertise, I attempted to increase the study's construct validity.

Generalizability is closely related to both reliability and validity, and Golafshani (2003) argues that high validity may lead to generalizability. Generalizability regards the ability to apply findings to wider groups and circumstances outside of the research context itself (Golafshani, 2003), and can be related to external validity. Generalization is usually something more of a concern or a goal when it comes to quantitative research. Qualitative research does not necessarily seek findings that are generalizable for a larger population, but rather a deeper understanding of a specific phenomenon. My qualitative research does not intend to say anything about all investigation boards, but rather give insight into exactly how the NTSB and the NSIA differ. Generalization is thus a more relevant issue for the quantitative data. Can the findings be applied to other accident investigations than those studied?

To answer this question, one can distinguish between statistical generalization and analytic generalization. Statistical generalization entails having a sample that can represent a population the researcher wants to say something about (Polit & Beck, 2010). My sample included all recent reports within similar categories from the two investigation boards. Thus, the sample should be representative of the two boards' most recent investigations and be statistically generalizable to the two boards' contemporary practices. The results were not intended to be statistically generalizable for all accident investigations, as this would need reports from a much larger pool of investigation boards.

Analytic generalization is often used in theory-driven quantitative research like this and can be used when trying to generalize from a particular group to a broader one. For ideal analytic generalization, the researcher uses conceptualizations of processes (Polit & Beck, 2010). To achieve this, a distinction was made between information that is relevant to all (or many) accident investigations, rather than information unique to each one. Then, my results could be generalized to a theory by identifying evidence that supports the theory, in line with Polit and Beck's criteria (2010). In this way, I have aimed for an analytic generalization from my results to accident investigations in general.

3.6 Methodological reflections

Throughout the process, I have had to make a lot of decisions, especially regarding how to code large amounts of information into simple categories. The categorizations are not an exact science, and the results from all analyses are influenced by my choices, my bias, and my understanding. I have tried to be as transparent as possible, so other researchers can be able to replicate the study and understand my thought process. This will increase reliability if future research finds similar results.

Furthermore, the research in this thesis is affected by the reports and the documents it is built on, especially regarding the nature of the investigation boards. If the investigation boards are not transparent and honest about their expertise, processes, and methods, my understanding of this will also be skewed, which in turn affects the results and conclusions. To avoid word-of-mouth, misunderstandings, and assumptions made by outsiders I chose to focus on the two investigation boards' statements and documents. I also maintained an awareness while analysing their reports, to make sure their statements were in line with the products they provide.

In terms of validity, I have done my best to analyse and categorize the content of the reports into their correct categories in line with the theories applied. It was a challenge to categorize safety recommendations at their correct socio-technical level, especially when they were addressed to different specific organizations or government agencies but contained advice regarding lower levels. However, I believe the potential safety measures that come out of the recommendations are the important aspect, rather than who the recommendations are addressed to.

While safety recommendation levels are used as an indicator of learning potential, it is not alone a valid indicator of learning potential as a whole. Learning potential is more nuanced than one variable. Instead, safety recommendation levels give insight into one aspect of learning potential and can still give valuable information that can help improve learning from accidents.

An additional challenge was to categorize long and elaborate reports into one of three accident models. Models are simplified understandings and do not represent reality accurately, so grouping real reports in line with such simplified models can have validity problems. While there were traces of several models in the same report, I saw prominent focus differences and chose to categorize the reports based on this main focus.

4. Results

In this chapter, the results from both the qualitative document analysis and the quantitative analyses will be presented. The results from the document analysis gaining insight into the investigation board's basic characteristics, methods, expertise, similarities and differences are presented first, as they provide a foundation for interpreting the quantitative results. Then, both descriptive statistics and the results from the quantitative analyses are presented.

4.1 Investigation boards

4.1.1 Standard procedures

The NTSB's Major Investigation Manual (2002) refers to aviation accident investigations, but investigations within the other modes of transportation are conducted in similar ways. The standard investigation process for the two boards is summarized in Table 4.

Table 4

Comparison of the standard investigation process

Stage	NSIA	NTSB
1. Preparation	- Accident inspector on duty	- Go Team on call
2. Notification and initial response	- Decision to investigate - Composition of team - On-scene briefing - Contacting affected parties	- Decision to investigate - Composition of Go Team - On-scene briefing - Contacting organizations
3. On-scene activities	- Gathering facts - Security investigation	- Gathering information and taking field notes - Off-scene materials laboratory examinations
4. Post-on-scene activities	- Assessing need for recommendations - Draft for report sent to affected parties - Adjustments to report in accordance with feedback - Final report	- Public hearings - Report draft circulated to all parties for feedback - Conclusions and recommendations made in public board meeting based on the draft report - Final report

Overall, their procedures are quite similar. One of the more apparent differences is NTSB's use of a so-called Go Team. The Go Team consists of different specialists depending on the probable scope of the investigation, the magnitude of the tasks, the number of injuries/fatalities, type of machinery involved, previous accidents, location, the extent of surrounding damage, weather, public interest, and specialist workload. This is similar to NSIA's investigation team, which is also established based on similar considerations. The NSIA does however not have an established team on call at all times. Instead, they have an inspector on duty and deploy more inspectors if needed. Their rail sector has six employed accident inspectors.

Another apparent difference is NTSB's focus on publicity and transparency. Their manual contains a plethora of information regarding public hearings and public board meetings, and even their conclusions and recommendations are often drawn in such a public meeting. For this reason, conclusions and recommendations are not included in their drafts – the draft is instead used as the basis to draw the conclusions. NSIA on the other hand seems to draw their conclusions more actively while writing the draft, and also contain more privacy during the process.

The NSIA investigative team writes the reports themselves based on their findings. NTSB, on the other hand, has a separate division called the Writing and Editing Division, and an employee from this division writes their reports based on the Go Team's findings. Furthermore, NTSB has a separate office for Safety Recommendations and Communications, coordinating recommendations and conclusion suggestions from other organizations and the different offices. NSIA's investigative team thus has more responsibilities and a bigger influence on the outcome of the reports than NTSB's Go Team.

4.1.2 Investigation methods

The NSIA has developed their own specific investigation methodology named the NSIA method (Previously named the AIBN method) (AIBN, 2018). This method incorporates elements from three different popular analytical tools: STEP, AcciMap and MTO, described in part 2.4. This resulted in seven steps the investigators go through in their analysis process. Stage one to three concerns what happened, stage four to five examines why the accident happened and stages six to seven concerns how to prevent new accidents (AIBN, 2018):

1. Clarifying the sequence of events and circumstances
2. Identifying local safety problems

3. Barrier analysis
4. Identifying risk factors
5. Assessing causality and importance
6. Considering systemic safety problems
7. Assessing the need for safety recommendations

The NTSB is considered to have introduced multi-linear events sequencing concepts, explained in part 2.4 (Sothivana & Siddiqui, 2015). These types of diagrams are still found in their reports today, and as one of the first users of this type of accident modelling, it can be assumed that this is a common tool used in NTSB's accident investigations. The STEP analysis used by NSIA is an example of such multi-linear event sequencing. However, unlike NSIA, the NTSB does not have an overview of how exactly to conduct the investigations. Instead, they use more of a checklist for things their teams have to investigate, rather than procedure steps (NTSB, 2002). Their Go Team is split into groups, in charge of investigating different elements which can be seen in Table 5. NSIA does not have such a checklist but is freer to make delimitations and focus their investigation. A comparison of NTSB's most common investigative features is also included in Table 5.

NTSB's checklist is much more concrete than what NSIA operates with, but it is also much more technical and neglects to involve blunt-end factors. They are however expected to investigate all influencing factors connected to the elements on the list, for example, rules and regulations regarding structures or training of employees.

Their investigations show large overlaps, but also some differences. While both boards investigate injuries to everyone involved in the accident, NTSB has a more in-depth investigation into impact forces, evacuation, emergency response and rescue efforts. NISA simply investigates the injuries, while not evaluating the emergency response. NSIA in return investigates the organization's emergency preparedness and protocols, while NTSB does not have standard routines for this. Laws and regulations tend to be in a separate section in the Norwegian board's reports, while the United States' board includes relevant laws and regulations relating to each section instead, meaning they both include these factors.

The NSIA also includes research into hazardous materials, while NTSB does not. Investigations on transportation of hazardous materials have their own sector at the NTSB, which explains why it is not included in the railroad sector. It is however important to note that these are not NSIA's standard procedures, but investigations that they often optionally include. NTSB's list covers their standard procedures, but that does not mean that they won't conduct other investigations if necessary.

Table 5

Comparison of investigation elements

Category	NSIA	NTSB
Operations	Workplace rules and procedures	Employee's duties
Infrastructure	Accident scene and structures	Accident scene and structures
Materials	Damages to involved materials such as types of trains, trams, machines, or other vehicles	Engines, engine accessories
Systems	Control systems, electrical systems, instruments, and other associated elements	Control systems, electrical systems, instruments, and other associated elements
Traffic control	Traffic control, communication channels, traffic direction and signal systems	Traffic control services and transcripts of communications
Weather	Weather data from the broad area around the scene	Weather data from the broad area around the scene
Human performance	Employee information, certifications, health, training, schedules and other factors that could influence performance	Performance of the crew/employees and factors that could influence their performance, such as fatigue, training or work environment
Hazardous materials	Involved hazardous materials and their properties and potential dangers	-
Survival factors	Personal injuries	Impact forces, injuries, evacuation, community emergency response and rescue efforts
Safety management	Emergency and preparedness protocols, manuals, and use of emergency calls	-
Laws and regulations	Relevant laws and regulations	Relevant laws and regulations are included in each category, not a separate point

4.1.3 *Expertise*

To gain insight into the two investigation boards' investigators' expertise, I looked into both their official statements and recent job listings. Department director Ronny Ruud from NSIA also gave information over email regarding expertise. NTSB had an active listing for the railroad sector, as well as having some basic information on their web page.

On NTSB's webpage (n.d.), they inform that their railroad accidents are investigated with locomotive engineers, signal system specialists and track engineers in charge of the working groups. In a job listing from the Federal Railroad Administration (2021), they look for a railroad safety specialist to help investigate railroad accidents with the NTSB. Their listed qualifications include knowledge of the railroad industry, general safety and health principles and practices and knowledge of railroad accident investigation techniques. They ideally seek a technical expert within one of five railroad safety disciplines – hazardous materials, operating practices, locomotive power and equipment, track, signal and train control.

The qualifications wanted by the NSIA are similar, based on a job listing as an emergency inspector in the railroad sector (Statens Havarikommisjon, 2020), which their investigative teams are composed of. The position is listed as an engineer position, and their needed qualifications are a minimum of five years of train operation experience, a master's degree in technical or safety-related fields, and experience with accident investigations or other safety-enhancing work. These qualifications are confirmed by a department director from NSIA. Many of their inspectors have both theoretical and operational expertise, and they compose their teams to cover a variety of expertise (R. Ruud, personal communication, May 2021).

4.2 Quantitative analyses

After the data collection, it became apparent that both boards always grounded their safety recommendations in facts and observations. In fact, the NTSB is required to do so by the guidelines in the U.S. Code Title 49 (§ 1117). Furthermore, they researched human, technological, and organisational factors in all their reports. Thus, there was no reason to conduct any further analyses with these variables, as there were no differences to examine.

A one-way ANOVA investigating the different characteristics between the NTSB and the NISA showed significantly fewer fatalities in NSIA's reports than in NTSB's reports the past 10 years. There was also a difference in the number of safety recommendations, where NTSB tended to give a higher number of recommendations than NISA. The descriptive statistics are presented in Table 6, while the results of the ANOVA can be seen in Table 7. Not only did NTSB give more recommendations, but they also gave more specific recommendations. As much as 99% of NTSB's recommendations were specific, while 80% of NSIA's reports were.

Table 6

Descriptive statistics of aspects of NSIA and NTSB's investigation reports

		M	*SE*	95% CI
Number of safety recommendations	NSIA	1.33	0.12	[1.09, 1.58]
	NTSB	8.21	1.40	[5.34, 11.09]
	Total	3.97	0.67	[2.64, 5.30]
Fatalities	NSIA	0.27	0.12	[0.12, 0.42]
	NTSB	2.61	0.93	[0.69, 4.52]
	Total	1.16	0.38	[0.40, 1.92]
Specific recommendations (%)	NSIA	65.93	6.70	[53.43, 79.43]
	NTSB	99.29	0.71	[97.82, 100.75]
	Total	78.72	4.54	[69.67, 87.78]

Note: M = Mean, SE = Standard Error, CI = Confidence Interval

Table 7

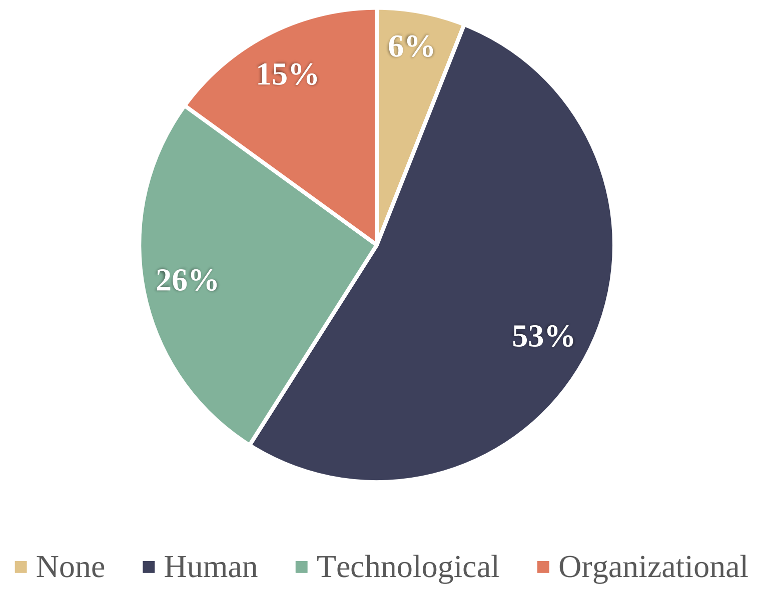
Effect of investigation board on different aspects of accident reports

	Sum of Squares	df	*F*	*sig*	Partial η^2
Number of safety recommendations	817.23	1, 71	38.26	.000	.350
Fatalities	94.55	1, 71	10.03	.002	.124
Specific recommendations (%)	19208.71	1, 71	15.28	.000	.177
Probable cause	0.84	1, 71	1.26	.266	.017
Accident model	0.02	1, 71	0.15	.705	.002

There were no significant differences between the boards' use of accident models (See Table 7). The distribution between the accident models was 0% sequential, 81% epidemiological and 19% systemic, showing a large preference for epidemiological accident models.

Figure 11

Combined distribution in concluded probable cause



The boards also showed no significant differences in whether they assign the probable cause to human, technological or organizational factors, also shown in Table 7, and the distribution of the causal factors is illustrated in Figure 11. However, they had a significant difference in their distribution of safety recommendation between the socio-technical levels. Their respective distribution is displayed in Figure 12, with statistics in Table 8.

NTSB had a significantly higher percentage at the government and industry level, while NSIA had more recommendations at the surroundings level. NSIA did not have a single recommendation directed at the government, while NTSB had 9% in this level. The effect was also highest at the government and industry level, with 19% of the variance being explained by board at the government and 14% at the industry level. In comparison, under 8% of the variance in surroundings was explained by investigation board.

The three remaining levels – company, management, and actors – were not significantly different.

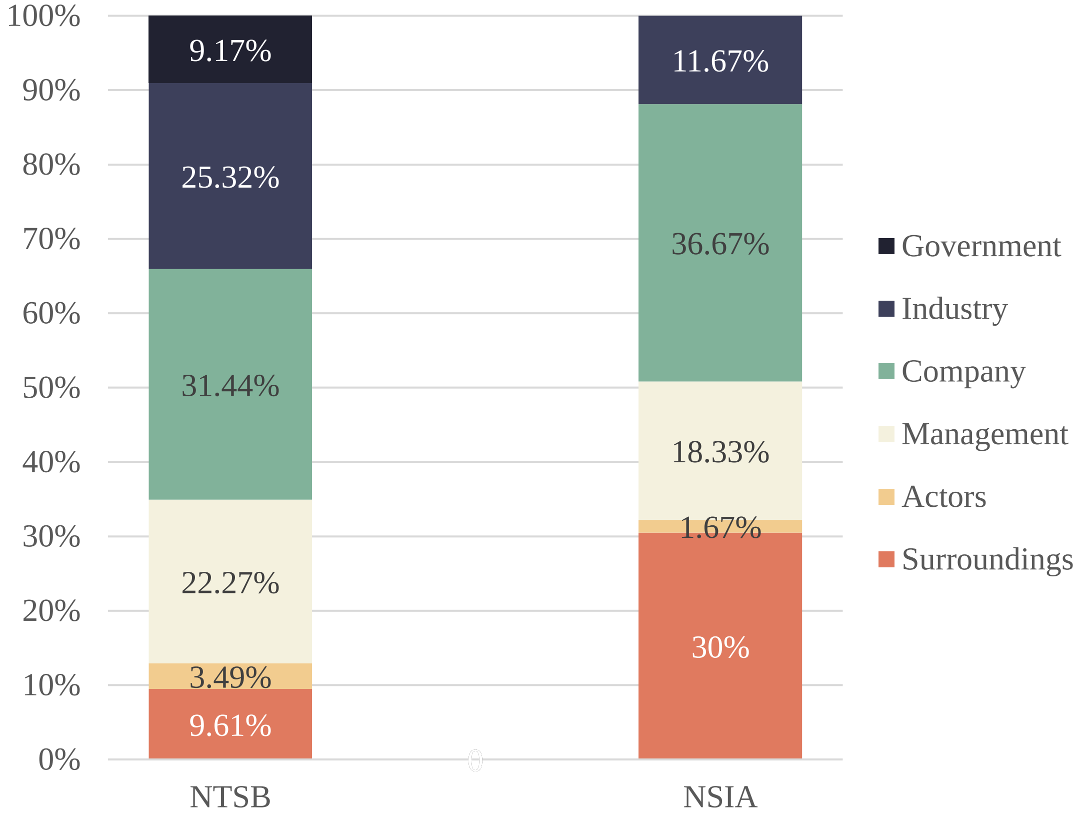
Table 8

Differences in socio-technical level based on investigation board.

	Sum of Squares	df	*F*	*sig*	Partial η^2
Government	2984.04	1, 71	17.08	.000	.194
Industry	6751.86	1, 71	11.64	.001	.141
Company	0.015	1, 71	0.00	.995	.000
Management	45.42	1, 71	0.06	.802	.001
Actor	32.57	1, 71	1.23	.271	.017
Surroundings	6326.89	1,71	5.69	.018	.077

Figure 12

Distribution of recommendations at different socio-technical levels



The results of a one-way ANOVA showed that there were significant differences in the probable cause determined while using different accident models ($F[1, 71] = 61.30$), $p < .001$).

The results in Table 9 show that as much as 46% of the variance in the probable cause can be explained by the use of different accident models.

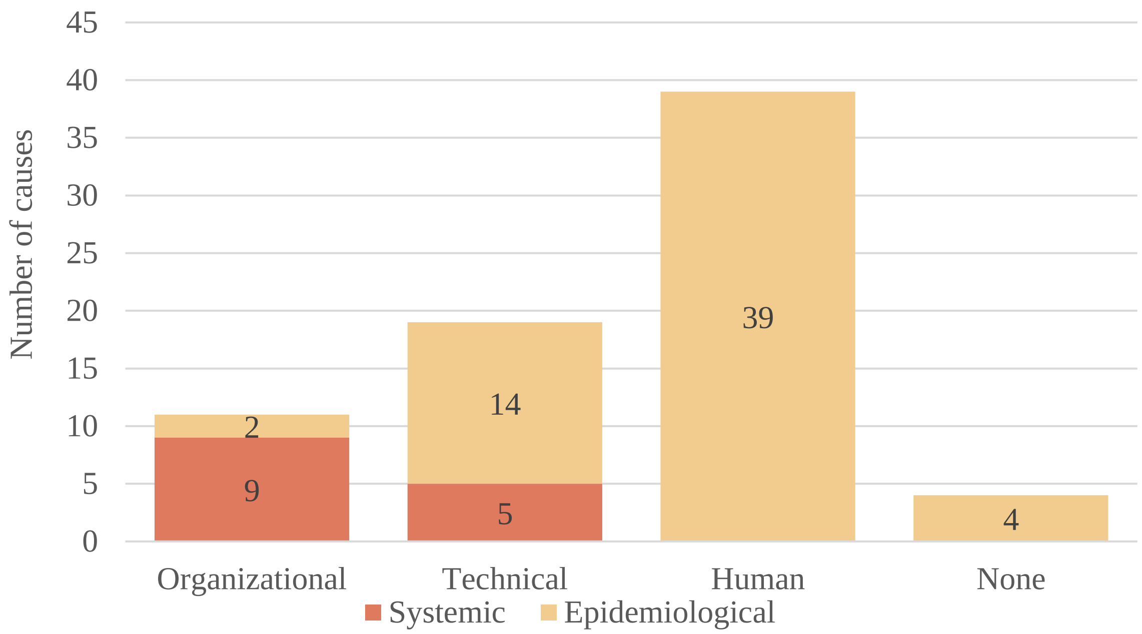
Table 9
Differences in probable cause based on accident model.

	Sum of Squares	df	Mean Square	*F*	*sig*	Partial η^2
Intercept	170.35	1	170.35	467.13	.000	.868
Probable cause	22.35	1	22.35	61.299	.000	.463
Error	25.89	71	0.37			

While using systemic accident models, the reports more often conclude that the cause is organizational factors, while epidemiological accident models blamed human factors the most. As much as 64% of the systemic models determined that the cause was organizational factors, and the remaining 36% were all determined to be technical factors. In contrast, epidemiological models blamed human factors 66% of the time, technological factors 24% of the time, and organizational factors only 3% of the time. The remaining reports, 7%, said that they could not determine a cause. The distribution can be seen in Figure 13.

Figure 13

Distribution of causal factors by accident model



Despite influencing probable cause, the accident models did not have any significant differences in socio-technical levels associated with the safety recommendations. They also showed no difference in the number of recommendations when corrected for unequal variances (p = .20) Similarly, probable cause showed no difference in the number of safety recommendations (p = .07), while also not having any influence on the socio-technical levels.

To investigate the relationship between the six socio-technical levels, Pearson's correlation coefficient was used. The results showed a negative correlation between the industry and management level (r = -.23, p = .04), between company and surroundings (r = -.43, p < .001), and between management and surroundings (r = -.24, p = .04). Neither the government nor actor level correlated with any of the other levels. Furthermore, all significant correlations were negative, meaning that if there are more recommendations made on one level, it reduces the number of recommendations in the other. Consult Table 10 for the statistics.

Table 10

Correlation between recommendations aimed at the six socio-technical levels

	1	2	3	4	5	6
1. Government		.011	-.169	-.069	-.045	-.122
2. Industry			-.199	-.231*	-.009	-.221
3. Company				-.170	-.012	-.432**
4. Management					-.053	-.244*
5. Actor						-.013
6. Surroundings						

Note: ** = Significant at p < .001, * = Significant at p < .05

5. Discussion

In this chapter, I will discuss my research findings with the help of the research questions considering the presented theoretical framework. The research questions concerned the different influences that accident models, investigator expertise and investigation methods can have on the outcomes of accident investigations. Ultimately, these questions are a tool to answer my main research problem: *How do different aspects of accident investigations by the NSIA and the NTSB influence learning potential?* The implications for the main research problem will also be discussed.

5.1 How can investigation procedures and methods affect the outcome of investigations?

5.1.1 Standard procedures

At face value, there are both obvious similarities and differences between the NSIA and NTSB. NTSB is a much larger organization than NSIA, with approximately 400 employees up against NSIA's 54 employees. This has caused NTSB to have a larger and more elaborate organizational chart (see Figure 2 and Figure 3), but ultimately their structure is similar at the bottom levels. They have both chosen to split the different modes of transportation into separate departments, with their respective investigators and directors.

The standard investigation process for the two teams is quite similar. An obvious difference is the number of investigators involved, and NSIA's use of several teams to investigate different parts. Further differences became most apparent regarding the post-on-scene activities (see Table 4), where the NTSB has more public processes with recommendations and conclusions drawn in a public board meeting. Their report draft includes no indications of probable cause or recommendations. In contrast, the NSIA adds conclusions and recommendations more continuously while writing the draft and has this included when sending the draft to affected parties.

As a result, NTSB allows the different boards, affected parties, and organizations to give recommendations and input without being primed or biased by already suggested conclusions and recommendations to a bigger extent than the NISA's process allows. Thus, NTSB's approach might open for a wider perspective. If this has an effect, it would be expected that NTSB has a more even distribution between the socio-technical levels in their recommendations. The results show exactly that: NTSB has significantly more recommendations at the top two levels (government and organization) while having fewer recommendations at the lowest level (surroundings). In turn, their recommendations are more evenly distributed between all the socio-technical levels. Furthermore, NSIA had no recommendations at the government level at all. This could be a sign that the NTSB has achieved a higher learning potential in this aspect.

The fact that NSIA often draws their conclusions and recommendations in public could possibly steer their recommendations to higher levels, as it is seen as more ‘proper’ and in line with modern thinking to aim blame away from individuals and lower levels. However, if this were the case, one would expect that the difference in deemed probable cause would also be significantly different, where NTSB had given more blame to organisational factors. This is not the case, and both boards assign the probable cause to human factors in over half of the reports (53%) while only assigning it to organizational factors 15% of the time (see Figure 11).

An alternate explanation for the difference in socio-technical levels is differences in responsibilities. The NTSB seem to have standard procedures to aim their recommendations to federal and state agencies, transportation providers and manufacturers. Thus, they aim recommendations at higher levels by default. No matter which other methods, perspectives, or expertise they have, this could overrule all of it. NSIA on the other hand, aims their recommendations at the Norwegian Railway Authority, railway providers and manufacturers. To give recommendations to the government level, the Norwegian investigation board would likely have to get adjusted standard procedures and responsibilities.

Further relevant procedure differences were found in the boards’ responsibilities. For instance, the NSIA is instructed to make recommendations *without underlying concrete solutions* (Samferdselsdepartementet, 2020, p. 4). This is reflected in the results regarding whether the recommendations are specific or not: 99% of NTSB’s recommendations were specific, while 80% of NSIA’s were, meaning that NTSB goes more into detail. This might also explain why NTSB has more safety recommendations (8.2 on average, compared to NSIA’s average of 1.3), as they try to give much more specific solutions and guidance.

The American approach gives more guidance for how to avoid or reduce future accidents or negative consequences, and thus also less room for error and misunderstandings. The recommendations are arguably there to give guidance and solutions. However, the Norwegian approach of being a less concrete solution gives the responsibility to the involved actors and organizations to solve their own problems. The recommendations work more as a guide for where to look. This could increase learning potential through working on, taking responsibility for, and understanding their own dilemmas. This could arguably increase double-loop learning.

Despite having firm rules regarding how concrete recommendations can be, the NSIA has expressed a wish to start giving more safety recommendations, as part of their goal to do "deeper" investigations (SHT, 2019). Despite having an increase, they still do not give safety recommendations after all their investigations, and rarely give more than one or two recommendations. Does this indicate that the NTSB does deeper and more thorough investigations? It is likely, especially considering their larger resource pool, that they can do so. However, there might also be other factors contributing to this difference in the numbers of safety recommendations.

The NTSB appears to investigate exclusively more serious accidents than the NSIA does, which might naturally result in more recommendations being necessary. This is apparent through their instructions, as NTSB investigates accidents where there is a fatality, substantial property damage, or a passenger train is involved. The NSIA on the other hand also investigates 'serious incidents'. During the data collection, some of NSIA's reports had to be excluded because they entailed near-accidents, where no damage was actually done, which NTSB does not investigate. However, even while excluding these reports, the results show that NTSB investigates reports with significantly more fatalities, indicating a difference in severity.

Differences in the accidents themselves, such as severity, might have implications for other aspects than just numbers of safety recommendations. If the two boards systematically choose to investigate different accidents, or if the two countries simply experience different types of accidents, this might explain further differences in results between them. For instance, it might be deemed more necessary with recommendations at higher socio-technical levels in response to more severe accidents.

On the other hand, the choice of which types of accidents to investigate could also have implications for learning potential. For the NSIA, more accidents will qualify for investigations, due to their lower bar. This increases the numbers of investigations done, increases the learning potential, and increases the possibility of doing risk-mitigating actions *before* a more serious accident occurs. Despite not being able to include all NSIA's accident reports in the analyses, the NSIA still had more accident reports than the NTSB did (62%). However, with higher numbers of accident reports (and as a smaller investigation board), their investigations will inevitably have to sacrifice some depth for them to get through all the reports required, which again could lower learning potential.

5.1.2 Investigation methods

NTSB's use of multi-linear events sequencing like STEP has its advantages. It is a suitable method for complex accidents where several activities happen at the same time, due to the use of multiple lines. It can be argued to be most suitable for epidemiological accident models, which is also the accident model that both boards seem to use the most – as much as 81% of the time. However, STEP lacks the barrier analysis used in an MTO analysis, as well as the interactions between socio-technical levels from AcciMap. Using such a method might also increase their use of epidemiological accident models.

NSIA's conscious use of all three of these methods gives them the advantage of having the strengths from each one. With the help of all methods, they go through a list of seven steps in their analyses. Regardless of which accident models their investigators might be influenced by, and their perspectives and experiences, they will always still consider local safety problems, risk factors, barriers, and systemic safety problems. Their investigation method then assures them that no matter who the investigators are, they will consider human, technological, and organizational factors, as well as both blunt- and sharp-end factors. This could increase the NSIA's learning potential through double-loop learning.

Rather than having a list of procedures, NTSB operates with a checklist of elements that need to be investigated. This checklist is largely consisting of sharp-end factors, such as human performance, infrastructure, and materials. However, for each sharp-end factor they investigate, they are expected to investigate all blunt-end factors that might influence them, such as laws and regulations and organizational factors. Due to their list focusing mainly on sharp-end factors, the investigators will have more influence on which blunt-end factors get considered. However, that is similar to NSIA's method: they are told to analyse systemic safety problems, but they have to determine which systemic factors are relevant.

From the quantitative analyses' results, it is evident that their use of different investigation methods did not affect what they deemed the probable cause and had no effect on their use of accident models since there were no significant differences between the two boards.

Furthermore, both methods provided sufficient guidance for both boards to always investigate human, technological, and organizational factors. The NSIA's inclusion of AcciMap and the following consideration of socio-technical levels was expected to lead to increased use of systemic accident models, and organizational factors to be deemed the probable cause more often. Since this is not the case, NTSB's method either leads to the same consideration of systemic factors, or NSIA's use of STEP and MTO gives similar increased focus on barriers.

The high use of epidemiological models might also reflect society's way of thinking of accidents in modern times. Sequential models used to be more common before epidemiological models took over, and even though systemic accident models have been on the rise and a popular option in theories and literature, the practical use might still lag behind. It takes time to shift the way we think about causation and how accidents happen, and these results could reflect that. Even though both investigation boards' methods include considerations of systemic factors, they are included as specific latent conditions in line with epidemiological thinking.

NSIA's use of all three accident models was hypothesized to influence the safety recommendations' socio-technical levels. The consideration of interacting socio-technical levels through AcciMap, barriers through MTO, and interactions between actors through STEP were expected to give a more even distribution between the recommendations' levels. The results showed the opposite: NTSB had a more even distribution between the socio-technical levels than NSIA had. This indicates that the chosen investigation methods might not have an as large effect on the outcome of accident models as other influences, such as the standard procedures discussed earlier. If other influences tip the scale in other directions, then it could outweigh any influence the methods might have. Alternatively, the accident method does not have any effect at all.

Another potential explanation could be that the investigation boards' methods are not sufficiently different to cause a significant effect. After all, both boards utilize the STEP model, as well as additional steps to make sure systemic factors are considered. This causes more of an overlap between their methods, rather than being two distinct approaches.

5.2 How can the investigation boards' expertise influence the investigation?

Even though the differences between the investigation boards can be explained by their procedures, it may not be the only influencing factor. There are other differences between the investigation boards, and it is hard to tell which factors play a bigger role and whether any of the factors are simply confounding variables. In this case, the use of expertise might be another relevant factor.

The NTSB's Go Teams execute the investigations and consist mainly of technical experts. Safety knowledge is also a requirement, but their safety knowledge is generally regarding technical aspects of the railroad industry. Similarly, the NSIA requires a minimum of five years of train operation experience, focusing their team's expertise on technical and operational knowledge. They do however also require experience with accident investigations or safety-related work. Additionally, their experts are qualified with a master's degree in safety-related fields, not only technical fields. Through communications with NSIA, they make it clear that they build their investigation teams to consist of a variety of qualifications with both technical and theoretical backgrounds (R. Ruud, personal communication, May 2021). NTSB on the other hand does not qualify people with an academic background in purely safety-related fields and focus more strongly on technical and operational backgrounds.

Based on Cedergren and Petersen's (2011) research, technical and operational expertise should increase the focus on micro-level factors. This is equivalent to the 'actor' and 'surroundings' socio-technical levels in this thesis. The results reflect this tendency in NSIA's reports, as they have the second most recommendations at the surroundings level (30%), which includes any technical recommendations. However, under 2% of their recommendations are at the actor level. This could reflect the new view, where human error is considered a symptom rather than a cause (Dekker, 2014), and safety interventions aimed at involved actors would not be considered very fruitful. NTSB similarly has the least number of recommendations at this level.

NTSB does not show the same large focus on the surroundings level, thus not reflecting the expectation that technical expertise increases focus on technical recommendations. However, there is an important difference between the NTSB and NSIA that might provide an explanation, resulting from the boards' organization discussed earlier. NSIA's investigation teams 'on the ground' are responsible for far more than the NTSB's Go Teams are – such as making the conclusions, recommendations, and writing the report. The NTSB has a Writing and Editing Division that writes the report, as well as an Office of Safety Recommendations and Communications (See Figure 3). This office coordinates all suggestions for safety recommendations and conclusions given to them by different involved parties. Thus, the Go Team does not draw conclusions, give recommendations, or write the report, they simply provide factual and technical information regarding the accident.

With this difference in expertise when it comes to the report, safety recommendations, and conclusions, the results from the two boards should be significantly different to be in line with Cedergren and Petersen's (2011) predictions. They suggest that larger cross-sectoral boards can minimize the effect of the investigators' expertise. Both NTSB and NSIA's investigations should be influenced by their investigators largely having technical backgrounds, but since the NTSB's recommendations and conclusions are not drawn by the Go Team, the influence of the investigators' expertise should not be as visible in the NSIA reports as in the NTSB reports. The other involved offices' expertise might even have a larger influence, due to them being responsible for the outcome of the investigations.

This is in line with the results in this study: the NSIA has a significantly higher number of recommendations at the surroundings level, while the NTSB has significantly higher numbers at both the government and industry level. This adds additional support to the theory that a larger cross-sectoral investigation board with more diverse competencies might increase learning potential and widen the focus of the reports (Cedergren & Petersen, 2011; Stemn et al., 2020; Svenson et al., 1999). The variance at the government level and industry level explained by investigation board are also rather high, as much as 19% and 14% respectively. The effect on the surroundings level is lower, and less than 8% of the variance is explained by investigation board. The difference between the investigation boards thus explains a bigger difference at the higher levels. This could be explained by them both having technical expertise, while the NSIA is lacking the competencies that come with a larger cross-sectoral board.

The two boards had no significant differences at either the company, management or actor level. This indicates that regardless of any differences in expertise (or methods and procedures), they still aim for a consistent amount of recommendations at these levels. These numbers are rather high for both the company and management levels and low for actor levels. This could again reflect modern thinking of accidents, and how human error is seen as a symptom (Dekker, 2014). It also reflects Leveson's (2004) findings that the most important factor in the occurrence of accidents is related to management and commitment to safety culture. To increase learning from accidents, aiming recommendations at these levels would be productive if Leveson is correct.

A Pearson's correlation on the socio-technical levels showed that the industry and management levels were negatively correlated, as well as surroundings being negatively correlated with both the company and management levels. Increasing focus on the industry level decreases focus at the management level and vice versa. Similarly, increasing focus on surroundings decreases recommendations at company and management levels.

These negative correlations show that these socio-technical levels take focus away from each other, rather than complimenting each other or simply being added in addition to each other. The government level, on the other hand, is not correlated with any other of the socio-technical levels. The way NTSB adds recommendations at this level shows that it is done as an addition, without removing any focus from other levels. Their government level recommendations are often regarding the same solutions and recommendations as they give to lower levels, except they ask for changes in laws to increase the reach and the effect of the recommendations.

The use of different expertise for different accidents might explain the negative correlations; If a team is put together with increased technical and operational knowledge, it could reduce positions for investigators with other types of expertise, thus reducing recommendations at higher socio-technical levels. Because this would depend on each individual investigation, also within each board, the difference in expertise between the two boards would not reveal any possible effects. Ultimately, the relationship between the other socio-technical levels enhances the importance of trying to give an even distribution of attention to each level. Both blunt- and sharp-end factors are important to attain learning for each level.

Despite NTSB's reports, conclusions, and recommendations all being written and decided by others than the Go Team, Cedergren and Petersen (2011) argue that the investigation itself has large implications for these later decisions. Both boards' investigation teams gather the information they deem necessary and relevant, which naturally sets limits for the scope of the report, conclusions, and recommendations. If the influence of the investigation itself is as large as Cedergren and Petersen's study suggest, then there should be indications in the reports regardless of who wrote it.

The results regarding probable cause and use of accident models showed no significant differences between the two investigation boards. This could indicate that the technical expertise of the investigators sets stronger frames for the accident models and the causal factors than it does for safety recommendations. The accident models are closely connected to the scope of the investigation, and since the investigation teams themselves have such similar expertise the results are consistent with Cedergren and Petersen's (2011) theory. The scope of the investigation is further expected to influence the conclusion regarding probable cause –

especially considering the WYLFIWYF effect hypothesised by Lundberg et al. (2009). The safety recommendations are naturally connected to these, but the results indicate that different perspectives might lead to different solutions for the same problem.

5.3 How do different accident models influence conclusions and safety recommendations?

Past literature suggests that accident causes are something we construct rather than find (Lundberg et al., 2009). This would be represented through the reports' conclusions regarding probable cause, as well as being reflected in the safety recommendations. The WYLFIWYF phenomenon entails how perspectives on accidents and their complimenting accident models can be a strong indicator of "what we look for", and as a result also an indicator of what we find.

The results in this study support the WYLFIWYF phenomenon, as there was a significant difference in the probable cause when using different accident models. As much as 46% of the variance in probable cause could be explained by the use of different accident models, which is a strong effect. Because the reports only used systemic and epidemiological accident models, the effect of sequential accident models could not be determined, and the results only reflect a difference between epidemiological and systemic accident models.

As hypothesized, organizational factors were determined to be the probable cause as much as 64% of the time when using systemic accident models, and only 7% when using epidemiological accident models. This supports the possibility that using systemic models might lead to the construction of systemic or organizational causal factors. Not even a single report using systemic accident models concluded that human factors were the probable cause.

Since epidemiological accident models focus on the interaction between active failures and latent conditions, with a strong barrier focus, it might be expected that technological factors would be blamed the most, or that the distribution between them would be rather even. That is not the case, and human factors were determined to be the cause as much as 66% of the time, in contrast with 24% being technological factors (see figure 13). This could indicate that epidemiological models in practice put more focus on human error than expected. Alternatively, it shows a trend of sharp-end factors being the determined probable cause, while blunt-end factors are seen as contributing factors. Neither board (despite their high use of epidemiological models and human-related probable causes) aim many safety recommendations at the actor level.

The fact that none of the boards utilize sequential accident models shows that they consider contributing factors rather than simple cause-effect links. Arguably, this shows signs of higher learning potential in terms of more double-loop learning than single-loop learning (Argyris & Schön, 1978), as solutions to latent conditions, will have to be solutions at deeper organizational levels. However, not everyone agrees that the focus on latent conditions is positive. Le Coze (2013) is critical to the use of accident models that focus on latent conditions and he specifically criticizes the Swiss cheese model. He calls it a 'Swiss Cheese philosophy' that causes investigators to focus all their attention on blunt-end factors while leaving sharp-end factors insufficiently explained and investigated. With this in mind, it is important to find a balance and give attention to both latent conditions and active failures to get the most out of epidemiological models. Thus, results in this study showing a focus on sharp-end causal factors, as well as blunt-end safety recommendations might reflect a healthy balance of focus.

Even though the results in this study are well in line with theories and previous research on the effect of accident models (cf. Benner, 1985; Le Coze, 2013; Leveson, 2004; Lundberg et al., 2009), it is important to consider that the accident models used were determined based on the already finished and published investigation reports. Even though the probable cause was not used to determine the accident model, there is still a chance that the drawn conclusions influenced the focus of the finished reports. Adjustments of focus in the writing could have been made to enhance and back up their conclusions, which in turn would have influenced the data collection in this thesis. Thus, the results could show an effect in the other direction, where the conclusions and outcomes of their investigation influenced the report, and thus the determined accident model. However, especially since the NTSB writes their draft before including any indication of conclusions and recommendations, it is more likely that the effect is the other way around.

If the WYLFIWYF phenomenon is not real, the results might show that they are finding the 'true' cause and that it is most often human factors. Human error is in fact recognized as one of the leading causes of high-speed railway accidents (Wang et al. 2018). On the other hand, that does not explain why not a single report using systemic accident models had human factors as the probable cause. The fact that human error is largely recognized as a lead cause might instead be explained by the large use of epidemiological accident models. Alternatively, there is a combination of human errors playing a large role in many accidents, and epidemiological accident models also putting more focus on them.

Accident models focus on how an accident happens and causal factors, and they are strongly linked with perspectives on accidents. Perspectives on accidents also focus on why accidents happen, and how to prevent them. Because of accident models' link with these broader perspectives, the accident models were expected to also influence the safety recommendations, as they reflect the boards' understanding of how to prevent future accidents. Despite this, the use of different accident models had no significant effect on neither the number of safety recommendations nor the recommendations' socio-technical levels.

Perhaps accident models simply do not give a good indication of perspectives on how to prevent accidents. As argued earlier, the same problem (cause) could be solved with different approaches. Simply identifying how someone perceives a problem would not automatically identify their solutions. Thus, it might seem that the accident models and perspectives on major accidents are useful in various combinations, but not necessarily a good indicator of each other.

5.4 How do different aspects of accident investigations by the NSIA and the NTSB influence learning potential?

The research questions discussed above also help answer the main research problem regarding how different aspects of accident investigations by the NSIA and the NTSB influence learning potential. This requires a summary of the implications for each of the different aspects.

5.4.4 Standard procedures

NTSB does not include recommendations and conclusions in the draft before sending it to their other offices, and involved parties and organizations, while NSIA does. It appears that this has influenced a significant difference in the socio-technical levels of their safety recommendations. NTSB has a more even distribution between the socio-technical levels, and significantly more recommendations at the government and industry level. In contrast, the NSIA showed a significantly higher number at the surroundings level. This might be a result of NTSB's approach allowing a higher number of different perspectives, without priming their opinions with already-made conclusions and recommendations. Based on Cedergren and Petersen (2011), the levels that get the most focus also achieve most learning. Thus, because NTSB has a more even distribution of focus on all levels, it can be argued that they also have achieved a higher learning potential. Additionally, they might achieve increase learning potential by higher levels contributing to double-loop learning.

The NTSB also seems to be expected to aim their recommendations at higher levels than the NSIA is. This difference in expectations and procedures thus could contribute to the same effect on learning potential, and NSIA might want to consider changing their procedures to involve recommendations aimed at the government level. As of right now, they had 0% recommendations aimed at this level, in contrast with NTSB's 8%.

The NSIA is not allowed to provide concrete solutions, while the NTSB's recommendations are highly specific. They also had a significantly higher number of recommendations, further illustrating their increased detail compared to NSIA's recommendations. NSIA's approach gives more responsibility to the involved actors and organizations, rather than simply giving them a list that they can follow. It also allows them to find solutions that are tailored to their organization. This increased responsibility could also increase their learning and understanding, and NTSB might benefit from adopting a similar approach.

The two boards have different approaches regarding what to investigate, where NSIA does more investigations for less serious incidents. NTSB on the other hand does more in-depth investigations for only the most serious accidents, reflected by their higher number of recommendations. NSIA's approach allows for learning from near-accidents too, which could result in necessary interventions *before* any accidents happen. NTSB's approach allows for increased learning through more detailed investigations. Which approach is most beneficial is unclear, and further studies might be necessary to uncover which approach has the most learning potential.

5.4.5 Investigation methods

The investigation method did not seem to have a significant effect on the outcome of the investigations, and thus not the learning potential. This contrasts with previous studies, which have found a big difference in the use of different investigation methods (Benner, 1985; Sklet, 2004). The lack of any effect could be because the methods were highly similar between the two boards, with overlap between their methods. Alternatively, it could be because the use of different methods simply does not have an effect, or because other aspects have a larger influence that outweighs any effect that investigation methods might have.

5.4.6 Expertise

The two boards investigation teams had highly similar expertise, with a focus on technical and operational backgrounds. This is theorized to increase focus on technical aspects of the accident, as investigators will focus on aspects from their own fields. However, NTSB has a larger cross-sectoral investigation board, which is hypothesized to minimize the effect of the investigator's expertise (Cedergren & Petersen, 2011; Stemn et al., 2020; Svenson et al., 1999). Additionally, different offices are responsible for writing the reports, conclusions, and recommendations. The results in this study support these theories, as NSIA has a higher portion of recommendations aimed at the surroundings level, and NTSB has more recommendations at the government and industry level. The NTSB's learning potential might be higher in this aspect as a result of their larger cross-sectoral investigation board, as well as their investigators having less responsibility for the outcome of the investigations.

5.4.7 Accident models

Accident models had a strong significant effect on probable cause, supporting the WYLFIWYF phenomenon. Systemic accident models more often find organizational causal factors, while epidemiological accident models blame human factors the most, as well as technical factors. There were no significant differences between the two boards' use of accident models, so both boards might benefit from considering their use of accident models and how this influences their conclusions. Focusing strongly on one accident model might limit their scope, and thus limit their learning potential. As shown by the relationship between the safety recommendations socio-technical levels, increased focus on certain levels decreases the focus on others. Giving a more even focus between the different causal factors is thus increasingly important and might increase learning potential by considering all alternatives.

5.5 Future research

There are many possibilities for further research. This study has looked into accident models in a very simplistic way, simply giving each report one accident model based on the most prevalent perspective. There is however a possibility to look more nuanced into major accident perspectives. There are many aspects to these perspectives, and each aspect could be analysed before going through a factor analysis. This would give a more nuanced look into which accident models and perspectives are used in practice.

Further research is also needed on accident investigations in different countries, especially outside of western cultures. This study only investigated reports within one specific transportation sector, and there is a lack of research into how investigations are done in different sectors and fields, meaning there might be differences based on fields yet to be explored.

Because the findings indicated that the investigation board's expertise as a whole might have a significant effect, rather than just the investigation team, it would be relevant for further studies to include data on the whole board's expertise, and its possible effect. This is especially relevant for larger and cross-sectional investigation boards like the NTBS.

Finally, research into whether more or less concrete safety recommendations have the most learning potential could provide valuable insight into the best approach when it comes to safety recommendations.

6. Conclusion

This thesis has investigated several different aspects of accident investigations, to find out *how different aspects of accident investigations by the NSIA and the NTSB influence learning potential*. The aspects in focus were related to standard procedures, investigation methods, accident models and expertise. The results have provided support for earlier studies and theories, as well as some conflicting results. Results and conclusions may not be a final answer for these questions but are logical conclusions considering the theoretical context.

The results from this study imply that different standard procedures, expertise, and accident models all influence the outcome of the investigations, and thus have implications for learning potential. In contrast, investigation methods did not have any apparent effect, but this might be explained by the investigation boards' methods simply being too alike to uncover any effect.

Having a larger cross-sectoral investigation board seems to increase learning potential by giving a more even focus between socio-technical levels. No effect was found on probable cause, which could indicate that the investigation team's expertise has a larger influence on the probable cause, while the wider investigation board has a larger influence on the safety recommendations. Further standard procedures at the NTSB also seem to give more focus to higher socio-technical levels, such as sending out the report draft before including any recommendations, only investigating more serious incidents, and expecting recommendations at the government level.

The NSIA is not allowed to underline concrete solutions in their safety recommendations, which promotes personal growth and learning for the involved actors by giving them more responsibility. Furthermore, they investigate both less serious accidents and near-accidents which can promote learning before a serious accident has to occur. The NTSB on the other hand, does fewer investigations and thus has the capacity to be more thorough and give out a larger number of safety recommendations. The balance between quality and quantity has to be considered, as both approaches can promote learning.

Accident models had a strong significant effect on probable cause, supporting the WYLFIWYF phenomenon. Systemic accident models more often find organizational causal factors, while epidemiological accident models blame human factors the most, as well as technical factors. Both boards use epidemiological accident models most frequently, and they could increase their learning potential by considering other models and widening their scope.

Both the NSIA and the NTSB have aspects that could increase and decrease learning potential, as well as some different aspects where it is unclear which approach is best. Both boards can benefit from adapting aspects from the other, as well as being conscious of how the different aspects may influence their results.

References

Accident Investigation Board Norway (2018). *The AIBN Method: Framework and analysis process for systematic safety investigations.* (2nd ed). AIBN. https://havarikommisjonen.no/?nid=13467&lcid=1033&iid=29135&pid=SH-Modules-FilesModule-Files.SH-FileLine-File

Argyris, C. & Schön, D. A. (1978). *Organizational learning: A theory of action perspective.* Addison-Wesley.

Benner, L. J. (1985). Rating accident models and investigation methodologies. *Journal of safety research, 16*(3). 105-126. https://www.sciencedirect.com/science/article/abs/pii/0022437585900386

Blaikie, N. (2000). *Designing social research: The logic of anticipation.* Polity Press

Brath, O. B. (2020). *Learning from incidents. Evaluating safety interventions after incident investigations* [Master's thesis, University of Stavanger]. UiS Brage. https://uis.brage.unit.no/uis-xmlui/handle/11250/2681460

Cedergren, A. (2013). Implementing recommendations from accident investigations: A case study of inter-organisational challenges. *Accident analysis and prevention, 53*(2013). 133-141. https://doi.org/10.1016/j.aap.2013.01.010

Cedergren A. & Petersen, K. (2011). Prerequisites for learning from accident investigations: A cross-country comparison of national accident investigation boards. *Safety science, 49*(8-9). 1238-1245. https://doi.org/10.1016/j.ssci.2011.04.005

Danermark, B., Ekström, M., Jakobsen L., & Karlsson, J. C. (2002). *Explaining Society: Critical realism in the social sciences.* Routledge

Dechy, N., Dien, Y., Funnemark, E. Roed-Larsen, S., Stoop, J. Valvisto, T. & Arellano, A. L. V. (2012). *Safety science, 50*(2012). 1380-1391. http://dx.doi.org/10.1016/j.ssci.2009.10.004

Dekker, S. (2014). *The field guide to understanding 'human error'* (3rd. ed). Ashgate Publishing.

Drupsteen, L., Groeneweg, J. & Zwetsloot, G. I. J. M. (2013). Critical steps in learning from incidents: Using learning potential in the process from reporting an incident to accident prevention. *International journal of occupational safety and ergonomics: JOSE, 19*(1). 63-77. https://doi.org/10.1080/10803548.2013.11076966

Federal Railroad Administration. (2021). *Railroad safety specialist: Department of transportation.* [Job advertisement]. Retrieved 21.04.21 from https://www.usajobs.gov/GetJob/ViewDetails/598632600

Filho, A. P. G., Jun, G. T. & Waterson, P. (2019). Four studies, two methods, one accident: An examination of the reliability and validity of Accimap and STAMP for accident analysis. *Safety science, 113*. 310-317. https://doi.org/10.1016/j.ssci.2018.12.002

Fiol, C. M. & Lyles, M. A. (1985). Organizational learning. *The academy of management review, 10*(4). 803-813. https://doi.org/10.5465/amr.1985.4279103

Fu, G., Xie, X., Jia, Q., Li, Z., Chen, P. & Ge, Y. (2020). The development history of accident causation models in the past 100 years: 24Model, a more modern accident causation model. *Process safety and environmental protection, 134* (2020). 47-82. https://doi.org/10.1016/j.psep.2019.11.027

Fukuoka, K. & Furusho, M. (2017). Convergence of the epidemiological and systemic accident models. *Journal of maritime research, 7*. 1-11. https://core.ac.uk/download/pdf/132511189.pdf

Gibb, A., Lingard, H., Behm, M. & Cooke, T. (2014). Construction accident causality: Learning from different countries and differing consequences. *Construction management and economics, 32*(5). 446-459. https://doi.org/10.1080/01446193.2014.907498

Golafshani, N. (2003), Understanding reliability and validity in qualitative research. *The qualitative report, 8*(4). 597-606. https://doi.org/10.46743/2160-3715/2003.1870

Goode, N., Shaw, L., Finch, C. F. & Salmon, P. M. (2019). Challenges of translating Rasmussen's Accimap into a usable, sustainable, and useful incident reporting system: End-user attitudes following 12-month implementation. *Cognition, technology & work* (2019). https://doi.org/10.1007/s10111-019-00596-x

Hale, A. R., & Glendon, A. I. (1987). *Individual behaviour in the control of danger*. Industrial safety series (vol. 2). Elsevier Science.

Harms-Ringdahl, L. 2004. Relationships between accident investigations, risk analysis and safety management. *Journal of hazardous materials, 111*(2004). 13-19. https://doi.org/10.1016/j.jhazmat.2004.02.003

Hopkins, A. (2001). Was Three Mile Island a 'normal accident'? *Journal of contingencies and crisis management*, *9*(2). 65-72. https://doi.org/10.1111/1468-5973.00155

Hovden, J., Albrechtsen, E. & Herrera, I. A. (2010). Is there a need for new theories, models and approaches to occupational accident prevention? *Safety science*, *48*(8). 950-956. https://doi.org/10.1016/j.ssci.2009.06.002

Hovden, J., Sklet, S. & Tinnmannsvik, R. K. (2004). I etterpåklokskapens klarsyn: Gransking og læring av ulykker. I Lydersen, S. (Red.), *Fra flis i fingeren til ragnarok: Tjue historier om sikkerhet.* 163-182. Trondheim: Tapir akademisk forlag.

Hyungju, K., Haugen, S. & Bouwer, I. U. (2016). Assessment of accident theories for major accidents focusing on the MV SEWOL disaster: Similarities, differences and discussion for a combined approach. *Safety science, 82*(2016). 410-420. https://doi.org/10.1016/j.ssci.2015.10.009

Haddon, W. (1970). On the escape of tigers: An ecological note. *American journal of public health*, *60*(12). 2229-2234. https://doi.org/10.2105/AJPH.60.12.2229-b

Haddon, W. (1980). Advances in the epidemiology of injuries as a basis for public policy. *Public health reports*, *95*(5), 411-421.

IBM Corp. (Released 2019). IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp

Jernbaneundersøkelsesloven. (2005). *Lov om varsling, rapportering og undersøkelse av jernbaneulykker og jernbanehendelser m.m.* (LOV-2005-06-03-34). Lovdata. https://lovdata.no/dokument/NL/lov/2005-06-03-34

Karanikas, N. Soltani, P., De Boer, R. J. & Roelen A. (2015). Evaluating advancements in accident investigations using a novel framework. *Proceedings of the 5th air transport and operations symposium*. 1-10. Delft University of Technology.

Katsakiori, P., Skallaropoulos, G. & Manatakis, E. (2009). Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models. *Safety science, 47*(7). 1007-1015. https://doi.org/10.1016/j.ssci.2008.11.002

Kim, H. & Haugen, S. (2015). Titanic viewed from different perspectives on major accidents. In Nowakowski, T., Mynczak, M., Jodejko-Pietruczuk, A. & Werbinska-Wojciechowska, S. (ed.). *Safety and reliability: Methodology and applications* (p. 65-73). CRC Press.

Kjellén, U. & Albrechtsen, E. (2017). *Prevention of accidents and unwanted occurrences: Theory, methods and tools in safety management* (2nd ed). CRC Press.

Klaveness, I. (2012). *Granskingsmetodikk i petroleumsindustrien* [Master's thesis, University of Stavanger]. UiS Brage. https://uis.brage.unit.no/uis-xmlui/handle/11250/184742

Kongsvik, T. Ø., Schiefloe, P. M., Hovden, J., Herrera, I., Antonsen, S. & Albrechtsen, E. (2018). *Sikkerhet i arbeidslivet*. Fagbokforlaget.

Krogsæter, S. (2016). *Granskingsdilemmaer i tilsyn og kommisjon: En komparativ dokumentanalyse av Statens havarikommisjon for transport og Petroleumstilsynet* [Master's thesis, University of Tromsø]. UiT Munin. https://munin.uit.no/handle/10037/10034

La Porte, T. R. & Consolini, P. M. (1991). Working in practice but not in theory: Theoretical challenges of "High-Reliability Organizations". *Journal of public administration research and theory: J-Part, 1*(1). 19-48. https://www.jstor.org/stable/1181764

Le Coze, J. C. (2013). What have we learned about learning from accidents? Post-disasters reflections. *Safety Science, 51*(1). 441-453. https://doi.org/10.1016/j.ssci.2012.07.007

Lehto, M & Salvendy, G. (1991). Models of accident causation and their application: Review and reappraisal. *Journal of engineering and technology management, 8*(2). 173-205. https://doi.org/10.1016/0923-4748(91)90028-P

Leveson, N. (2004). A new accident model for engineering safe systems. *Safety science, 42*(4). 237-270. https://doi.org/10.1016/S0925-7535(03)00047-X

Lundberg, J., Rollenhagen, C. & Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find: The consequences of underlying accident models in eight accident investigation manuals. *Safety science*. *47*(10). 1297-1311. https://doi.org/10.1016/j.ssci.2009.01.004

Lundberg, J., Rollenhagen, C., Hollnagel, E. & Rankin, A. (2012). Strategies for dealing with resistance to recommendations from accident investigations. *Accident analysis & prevention*, *45*(2012). 455-467. https://doi.org/10.1016/j.aap.2011.08.014

Malterud, K. (2001). Qualitative research: standards, challenges, and guidelines. *The Lancet*, *358*(9280). 483-488. https://doi.org/10.1016/S0140-6736(01)05627-6

National Transportation Safety Board. (2002). *Aviation investigation manual: Major team investigations*. https://www.ntsb.gov/investigations/process/Documents/MajorInvestigationsManual.pdf

National Transportation Safety Board. (2017). *Fiscal years 2018-2022 strategic plan*. OMB final submission. https://www.ntsb.gov/about/reports/Documents/FY2018-2022strategicPlan.pdf

National Transportation Safety Board. (2020). *About The National Transportation Safety Board Organization*. Retrieved 30.04.21 from https://www.ntsb.gov/about/organization/Pages/default.aspx

National Transportation Safety Board. (n.d.). *The investigative process*. Retrieved 21.04.21 from https://www.ntsb.gov/investigations/process/Pages/default.aspx

Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Person Education Limited

Norwegian Safety Investigation Authority. (n.d.). *About us*. Retrieved 15.03.21 from https://havarikommisjonen.no/About-us

Perrow, C. (1984). *Normal accidents: Living with high risk technologies*. Basic Books

Pidgeon, N. & O'Leary, M. (2000). Man-made disasters: Why technology and organizations (sometimes) fail. *Safety science, 34*(2000). 15-30. https://doi.org/10.1016/S0925-7535(00)00004-7

Polit, D. F. & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. https://doi.org/10.1016/j.ijnurstu.2010.06.004

Qureshi, Z. (2007). *A review of accident modelling approaches for complex socio-technical systems* [Paper presentation]. 12th Australian workshop on safety related programme systems (SCS'07), Adelaide. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.294.4125&rep=rep1&type=pdf

Rasmussen, J. (1997) Risk management in a dynamic society. *Safety science, 27*(2/3). 183-213. https://doi.org/10.1016/S0925-7535(97)00052-0

Reason, J. (1997). *Managing the risks of organizational accidents*. Ashgate Publishing

Rochlin, G. I., La Porte, T. R. & Roberts, K. H. (1987). The self-designing high-reliability organization: Aircraft carrier flight operations at sea. *New war college review, 40*(4). 76-92. https://www.jstor.org/stable/44637690

Roed-Larsen, S. & Stoop, J. (2012). Modern accident investigation: Four major challenges. *Safety science, 50*(2012). 1392-1397. https://doi.org/10.1016/j.ssci.2011.03.005

Rosness, R., Guttomersen, G., Steiro, T., Tinmannsvik, R. & Herrera, I. A. (2004). *Organisational accidents and resilient organisations: Five perspectives*. (SINTEF Report STF38 A 04403). https://www.sintef.no/globalassets/upload/teknologi_og_samfunn/sikkerhet-og-palitelighet/rapporter/stf38-a04403.pdf

Samferdselsdepartementet. (2020). *Instruks for Statens havarikommisjon*. https://havarikommisjonen.no/filmodul?iid=31806&pid=SH-Modules-FilesModule-Files.SH-FileLine-File

Statens Havarikommisjon. (2021). *Havarikommisjonen: Årsrapport 2020.* Retrieved from: https://havarikommisjonen.no/?nid=13473&lcid=1044&iid=33440&pid=SH-Modules-FilesModule-Files.SH-FileLine-File

Statens Havarikommisjon. (2020). *Statens havarikommisjon for transport søker havariinspektør* [Job Advertisement]. Retrieved 23.04.21 from https://e24.no/ledige-stillinger/s/3JbrlA/statens-havarikommisjon-for-transport-sker-havariinspektr

Statens Havarikommisjon. (n.d.) *Organisering*. Retrieved 20.04.21 from https://havarikommisjonen.no/Om-oss/Organisering

Statens Havarikommisjon for Transport. (2019). *Årsrapport 2018*. Retrieved from: https://www.regjeringen.no/contentassets/2ff4d41f5ff648afb9ae66994863d1f7/statens-havarikommisjon-arsrapport-for-2019.pdf

Stemn, E. Hassall, M. E. & Bofinger, C. (2020). Systemic constraints to effective learning from accidents in the Ghanaian mining industry: A correspondence analysis and AcciMap approach. *Safety Science, 123*(2020).

Stormo, S. (2011). *Gransking: Læring i organisasjoner eller pulverisering av skyld?* [Master's thesis, University of Stavanger]. UiS Brage. https://uis.brage.unit.no/uis-xmlui/handle/11250/184576

Sothivanan, S. & Siddiqui, N. A. (2015). Laconic study on incident/accident investigation technique: Sequentially timed event plotting (STEP). *International journal for scientific research & development, 3*(8). 680-683

Svenson, O., Lekberg, A. & Johansson, A. E. L. (1999). One perspective, expertise and differences in accident analyses: Arguments for a multidisciplinary integrated approach. *Ergonomics, 42*(11). 1561-1571. https://doi.org/10.1080/001401399184893

Tinmannsvik, R.K. & Kjellen, U. (2018) Gransking etter hendelser: SIBA- sikkerhetsstyring i BA- bransjen

Toft, Y., Dell, G., Klockner, K. K. & Hutton, A. (2012). Models of causation: Safety. In Health and Safety Professionals Alliance (HaSPA). *The core body of knowledge of generalist OHS professionals.* Safety Institute of Australia.

Turner, B. A. & Pidgeon, N. F. (1997). *Man-made disasters* (2nd ed). Butterworth-Heinemann.

49 CFR. (2021). *Code of Federal Regulations Title 49.* Subpart D: Authority of NTSB in railroad, pipeline and hazardous materials investigations. § 831.40 https://www.ecfr.gov/cgi-bin/text-idx?SID=f1e35b933faecab96a2d3f2974ee50a3&mc=true&node=se49.7.831_140&rgn=div8

49 U.S.C. (2012) *United States Code Title 49: Transportation*. Subtitle II: Other Government Agencies, Chapter 11: National Transportation Safety Board. §§ 1101-1155. https://www.law.cornell.edu/uscode/text/49/subtitle-II/chapter-11

U.S. Department of Energy. (2000). *DOE workbook: Conducting accident investigations.* U.S. Department of Energy, Washington, DC.

Wang, W., Liu, X. & Qin, Y. (2018). A modified HEART method with FANP for human error assessment in high-speed railway dispatching tasks. *International journal of industrial ergonomics*, *67*(2018). 242-258. https://doi.org/10.1016/j.ergon.2018.06.002

Woods D. D., Dekker, S., Cook R. I., Johannesen L. & Sarter, N. (2010). *Behind Human Error* (2nd ed.). CRC Press.

Appendix A

Table 11

Included NTSB accident reports

Report nr.	Title	Accident date	Report date
RAR-20-05	Collision of Union Pacific Railroad Train MGRCY04 with a Stationary Train	04/10/2018	29/12/2020
RAR-20-04	CSX Train Derailment with Hazardous Materials Release	02/08/2017	23/11/2020
RAR-20-03	Collision of Two CSX Transportation Freight Trains	12/08/2019	15/09/2020
RAR-20-02	CSX Transportation Railway Maintenance Machine Operator Fatality	12/03/2018	24/07/2020
RAR-20-01	Long Island Rail Road Roadway Worker Fatality	10/06/2017	29/04/2020
RAR-19-02	Amtrak Passenger Train Head-on Collision With Stationary CSX Freight Train	04/02/2018	23/07/2019
RAR-19-01	Amtrak Passenger Train 501 Derailment	18/12/2017	21/05/2019
RAR-18-02	Derailment and Hazardous Materials Release of Union Pacific Railroad Unit Ethanol Train, Graettinger, Iowa, March 10, 2017	10/03/2017	30/10/2018
RAR1801	BNSF Railway Roadway Worker Fatalities, Edgemont, South Dakota, January 17, 2017	17/01/2017	27/08/2018
RAR1702	Amtrak Train Collision with Maintenance-of-Way Equipment, Chester, Pennsylvania	03/04/2016	14/11/2017
RAR1701	Highway-Railroad Grade Crossing Collision, Commerce Street, Valhalla, New York, February 3, 2015	03/02/2015	25/07/2017
RAR1603	Railroad Accident Report: Collision of Two Union Pacific Railroad Freight Trains	17/08/2014	19/12/2016
RAR1602	Derailment of Amtrak Passenger Train 188	12/05/2015	17/05/2016
RAR1601	Washington Metropolitan Area Transit Authority L'Enfant Plaza Station Electrical Arcing and Smoke Accident	12/01/2015	03/05/2016
RAR1502	Collision Involving Three BNSF Railway Freight Trains near Amarillo, Texas, September 25, 2013	25/09/2013	25/06/2015
RAR1501	Chicago Transit Authority Train Collides with Bumping Post and Escalator at O'Hare Station	24/03/2014	28/04/2015
RAR-14-02	Collision of Union Pacific Railroad Freight Train with BNSF Railway Freight Train	25/05/2013	17/11/2014
RAR1401	Conrail Freight Train Derailment with Vinyl Chloride Release	30/11/2012	29/07/2014
RAR-13-02	Head-On Collision of Two Union Pacific Railroad Freight Trains	24/06/2012	18/06/2013
RAR-13-01	Collision of Two Canadian National Railway Freight Trains	30/09/2010	12/02/2013

RAR-12-05	Collision of Port Authority Trans-Hudson Train with Bumping Post at Hoboken Station	08/05/2011	05/11/2012
RAR-12-04	Washington Metropolitan Area Transit Authority Hi-Rail Maintenance Vehicle Strikes Two Wayside Workers Near the Rockville Station	26/01/2010	17/05/2012
RAR-12-03	Collision of Dakota, Minnesota & Eastern Railroad Freight Train and 19 Stationary Railcars	14/07/2009	30/04/2012
RAR-12-02	Collision of BNSF Coal Train With the Rear End of Standing BNSF Maintenance-of-Way Equipment Train	17/04/2011	24/04/2012
RAR-12-01	Derailment of CN Freight Train U70691-18 With Subsequent Hazardous Materials Release and Fire	19/06/2009	14/02/2012
RAR-11-01	Miami International Airport, Automated People Mover Train Collision with Passenger Terminal Wall	28/11/2008	08/11/2011
RAR-10-02	Collision of Two Washington Metropolitan Area Transit Authority Metrorail Trains Near Fort Totten Station	22/06/2009	27/07/2010
RAR-10-01	Collision of Metrolink Train 111 with Union Pacific Train LOF65-12	12/09/2008	21/01/2010

Table 12

Included NSIA accident reports

Report nr.	Title	Accident date	Report date
2021/01	Rapport om sammenstøt mellom tog 64 og personbil på planovergang på Vikersund stasjon 28. januar 2020	28/01/2020	26/01/2021
2020/09	Rapport om dødsulykke ved Storforshei, Nordlandsbanen 7. desember 2019	07/12/2019	01/12/2020
2020/08	Rapport om avsporing på Bryn stasjon, Hovedbanen 25. september 2019	25/09/2019	28/08/2020
2020/06	Rapport om sammenstøt mellom tog 1859 og 1860 på Berekvam stasjon, Flåmsbana 31. juli 2019	31/07/2019	10/06/2020
2020/05	Rapport om avsporing på Oslo S 8. mai 2019	08/05/2019	05/05/2020
2020/04	Rapport om avsporing ved Bjørnstad, km 303,3 mellom Majavatn og Namskogan, Nordlandsbanen, tog 5790 den 25. juli 2019	25/06/2019	30/04/2020
2020/02	Rapport om personulykke på Filipstad driftsbanegård søndag 24. februar 2019	24/02/2019	21/01/2020
2019/10	Rapport om brann og eksplosjon i T-banetunnelen mellom Ensjø og Helsfyr 17. desember 2018	17/12/2018	11/12/2019
2019/09	Rapport om planovergangsulykke på Bjøråneset planovergang, Rørosbanen 29. november 2018	29/11/2018	27/11/2019

2019/03	Rapport om brann i lokomotiv El 14.2188 ved Oslo S 29. mai 2018, tog 5301	29/05/2018	04/04/2019
2019/02	Rapport om avsporing på Blaker stasjon, Kongsvingerbanen 12. mars 2018	12/03/2018	11/03/2019
2019/01	Rapport om personulykke ved Alna holdeplass på Hovedbanen 2. mars 2018	02/03/2018	19/02/2019
2018/11	Rapport om sammenstøt ved Grorud stasjon 16. desember 2017	16/12/2017	05/12/2018
2018/10	Rapport om sammenstøt mellom tog 135 og veiskrape på Høium planovergang, Østfoldbanen 23. januar 2018	23/01/2018	05/11/2018
2018/08	Rapport om avsporing på Trondheim sentralstasjon 26. oktober 2017	26/10/2017	15/08/2018
2018/07	Rapport om avsporing ved Svorkmo stasjon på Thamshavnbanen 3. oktober 2017	03/10/2017	01/08/2018
2018/06	Rapport om personulykke på Fetsund stasjon, Kongsvingerbanen 22. februar 2018, tog 1015	22/02/2018	03/07/2018
2018/03	Rapport om avsporing på Loenga stasjon 14. april 2017	14/04/2017	11/04/2018
2018/01	Rapport om personulykke på Holstein T-banestasjon 2. februar 2017	02/02/2017	24/01/2018
2017/08	Rapport om sammenstøt mellom tog og skinnegående gravemaskin ved Dallerud på Dovrebanen 11. mars 2017	11/03/2017	18/12/2018
2017/03	Rapport om avsporing nord for Bøn stasjon på hovedbanen 31. mai 2016	31/05/2016	22/05/2017
2017/01	Rapport om alvorlig jernbanehendelse, Hovedbanen, Alnabru skiftestasjon 29. april 2016	29/04/2016	23/03/2017
2016/06	Rapport om jernbaneulykke på Fauske stasjon 30. mars 2016 med tog 471	23/03/2016	08/12/2016
2016/04	Rapport om jernbaneulykke ved Oppegård stasjon 20. mai 2015 med tog 45958	20/05/2015	19/05/2016
2016/02	Rapport om togavsporing ved Grytå, km 285,28 på Bergensbanen 9. februar 2015, tog 5502	09/02/2014	08/02/2016
2015/08	Rapport om jernbaneulykke med trikk og buss i krysset Cort Adelers gate / Munkedamsveien 11. november 2014	11/11/2014	XX/11/2015
2015/07	Rapport om avsporing med tog 5932 Eidsvoll stasjon, Dovrebanen 3. november 2014	03/11/2014	10/09/2015
2015/05	Rapport om alvorlig jernbanehendelse Bybanen, Kaigaten 26. mai 2014	26/05/2014	21/05/2015
2015/04	Rapport om avsporing med tog 5790 ved Trofors stasjon, Nordlandsbanen 30. mai 2014	30/05/2014	14/04/2015
2015/03	Rapport om avsporing ved Svene på Numedalsbanen 15. april 2014	15/04/2014	27/03/2015
2015/02	Rapport om jernbaneulykke på Steinerud planovergang på Holmenkollbanen 9. februar 2014	09/02/2014	19/01/2015

2015/01	Rapport om alvorlig jernbanehendelse ved Jar stasjon 23. januar 2014 tog 1308	23/01/2014	08/01/2015
2014/07	Rapport om avsporing mellom Dovre og Dombås stasjoner på Dovrebanen 4. november 2013 tog 5910	04/11/2013	03/11/2014
2014/06	Rapport om påkjørsel ved Mårdalen holdeplass, Bybanen Bergen 17. oktober 2013	17/10/2013	09/10/2014
2014/05	Rapport om sammenstøt mellom tog og bil på Enebekk planovergang på Østfoldbanen vestre linje 1. august 2013	01/08/2013	01/07/2014
2014/04	Rapport om jernbaneulykke på Lambertseterbanen, Høyenhall stasjon 5. juli 2013 tog 409	05/07/2013	18/06/2014
2014/03	Rapport om avsporing med godstog 41631 ved km 281,5 mellom Kvam og Sjoa på Dovrebanen 22. juli 2013	22/07/2013	17/06/2014
2014/01	Rapport om sammenstøt mellom et skift og materiellet til tog 5509 på Alnabru stasjon 9. januar 2013	09/01/2013	08/01/2014
2013/02	Rapport om togavsporing ved Nykirke stasjon, Vestfoldbanen, 15. februar 2012 tog 12926	15/02/2012	12/02/2013
2012/06	Rapport om togavsporing tog 2378 ved Krokegga, Opphus, Rørosbanen 5. september 2011	05/09/2011	16/07/2012
2012/05	Rapport om jernbaneulykke Bergensbanen, Hallingskeid stasjon 16. juni 2011, tog 62	16/06/2011	15/06/2012
2011/08	Rapport om avsporing med tog 55 på Skotterud stasjon Kongsvingerbanen 1. oktober 2010	01/10/2010	22/09/2011
2011/01	Rapport om sammenstøt mellom trikk og buss ved Ilevollen i Trondheim 5. november 2009	05/11/2009	02/02/2011
2010/08	Rapport om sammenstøt mellom persontog 2387 og godstog 5741 på Koppang stasjon Rørosbanen 17. desember 2009	17/12/2009	22/11/2010
2010/02	Rapport om jernbaneulykke ved Nationaltheatret stasjon den 7. juni 2009, tog 3744	07/06/2009	11/03/2010