Exploiting control attempt effectiveness in the development of safety events

Exploiting control attempt effectiveness in the development of safety events
The study described in this thesis has been performed at the Aviation Academy of the Amsterdam Universty of Applied Sciences for the bachelor programme Aviation Studies, Amsterdam, The Netherlands.

Cover art: “*Human controlled safety performance indicator airplane*”
by Jeffrey Nederend

Amsterdam University
of Applied Sciences

Exploiting control attempt effectiveness in the development of safety events

A thesis

Bachelor of Science,
Aviation Studies
at the Amsterdam University of Applied Sciences

Jeffrey Nederend
500661672
4 July 2016

Supervisors: dr. N. Karanikas
S. Piric

“Never regard study as a duty, but as the enviable opportunity to learn to know the liberating influence of beauty in the realm of the spirit for your own personal joy and to the profit of the community to which your later work belongs.”

- Albert Einstein.

Table of contents

	Acknowledgements	viii
	Executive summary	ix
	List of abbreviations	xii
Chapter 1	*General introduction*	1
1.1	Research background	3
1.2	Research aim	6
1.3	Research questions	7
1.4	Research scope	8
1.5	Research approach	9
1.6	Outline of this thesis	9
Chapter 2	*Literature review*	11
2.1	International standard in safety investigations	13
2.2	The safety investigation process	16
2.3	Data considerations	20
2.4	Safety performance measurement in civil aviation	23
2.5	Safety event's controllability as performance indicator	30
2.6	Concluding remarks	36
Chapter 3	*Methodology*	39
3.1	Research sample	41
3.2	Framework development	42
3.3	Quality assurance	51
3.4	Data analysis methods	54
3.5	Summary	56
Chapter 4	*Results*	57
4.1	Frequency analysis	59
4.2	Associated factors to safety event controllability	60
4.3	Severity induced bias of safety event's outcomes	73
4.4	Correlation statistics and exploration of relations	79

Chapter 5 *Discussion* 83
5.1 Associations with controllability and intervention effectiveness in safety events 84
5.2 Controllability and severity classifications 87
5.3 Safety investigation bias in safety events 87
5.4 Selection priority criteria 89
5.5 Limitations of the study 90

Chapter 6 *Conclusions and recommendations* 93
6.1 Conclusions of the study 94
6.2 Future research considerations 97
6.3 Recommendations 98

Addendum References 101
Nederlandse samenvatting 106
Appendices 110

Acknowledgements

After an interesting period of four months and reflecting back at the first moment I read my assignment, I can state that my period at the Aviation Academy has been interesting. One confession should be made however: this assignment was my very last hope. For days on end, I have tried to make contact with companies in the aviation industry, but not even one expressed interest in my acquisition as intern. Rest assured, this particular assignment was placed on the DLWO portal. Every other student would say that this 'opportunity' is really the last option they would even want to consider, me included. As harsh as it may sound, I have learnt my lessons eventually. Of course there are no fancy aircraft, no trips abroad or interesting sights to behold, but what is the one thing that distinguishes students in their graduation? Achieving new knowledge and/or specialties in the field of study. My achievement? A new understanding and motivation in doing research. Don't get me wrong, I am no expert, just competent. However, I can state that I have learnt new ways of thinking and approaching problems, which made me motivated for future endeavours. One (overused, but appropriate) quote encapsulates my situation perfectly: "Don't judge a book by its cover". Be satisfied with the opportunities presented, and make the best of it. This message really made me realise that, although I am still a little disappointed that I did not had a chance to work in the 'real' aviation industry, good things can come from lesser options.

This change in perspective is mostly due to one professor in specific, Dr. Nektarios Karanikas. My gratitude is sincere. I have always admired your determination in your field of work and looked up to your work in general. As it should, you only helped me when it was really needed, demanding me to independently reach my goals. I would like to thank you for your support in my graduation period and look forward to continue this collaboration once this thesis is accepted.

One last acknowledgement is for my supervisor Selma Piric. Although the communication has been minimal, the effort and help I have received was always in the best interest of me. As a supervisor you stood up for me to defend my rights as a graduation student. Therefore, I would like to thank you for your support, it is very appreciated.

Executive summary

Today, the aviation industry is safer than it has ever been. However, in order to state such a significant claim, the outcomes of safety events are mostly used to compute safety performance. Moreover, the number of accidents has declined to a significantly low level in which safety cannot be representatively expressed in terms of adverse events. Safety events are most often contributed by the actions of the human element, however, this critical part in the development of safety events is not found in today's computation of safety performance. It is a rather deterministic approach to express safety performance on adverse outcomes is as it might not represent the real safety level of organisations and the aviation industry on the whole since it does not count for the controllability of the event. Besides, investigative resources of safety investigation authorities are allocated according the severity of these outcomes. Accidents and serious incidents are required by regulation to be supported by safety investigations. Incidents, on the other hand, can provide better information for the prevention of safety events. Nevertheless, incidents are not required for investigations nor are there enough resources available to investigate all incidents.

A new taxonomy was published that considers if an occurrence was (1) controlled by the operator, meaning that there was a chance to intervene the development of the event and alleviate the outcome, (2) neutrally controlled by the operator, indicating a control attempt that was reactionary in nature or standard procedure or (3) uncontrolled, meaning that there was no control attempt by the operator and as a result developed without intervention. Taking this taxonomy into consideration, it was exploited to present an alternative approach, other than adverse events, for the indication of safety performance and to allocate resources of safety investigation authorities more appropriately with the issues raised in controllability rather than the impact by outcome severity. As such, the controllability in safety performance might be more representative of safety performance and priority criteria can be generated to appropriately drive the safety investigation's resources and support the resource restraints in incident investigations.

This study applied the new taxonomy on a sample of 297 safety investigation reports among five distinguished safety investigation authorities. These reports were the main source of information as it is closely related to the expected practical use of the taxonomy: occurrence data. Applying the taxonomy on these reports while focussing

on the taxonomy's associated factors provided certain areas in which the utilisation is beneficial or difficult to exploit. These areas were determined and supported by scientific literature and found in the analysis by associated factors in the controllability of safety events. Analysis methods in this study included frequency, Chi-square and Spearman's rho analyses.

It was found that the factors affecting the controllability included differences in air crew nationalities, aircraft generation, type of operation and the human performance related to experience, rest periods, duty time and fatigue. There were, however, no variations in controllability over time, which are depicted in terms of accident rates. Safety investigation authorities were found to have bias towards the severity of safety events. However, it was found more interesting that investigation focussed on control attempts that did go wrong, rather than focussing on those that did go right. Improvements in safety are therefore expected by these authorities to be achieved by investigating adverse events, while improvements can be achieved through the control attempts that do go right. With the knowledge of an event's controllability, priority criteria can provide safety investigation authorities with information to allocate resources accordingly and more evenly across the accident control classes, without bias towards an event's severity. Accidents and serious incidents are required to be supported with an investigation and the exploitation of controllability information can therefore only be used as a means to manage resources more effectively. In addition, in order to obtain controllability information a small inquiry is mandatory at the initial stage of investigating.

On the other hand, incidents are not required to be investigated and due to resource restraints are not as often or as thorough investigated. Obtaining controllability information for incidents was comprised of two options: (1) initial inquiry and (2) occurrence reporting scheme. Identical to accidents and serious incidents, an initial inquiry requires effort from the safety investigation authority. The second option relied on the live witnesses of incidents. Implementing a control field in occurrence reporting schemes allows safety investigations to select and prioritise resources for incident investigation, without additional effort. In addition, with a common agreed 'controllability' field in occurrence reporting schemes, larger amounts of data can reveal trends in incident safety performance and indicate areas of special attention that, in return, can be utilised to allocate investigative resources accordingly. With respect to resource restrains, the selection of incidents was aimed at the rare nature of one control class in particular, uncontrolled. Investigations for other control classes are important nonetheless, however, these events are also investigated in obligatory events. These rare uncontrolled incidents are hence prioritised for selecting incident investigations. That is, an investigation authority should conduct an investigation in uncontrolled incidents when presented with one.

The taxonomy indicates whether events were dependent on chance, indicated by uncontrolled events, or actually comprised of potential controllability to intervene in the development of such events. It can furthermore be utilised to depict control attempt effectiveness to demonstrate control performance of implicated personnel. The associated factors with both elements can indicate areas that require additional attention to improve safety more focussed. Besides it was found that the safety performance in controllability for aviation did not change for a time period of 25 years, while improvement in safety is claimed for this time period in terms of accident rates. In order to improve safety by means of the controllability taxonomy, target levels must be set at higher ratios of controlled by uncontrolled states and positive by negative control attempts. Taking this all into consideration, safety performance as indicated by the controllability of safety events comprises as a more representative and effective manner to indicate safety performance than the commonly used accident rates.

Finally, it is recommended that the implementation of controllability in occurrence reporting schemes is explored in further research. The taxonomy's inclusion can alter incident selection for investigation in a drastic and more effective manner, as it allows depicting trends and revealing critical areas in the industry. It is furthermore recommended that any industry or institution that seeks to unveil its safety performance in terms of controllability and drive resources according the areas of special attention are encouraged to consider and implement this new taxonomy.

List of abbreviations

AAIB	Air Accidents Investigation Branch
AAIU (Be)	Air Accident Investigation Unit (Belgium)
ADREP	Accident/Incident Data Reporting
AFI	Africa
ALoSP	Acceptable Level of Safety Performance
ANSP	Air Navigation Service Provider
APAC	Asia Pacific
ASR	Air Safety Report
ATSB	Australian Transport Safety Board
BFU	Bundesstelle für Flugunfalluntersuchung
CFIT	Controlled Flight into Terrain
DSB	Dutch Safety Board
EASA	European Aviation Safety Agency
ECA	European Cockpit Association
ENCASIA	European Network of Civil Aviation Safety Investigation Authorities
EUR	Europe
FAA	Federal Aviation Administration
ICAO	International Civil Aviation Organization
ICC	Intraclass Correlation Coefficient
IQR	Interquartile Range
KPI	Key Performance Indicator
LOC-I	Loss of Control - Inflight
MAC	(near) Mid-Air Collision
MID	Middle East
MSA	Minimum Safe Altitude
NTSB	National Transportation Safety Board
PA	Pan America
PICAO	Provisional International Civil Aviation Organization
RASG	Regional Aviation Safety Groups
RE	Runway Excursion
SARP	Standard and Recommended Practice
SCF-NP	System or component failure - Non-Power Plant

SCF-PP	System or component failure - Power Plant
SHK (SAIA)	Statens Haverikommission (Swedisch Accident Investigation Authority)
SMS	Safety Management System
SPI	Safety Performance Indicator
SRM	Safety Risk Management
SSP	Safety State Programme

1

Chapter 1

General introduction

Contents

1.1. Research background
1.2. Research aim
1.3. Research questions
1.4. Research scope
1.5. Research approach
1.6. Outline of this thesis

General introduction

"Before anything else, preparation is the key to success."

- Alexander Graham Bell.

Safety investigations' classification and the allocation of investigative resources stand central in the elaboration of this work. One specific issue in today's investigation processes drives the necessity of this study: resources are allocated on events with *more* severe outcomes (Greenwell, 2003). In this thesis, the classification of occurrences in aviation, standard procedures of investigations and the selection criteria for deciding whether an investigation should be conducted formed the very foundation to perform this study. This research is utilising and analysing a new type of taxonomy, based on the controllability of a developing accident by the involved person(s) in trying to mitigate the outcome of such event, on a sample of safety investigation reports with respect to today's accident classification that is based on an occurrence's outcome to develop a selection criteria for conducting incident investigations. Furthermore, safety performance is widely measured through rates of adverse events (Bellamy & Sol, 2012), especially accidents, a practice that forms a deterministic approach and might not represent the real safety level of organisations and the aviation industry in overall since it does not count for the controllability of the event.

This study starts off with the description of today's practices in safety investigation to create a common ground for the understanding of the main problems in this study (Section 1.1). As the problems become more understood, the aim for this research is described (Section 1.2). In order to perform the study successfully, the right questions have to be asked first. In Section 1.3, the main and the most important sub research questions complement the research aim established in the section before. Aspects that will, but most importantly those that will not, be studied are described in the scope of the research (Section 1.4) and the approach to perform this research follows subsequently (Section 1.5). The concluding section will address the complete structure of this thesis (Section 1.6).

1.1. Research background

Safety stands top priority in the aviation industry and has been since the very beginning of aviation. It is consequently most often regarded as the safest mode of transport (Allianz, 2014). The constant drive to obtain absolute safety is apparent in today's safety performance of the industry. Accident rates (i.e. number of accidents per million departures or flight cycles) show a decreasing trend for, at least, the last 50 years. Moreover, the number fatal accidents remain stable throughout the same 50 year time period. This is a rather impressive achievement, considering that the risk exposure increases with the steady increase in air traffic (Boeing, 2014; Airbus, 2015). The year 2014 was praised with "one of the lowest [accident] rates on record." with an accident rate of 3.0 per one million departures (ICAO, 2015). Last year [2015] has been praised for being "by far the safest year ever" for fatal accidents in the aviation industry with a fatal accident rate of 0.2 per one million flights. With regards to the fatal accident rates in aviation, 2014 stands as the second best year behind 2015 (Flightglobal, 2016). Alongside the pursuance of safety are the accident investigations that seek to understand the causality of the event with the objective to prevent similar occurrences from happening. Accident and incident investigations are executed by the state of the occurrence's safety investigation board or authority (ICAO, 2010). Quoting that safety is indeed the safest it has ever been may not be entirely justified in terms of the origin of that statement. Accident rates are often used to illustrate the safety performance of a state or organisation, though these rates only focus on the outcomes and thus do not indicate related factors to why such accidents occurred (Roelen & Klompstra, 2012; Step change in safety, 2016). Moreover, today's accident rates are too low to provide sufficient information to state the safety performance in the industry (Airbus, 2015; SM ICG, 2013).

Occurrences in aviation are classified as accidents, serious incidents and incidents. Both ICAO (2010) and hence EASA (European Commission, 2010) define an accident as "an occurrence associated with the operation of an aircraft […] in which a person is fatally or seriously injured […], the aircraft sustains damage or structural failure which adversely affects the structural strength, performance or flight characteristics of the aircraft […] or the aircraft is missing or is completely inaccessible" (ICAO, 2010, p. 20). An incident is subsequently defined as an "occurrence, other than an accident, associated with the operation of an aircraft which affects or could affect the safety of operation" (ICAO, 2010, p. 21). And finally, a serious incident is defined as "an incident involving circumstances indicating that there was a high probability of an accident…" (ICAO, 2010, p. 21-22). All definitions of occurrences are based solely on the end-state or outcome of the event, without considering the development of such occurrences. ICAO hence states that the difference between an accident and a serious incident only lies in the outcome of the event (ICAO, 2010). States in which such occurrences occur are only obligated to conduct an investigation into an accident or serious incident. However if an incident can reveal safety-related information, then the applicable authority may conduct an investigation into the incident (ICAO, 2010; European Commission, 2010). Although investigating incidents is not required for the applicable authorities, incidents may provide better information to prevent (similar) accidents than actual accident investigations can (ICAO, 1993). Classifying an event solely on the outcome may be rather misleading in the fact that incidents can have the potential to become an accident, but fell short of the set criteria due to factors such as luck. In his research at a large aviation company, Karanikas (2015) created a new type of safety taxonomy that demonstrated the controllability of an event and with it the effectiveness of the involved personnel in trying to alleviate the outcome in the development of the event. The research showed that the outcomes of half of the occurrences were only a matter of chance (as indicated by uncontrolled events), but when an accident was controlled by the operator it showed that 87 percent of the time the outcome of the accident was alleviated. He concluded that this new classification scheme might function as a "more realistic" method to measure safety performance than the conventional outcome-oriented performance measurement (Karanikas, 2015).

When considering the human element in the causation of accidents in aviation, most studies agree that this element is by far the largest contributor. In fact, the same studies denote that about 70 percent of all aviation accidents involve a form of human-error (Shappell, et al., 2006; Shappell & Wiegmann, 2000; McRandle, et al., 2007; BASI, 1996). Despite the clear significance of the involvement of human error in aviation accidents, the performance and effectivity of the actions taken by the same people involved in such accidents are not included in the classification of these occurrences. In an article with Flight Global, Keith Conradi noted that today's investigation practices do not diagnose the latent factors (i.e. factors "present in the organization long before a specific incident is triggered" (Eurocontrol, 2006, p. 2)) that are important to the determination of the cause of human error in accidents. Conradi also stated that emphasis must be given on the successfulness of the air crew involved in these occurrences, rather than just only the human-error (Flightglobal, 2015).

The same structure of classifying occurrences based on the actual outcome lies within the public attention towards investigations and resource allocation of safety investigation authorities. A distinction can be made between two types of accident investigations: major accident investigations and incidents and non-major accidents investigations. The former requires a more comprehensive team with specialised investigators and consists overall of a larger investigation team than the latter, which can be conducted by "one or two investigators" (ICAO, 2015). These major accident investigations gain much more publicity and attention, due to factors such as blood priority (i.e. predominant attention due to a tragic event) by the media, grief for the loss of relatives and needs of manufactures for data whether corrective measures are necessary (Sarsfield, et al., 2000; ICAO, 2005; Dutch Safety Board, 2014; Flightglobal, 2015). Safety investigation authorities also have to manage and prioritise the allocation of limited resources for occurrence investigations. As a result, most investigation authorities base their prioritisation on the severity of the event. Hence incidents or accidents with less severe outcomes do not receive as much attention as major accidents do (Greenwell, 2003).

Before commencing investigations on accidents and incidents, the lessons expected to be drawn from them influence the resource allocation of the activity. Allocation of resources must therefore be justified prior to the conduction of the investigation. With the argumentation phrased before that incidents can provide better preventive measures than accidents, it seems that the investigation of incidents is indeed justified (Wise, et al., 2010). Moreover, occurrences of accidents are decreasing over time due to the improvements of safety in the aviation industry (Griffin, et al., 2015; Wise, et al., 2010). This means that the amount of lessons to be learned from such accidents have decreased over time. In addition, the amount of incidents is much greater than accidents; incidents should hence be regarded as the new form of accidents in occurrence investigations (Griffin, et al., 2015). Since the amount of incidents exceeds the amount of accidents, incidents can provide better information to detect trends and form a broader understanding of the problems at hand. Incidents do also consist of the same human behaviour that is present in accidents and can therefore provide such information before an accident occurs (Wise, et al., 2010).

Investigating accidents after they have happened is a *reactive* approach. Even though these practices have helped to understand the causes of such occurrences, an accident has to occur in order to identify the problem. Both Griffin et al. (2015) and Oster Jr et al. (2013) acknowledged the fact that preventative measures can be found by incident investigations. Incidents can hence be seen as the *proactive* measure for the prevention of accidents.

To summarise, classification of occurrences are solely based on the outcome of the event, which in return drives the publicity and resource allocation of safety investigations and is used to measure safety performance. Even though human-error is most often a contributing factor in accidents, it is not an explicit part in safety performance monitoring and measurement. These occurrences might have been a matter of luck or implicated an attempt to alleviate the severity of the outcome. Focussing (mostly) on accidents is rather misleading in the fact that incidents can provide better proactive measures than accidents. Nevertheless, the available resources to investigate incidents are still not used to the fullest today.

1.2. Research aim

The aim for this study is to utilise Karanikas' (2015) taxonomy, based on the controllability in the generation of a safety event, on a sample of investigation reports while comparing it with today's occurrence classification based on the severity of the outcome of such events. Contributing to the incentive to perform this research is the hypothesis that more severe events (i.e. defined by its outcome) are treated much more extensively than less severe ones. This is highly noticeable with events that are associated with significant fatality counts and catastrophic result, which *may* be amplified by the coverage from, for instance, the media. This overshadowing of incidents by more severe accidents is implicated in this study by the analysis of the length of safety investigation reports. Therefore will this study aim to confirm this hypothesis (i.e. the impact of the event based solely on the outcome) in the analysis of this research. In obtaining data from the controllability and length of safety investigation reports, further research shall be conducted to create selection criteria prior to conducting incident investigations.

To successfully accomplish this study, it starts off with a qualitative research by a literature review into the current classification system and its regulatory nature. Processes starting from selecting an occurrence to the completion of investigations are implemented in the review. The significance of incidents with respect to accidents will be discussed as well. Subsequently, the study continues into a quantitative research when investigation reports will be analysed.

The ultimate aims can finally be summarised into the following points:

- Determine the associated factors with the controllability of safety events
- Confirm the bias of safety investigation authorties towards the outcome of safety events
- Generate priority criteria for the selection of incidents before investigation
- Evaluate if accident controllability might be more representative of safety performance

1.3. Research questions

This study aims, in general, to indicate the factors related with the controllability taxonomy, as in whether the implicated person(s) attempted to alleviate the outcome of the event and the associated successfulness of the attempt, of Karanikas (2015) and the outcome of the respective occurrence. As this information stands central in the conduct of this study, it is expressed in the following **main** research question.

RQ *What factors are associated with the controllability of safety events in aviation?*

Following the results of the analysis, the extent of safety investigations shall be derived from the length of the investigation report. As major accidents gain more publicity and gets prioritised by means of the number of allocated resources and depth of the inquiry, the length of such investigation reports may support the underlying hypothesis. Identifying bias by means of the length of safety investigation reports is decided based on the availability of these reports and recordability of this information (i.e. word count is always recordable). Therefore will the length of the reports be analysed against the controllability taxonomy and severity classification of occurrences to reveal any relations. This is formulated in the following sub-research question.

RQ_1 *Is there a relation between the frequency analysis of the new taxonomy, severity classification of the occurrences and analysis of the length of accident investigation reports?*

The factors that contributed to the main research question ought to be further explored and applied for consecutive research. By exploiting the information obtained in the analysis of both taxonomies, the contributing factors may be applied to the selection process prior to the conduction of safety investigations. In particular, the investigation of incidents has been regarded as a proactive and preventive measurement for accident occurrence. As incidents are important to the prevention of recurrence/occurrence of accidents, priority criteria shall be created, or an attempt to it, by utilising the controllability taxonomy. The possibility must be explored followed by the implementation in today's investigation practices. This leads to the second sub-research question.

RQ$_2$ *What factors are associated with the frequency analysis and can be further explored to indicate priority criteria for conducting investigations?*

As Karanikas (2015) concluded, the controllability taxonomy might represent a company's or state's safety performance in a more realistic method. Half of the analysed occurrences were depicted as uncontrolled and hence dependent on 'luck' (p. 188). It is rather troubling when safety performance is measured on the outcomes of such occurrences when half of those did not implicate a control attempt and depended therefore on the development of these events. Whether this proportion holds true on a large sample of safety investigation reports from multiple sources, is yet to be examined and stated. Hence the possibility to introduce the controllability taxonomy for better measurement of safety performance in the aviation industry must be explored. Exploring the use of the taxonomy as a safety performance measurement is expressed in the final sub-research question of this thesis.

RQ$_3$ *Can safety performance be measured with the controllability of a safety event and does it enable for a better representation of safety performance?*

1.4. Research scope

Safety investigation is apparent in every major transport industry and practices are similar, yet this study will only focus on the safety investigation practices of the aviation industry. Aviation itself is a diverse industry with multiple interests, from leisure to commercial, with a wide array of different types of transport air vehicles. An occurrence to each of these different types of air vehicles and operation type can all be reported in a final report, which will function as the main source for the study. These investigation reports must obviously be made publicly available to incorporate them in this research. Therefore shall this research aim at all of types of air transport occurrences, regardless of its scale, but will exclude occurrences that only include military activity (due to its restrictive nature). Research of investigation is done solely with investigations' respective final reports.

As the very notion is to utilise both classifications (i.e. current industry classification and controllability taxonomy), all incidents and accidents investigation reports shall be incorporated in the research. Length of such reports will be a key factor in proving the hypothesis stating that more eventful accidents gain more attention and resources than smaller ones.

1.5. Research approach

This study consists of three complementary phases: The exploration phase, research phase and finalisation phase. The first phase entails the exploration of information regarding the execution of today's safety investigations. Literature concerning the preceding researches, current industry problems and opportunities provide the student with sufficient data to create a sound methodology before conducting the research.

Within the research phase, the safety investigation reports of various safety investigation authorities will be analysed with both safety taxonomies by utilising the methodology from the first phase.

Results from, and relevant factors to, this research must be further discussed to provide productive meaning to the findings. First, answers to the research questions will be discussed and concluded. The information that can be extracted from the meaning of these findings and controllability taxonomy will be used to generate priority criteria.

1.6. Outline of this thesis

Chapter one Introduction outlined the research background of this specific study. As the central problem is identified, the proposed means to rectify or answer this problem is discussed. This included the objective of the study which is solidified by means of research questions. And in order to attain quality, the research approach and its confinements were settled.

Chapter two Literature review introduces the study with scientific information concerning the practices related to safety investigations, safety performance measurement and indication and the controllability taxonomy itself. This chapter provides essential information for the evaluation of the new taxonomy, but also the incentive of instituting investigation for the creation of priority criteria.

Chapter three Methodology focusses on the methods used in this research to provide the answers to the hereinbefore mentioned research questions. Factors that might show association with the taxonomy and indicate bias of safety investigation authorities towards an event's severity are discussed in detail.

Chapter four Results highlights the results from the analysis of safety investigation reports. Chi-square analyses are performed to indicate statistical differences among the selected variables of the previous chapter.

Chapter five Discussion evaluates the results to provide meaningful information that will be used to assess the taxonomy's capability as a safety performance indicator and selection priority criteria.

Chapter six Conclusions and Recommendations is the final chapter of this report that evaluates the answers to the research questions, and statement on how the taxonomy can be utilised as safety performance indicator. It will furthermore present the selection priority criteria as new means to select events for investigations before assessing its severity. Finally, recommendations will be made with respect to the use of the taxonomy and priority criteria.

2

Chapter 2

Literature review

Contents

2.1. International standard in safety investigations

2.2. The safety investigation process

2.3. Data considerations

2.4. Safety performance measurement in civil aviation

2.5. Safety event's controllability as performance indicator

2.6. Concluding remarks

Literature review

> *"The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them."*
>
> - William Lawrence Bragg.

This chapter focusses on creating a general understanding of the purpose and practice of safety investigations, but also on how safety may be defined. It must be clear that safety investigations are subject to regulatory requirements. However, not all safety events in aviation are required to be supported with such inquiry, which is dependent on a statutory occurrence classification. Needless to say, safety investigations that are not required by legislation may still be performed. Before investigations are instituted however, they ought to be selected by means of selection criteria under various circumstances. As the new taxonomy shall provide the means to create selection criteria, this information is essential in the understanding what criteria may be present in today's selection processes. This taxonomy will also be applied to illustrate whether it comprises as a more representative measure of safety performance. In this respect, safety performance measurement utilised in aviation ought to be discussed.

The second chapter begins with what is important to the universal establishment of safety investigation requirements and practices that is set forth by the Chicago Convention (Section 2.1). These requirements are subsequently elaborated through the discussion of the processes of safety investigation, including the selection criteria for the determination of instituting investigations (Section 2.2). Data considered for this study is the frequency of safety events in which less severe incidents are more frequent than more severe one, hence are more considered by selection criteria. In addition, safety investigations are associated with (final) reports which will be found useful for the application and analysis of the new taxonomy (Section 2.3). Furthermore, safety performance practices are described in Section 2.4. Information from all preceding sections supports the understanding and aim for the new taxonomy, which is described in Section 2.5. Finally, chapter 2 will be concluded in Section 2.6 and further steps for the subsequent chapter shall be discussed.

2.1. International standard in safety investigations

This section incorporates the international establishment of safety investigations and its regulatory conditions. The classification of occurrences, as defined by the International Civil Aviation Organization (ICAO) and adopted by (most) of its contracting states, is discussed. In association with the presented standard, current European regulations regarding safety investigations and its ultimate objectives are elaborated subsequently.

2.1.1. International Civil Aviation Organization

On November 7, 1944, 54 States assembled in Chicago to attend the International Civil Aviation conference. This conference was set out to arrange an "immediate establishment of provisional world air routes and services" and constitute a council that supervises the international civil aviation with regards to safety and further improvements. At the end of the conference, on December 7, 1944, 52 states signed the Convention on International Civil Aviation, better known as the Chicago Convention. Awaiting the ratification of 26 contracting states, interim body the Provisional International Civil Aviation Organization (PICAO) was established on June 6, 1945 after an interim agreement on International Civil Aviation. On April 4, 1947, the permanent body the International Civil Aviation Organization (ICAO) was established after the 26th State ratified the Chicago Convention. On October 3, 1945, ICAO became a specialized agency of the United Nations (ICAO, 2016).

The International Civil Aviation Organization was set out to establish international cooperation of all contracting member states and uniformity in Standards and Recommended Practices (SARPs) concerning civil aviation. Pursuant to article 37 of the Chicago Convention, which implicates that ICAO publishes international SARPs to establish international uniformity (i.e. in procedures, standards, regulations and organization), ICAO publishes these Standards and Recommended Practices by means of 19 Annexes (ICAO, 2006). For this study in particular, Article 26 forms the basis for Annex 13: Aircraft Accident and Incident Investigation (ICAO, 2010), while Annex 19: Safety management provides SARPs specific to the safety management activities for both the aggregate state and individual level (ICAO, 2013a). It should be noted that the contracting states that signed the Chicago Convention are obligated to comply with it. In contrast, these contracting states are not required to implement the SARPs if the differences from the international standard are notified to the Council (ICAO, 2006).

2.1.2. Annex 13 a guide for independent safety investigations

Pursuant to article 37 and subject to article 26 of the Chicago Convention, Annex 13 provides standards, recommended practices and procedures for conducting investigations of accidents and incidents in ICAO's contracting member states. In order to prevent similar accidents and/or serious incidents from recurring, the causal factors ought to be defined. Annex 13 provides guidelines and practices for the investigations of such occurrences with the main objective to prevent future occurrences without apportioning blame or liability (ICAO, 2010; ICAO, 2006).

Article 26 of the Chicago Convention dictates that accidents involving death, serious injury or serious technical defect in the respective aircraft or air navigation facilities are to be investigated by the state in which the accident occurred (ICAO, 2006). This translates into the definition of three types of occurrences: accident, serious incident and incident. Therein are accidents and serious incidents compulsory for investigation, excluding the obligation to investigate incidents. ICAO (2010) defines an accident as "an occurrence associated with the operation of an aircraft […] in which:

a. a person is fatally or seriously injured as a result of:

- being in the aircraft, or
- direct contact with any part of the aircraft, including parts which have become detached from the aircraft, or
- direct exposure to jet blast,

Except when the injuries are from natural causes, self-inflicted or inflicted by other persons, or when the injuries are to stowaways hiding outside the areas normally available to the passengers and crew; or

b. the aircraft sustains damage or structural failure which:

- adversely affects the structural strength, performance or flight characteristics of the aircraft, and
- would normally require major repair or replacement of the affected component,

Except for engine failure or damage, when the damage is limited to a single engine (including its cowlings or accessories), to propellers, wing tips, antennas, probes, vanes, tires, brakes, wheels, fairings, panels, landing gear doors, windscreens, the aircraft skin (such as small dents or puncture holes), or for minor damages to main rotor blades, tail rotor blades, landing gear, and those resulting from hail or bird strike (including holes in the radome); or

c. The aircraft is missing or is completely inaccessible. (p. 20)"

A serious incident is defined as an occurrence that had the "high probability" of becoming an accident. Finally, an incident is defined as an occurrence that affects or could affect the safety of the operation. One can unquestionably separate accidents from (serious) incidents, as the difference lies in the outcome of the event. However the differences between an incident and serious incidents are less obvious. Annex 13 therefore provides a list of eighteen sample scenarios of serious incidents to help identify an occurrence its classification. Annex 13 furthermore provides clarity to whom an occurrence is responsible and which contracting state should conduct the investigation. The accident investigation authority of a contracting state must have independence in conducting investigations. These investigations (normally) consist of the collection of accident data that provides the investigation authority to analyse the occurrence to determine causal and/or contributing factors, the issuance of safety recommendations and a final report (ICAO, 2010).

2.1.3. The European Aviation Safety Agency

The European Aviation Safety Agency (EASA) is an agency of the European Union that provides among other things its member states with implementing rules to ensure a high level of safety. By better harmonisation of the member states, improvements in independence of investigation authorities, better protection of victims' rights and the improved implementation of safety recommendations, Regulation (EU) 996/2010 repeals Directive 94/56/EC and fills the need to improve the efficiency of safety investigations in Europe (Commission of the European Communities, 2009; European Commission, 2009; European Commission, 2010a). Alongside the new regulation is the establishment of the European Network of Civil Aviation Safety Investigation Authorities (ENCASIA) to increase the efficiency of the coordination of safety investigation authorities (ENCASIA, 2014). On November 15, 2015, this regulation has been amended by Regulation 376/2014 that enhances the 'just culture' reporting of occurrences by enabling open occurrence reporting without punishing the aviation professionals on his or her actions. Safety investigations must be conducted to prevent similar occurrences from happening and to ensure a high common level of safety in aviation without apportioning blame or liability, the human element is only a part in a complex system. These safety investigations are conducted by safety investigation authorities who stand as central (independent) investigative bodies who ultimately determine the causes of the occurrences. Important to this specific study, the scope of an investigation *should* be dependent on the expectations of the lessons to be learned from the occurrence with the underlying principle to improve safety thus considering the cost-efficiency of the utilisation of resources in the European Union (European Commission, 2014).

2.2. The safety investigation process

Significant safety investigations are performed by the state's independent safety investigation authority. Investigation authorities function independently to provide the causes of occurrences without the implication of other interests (e.g. to allocate blame or liability) and external interference in the investigative process (European Commission, 2010a). Some safety investigation authorities (e.g. ATSB, NTSB and DSB) have been established as multimodal safety boards that not only focus on the aviation industry, but expanded upon transportation modes such as marine and railroad (Stoop, 2003). ICAO (2011) identifies investigating practices of these safety boards in three distinct phases: collection of data, analysis of data and presentation of findings. Stoop (2003) identified five distinct phases for a successful execution of investigations. With respect to the three aforementioned phases, Stoop incorporated the initiation phase and the feedback phase, which can be put respectively prior and subsequent to the phases. However, before an investigation can be initiated, it should first be notified to and selected by the respective investigating body (i.e. initiation phase). Furthermore, after an investigation has presented and published its findings, investigation authorities monitor the actions set forth by the publication of its safety recommendations (i.e. feedback phase).

Before an investigation can be conducted, it must be notified by the investigation authority to initiate deciding whether the occurrence should be investigated or not. Occurrences are generally notified directly to the respective authority, but due to advances in technology authorities may be notified through sources as (social) media. Safety investigations' main purpose is to prevent similar occurrences from happening in the future. Effective allocation of resources is therefore spent on the investigations that will provide the most information concerning the prevention of recurrence and overall improvements in aviation safety. Safety recommendations are formulated, on the basis of the analysis and determination of the root causes of such occurrences, to effectuate this very objective. A finite number of investigators limits the possibility to investigate all events that may comprise of relevant information for the safety in aviation.

As a result, not all relevant occurrences that can be investigated will be investigated. Deciding whether an occurrence should be investigated can be based on the obligation and the authority's interest. As previously mentioned in paragraph 2.1.2, all accidents and serious incidents must be subject of an investigation by the respective state. Of more interest to this study is the selection of occurrences based on the interest of the respective authority. An investigation authority can also be requested to conduct an investigation. In the Netherlands, for example, an investigation can be initiated by the Minister, Commissioner of the King or the Mayor[1]. Furthermore, the authority may institute a thematic investigation that comprises a series of occurrences, generally incidents (Onderzoeksraad voor Veiligheid, 2012).

The Dutch Safety Board selects the investigation of occurrences based on (positively) answering questions including social commotion, public exposure and safety, signals of systemic errors and the authority's (avail)ability and its significance in the investigation. An investigation will be instituted if the respective occurrence yields enough favourable arguments to justify the means of conducting the investigation (Onderzoeksraad voor Veiligheid, 2012). A sample analysis has been performed by Lysias consulting group (2014) for the indication of the reasoning of the Dutch Safety Board for selecting occurrences to initiate safety investigations. The analysis showed that most of the arguments for selecting an occurrence were the severity of the event and the social impact. The severity is most often discussed by the number of fatalities or injured and the material damages. Severity is also associated with the classification as whether an investigation is obligated. Events that implicate social commotion, public exposure and safety are related to the reasons for selecting an occurrence based on social impact. It is also noticed that some of these cases implicated much attention of the media. These selection criteria are followed by the expected safety gain and the obligation to perform an investigation. The evaluation concluded that the Dutch

---

[1] Kingdom Act, 2 December 2004, instituting a Safety Investigation Board (Kingdom Act concerning Safety Investigation Board), Art. 43(1).

Safety Board most often focusses on the severity of the occurrence, without mostly relating the selection criteria with the pursuance of the increase in aviation safety. It becomes rather difficult to decide if an investigation should be initiated upon the expectation that it will provide gains in safety (Wise, et al., 2010). Selection is hence based on limited information of the occurrence, which makes it even more difficult to choose between multiple occurrences when determining which one provides the better information (Onderzoeksraad voor Veiligheid, 2012).

From a different perspective, the Australian Transportation Safety Board (ATSB) prioritises its accident investigations on the enhancement of public safety and matters that are of widespread public interest. The ATSB also acknowledges the fact that (serious) incidents may comprise of better information for the enhancement of aviation safety and does prioritise the investigations accordingly. Alongside the gain of learnable lessons and the public's safety and its interest from the investigation for the improvements in safety, the ATSB allocates its resources to the type of operation. These operation types are based on the aircraft's physical scale and potential public exposure. Large passenger aircraft stands as a result at the top, followed by small passenger aircraft and general aviation type of operations (e.g. aerial work, recreational, etc.) (ATSB, 2016).

Considering the National Transportation Safety Board (NTSB), this board is obligated to investigate all accidents and incidents involving civil aircraft or certain public aircraft (Code of federal regulations, 1997). However, the NTSB is limited in amount of resources to conduct all accidents in all of its responsible modes (e.g. aviation, highway, marine, railroad, pipeline and hazardous materials). Over 90 percent of all completed investigations concerned aviation, which affects the available resources of the remaining modes (GAO, 2006). In support of restrictive resources, the Federal Aviation Administration (FAA) may provide investigative resources by (partly) conducting investigations for the NTSB (Code of federal regulations, 1997). Considering the selection of accidents or incidents to investigate, a four-tier system is utilised to prioritise resources on risk based criteria for the highway transportation mode. These criteria incorporate the severity and property damage of the respective accident in allocating resources. The marine mode investigates major accidents involving, among other things, six or more fatalities and/or property damage equal to or more than $500,000. Other accidents may be selected based on safety risk of third parties, involving a fatality or property damage equal to or more than $75,000. Accidents in the remaining modes are only selected for investigation based on the judgement of the office director, considering factors such as involvement of a fatality, significant property damage or environmental damage (GAO, 2006).

The second phase entails the on-site investigation of the accident or incident, in which the main objectives is to record the accident (or incident) site and conduct interviews with witnesses and involved persons (Onderzoeksraad voor Veiligheid, 2012). All data relevant to the accident must be determined and obtained in the process, wherein data that is perishable must have priority. Four types of data are considered that are to be collected in this phase: accident specific, meteorological, technical and human factors. Analysis of the data is performed alongside the collection of data. As the analysis of data can generate more questions in the process, additional effort is needed in the collection of data to answer them. This back-and-forth process institutes discussions within the investigation team necessary for successful completion of both phases (ICAO, 2011).

Each investigation is completed by means of an accompanying report, though the format and extent of the report depends on the nature of the investigation. Before the actual final report is published, an interim report is sent to the applicable authorities to request feedback on errors and uncertainties in the report (Onderzoeksraad voor Veiligheid, 2012; ICAO, 2014a). Once a final report is finished it is sent to the applicable authorities (ICAO, 2010). The main purpose of safety investigation is to prevent future occurrences from recurring and thus enhancing the aviation safety and in order to effectuate this, safety recommendations are published (ICAO, 2014a).

Finally, the investigation authority enters the feedback phase once the final report and its recommendations have been published. The safety investigation authority monitors the actions taken by the addressee(s) of the safety recommendation(s) and the conception on the transportation safety (Stoop, 2003). Within 90 days on the reception of the safety recommendation, the respective addressee will reply the actions taken, actions still under consideration and the actions not taken. In response to this reply, the investigation authority will state the adequacy on the actions taken or disagreement within 60 days of receipt (European Commission, 2010a).

2.3. Data considerations

This study considers various data types and sources. Before analyses may be conducted on these types of data, the origin must be understood thoroughly. It has been noted that incidents may comprise as better preventative measure for the occurrence of more severe accidents (See Section 1.1). In support of this notion, incidents are more frequent than accidents. This section discusses the frequency relation of safety events based on their occurrence classification. In addition, safety investigations are associated by final reports that consist mostly of a standard format.

2.3.1. Frequency relation of safety events

Aviation accidents are “luckily” rare events and using these events as preventive measures for future occurrences is a reactive approach to what could have been prevented by the proactive approach of investigating incidents before the actual accident occurred. That is, incident investigations can provide better preventive measures for the occurrence of accidents (ICAO, 2005; Wise, et al., 2010). With respect to accidents, incidents are far more common in quantity, also often expressed in the accident triangle. First introduced by Heinrich in 1931 and later recalculated by Frank E. Bird Jr., the triangle indicates the ratio of one major or fatal injury with regards to less severe occurrences (Figure 1) (ICAO, 2005; Davies, et al., 2004; Heinrich, 1931; Bird, 1966).

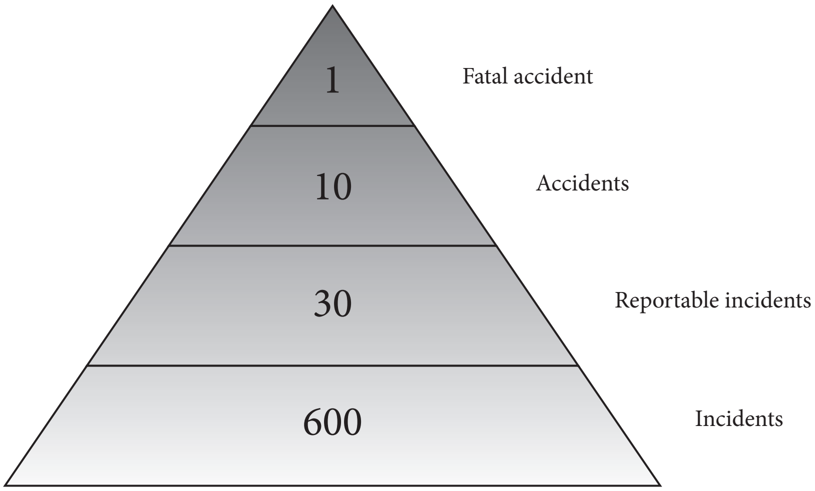

Figure 1. F.E. Bird’s accident triangle
Source: F.E. Bird. (1966). *Damage control.* Philidelphia: Insurance company of North America.

The accident triangle defines that for every 600 reported occurrences, 30 of which were incidents that involved property damage (i.e. severe enough to be accounted for), 10 of which were accidents that involved serious injuries and one of which that involved major or fatal injury (ICAO, 2005; Davies, et al., 2004). Incidents' significance is not only related to the sheer quantity, as incidents tend to have less publicity and may enable better obtaining of information as evidence is better preserved and available (e.g. little to no damage to evidence, live witnesses and involved persons). It provides the investigators of incidents with better circumstances to identify the causation of the incident and how the incident avoided from becoming an accident (ICAO, 2005). The quantity of incidents enables to create a wider understanding of problems and trends (Wise, et al., 2010). Though stating that all 600 incidents provide valuable information for the acquisition of knowledge of trends and problems is not entirely justified. To use incident data for the prevention of accidents, these occurrences need to have similar or common causes, which are discussed in the 'common cause hypothesis' (Davies, et al., 2004). If the causes of the less severe incidents are similar to the causes related to major accidents, then the prevention of major accidents may be expected by the preventive measures taken on the basis of less severe incidents. Davies et al. (2004) indicated that the common cause hypothesis is "largely" validated by the study of Wright (2002) which showed that the hypothesis holds strong and that near-misses (i.e. incidents) indeed can be used to identify more severe accidents (Davies, et al., 2004; Wright, 2002).

2.3.2. Safety investigations associated report file

The very purpose of conducting investigations is to prevent similar occurrences from recurring in the future. Concluding each investigation is the obligation to issue a final report that helps to prevent this recurrence. It should answer the 'what, how and why' questions to establish a sound foundation for the establishment of the occurrence's causality and future prevention. Formally, each final report contains safety recommendations from where preventive actions can be drawn from (ICAO, 2014b). Even though the publication of safety recommendations in a final report is not obligatory, a safety investigation authority shall communicate (preventive) recommendations to the applicable authorities for the improvement of aviation safety (European Commission, 2010a). Each final report should address the relevant facts, the analysis thereof, the conclusions and safety recommendations (ICAO, 2014b). The format of such final reports is not legally binding, but is found to be generally consistent over multiple final reports of various safety investigation boards (AAIB, 2014; De onderzoeksraad voor veiligheid, 2010; BFU, 2010; SHK, 2013; AAIU, 2012). As published in Appendix II of Annex 13 (2010), a final report's format contains six distinct sections. The core of the final report shall give answers to the 'what, how and why' questions and is structured as such:

1. Factual information
2. Analysis
3. Conclusions
4. Safety recommendations
5. Appendices

The first section contains the formal introduction to the investigation (report) with the title that should indicate the name of the operator, the manufacturer, model and registration of the aircraft, nationality, place and the date of the occurrence. Furthermore, this section should contain a synopsis detailing the relevant information in short format. In the first section of the informational body of the final report, the very foundation must be formed from where the analysis, conclusions and safety recommendations will be derived from and expanded upon in the subsequent chapters. Relevant facts are stated of the investigation of the occurrence. Factual information may include topics such as the flight history, injuries to persons and damages to aircraft or objects, personnel, organisational and aircraft information, navigation, communication, meteorology, flight recorders, medical and test data results.

Followed up by the stated facts of the investigation, the analysis connects the factual information with the ultimate conclusion(s). The facts and circumstances are discussed and further analysed in this section to determine the causal factors. In addition, the evidence from the prior section is evaluated to generate hypotheses which are in return tested against the evidence to cancel out unsupported hypotheses.

The third section contains a summative list of findings, the causal and/or contributing factors from the preceding sections of the investigation. As the objective of an investigation is to prevent similar occurrences from happening, safety recommendations provide information to obtain this very intention. Similar to the conclusions, safety recommendations are concisely stated in the fact that it is in listed format. And finally all relevant data that complements information in the report and is helps the understanding of the investigation is appended (ICAO, 2010; ICAO, 2014b).

2.4. Safety performance measurement in civil aviation

The aviation industry is notoriously known for its impressive safety performance as depicted in statistics such as low accident and fatality rates (Allianz, 2014; Flightglobal, 2016; ICAO, 2015a). However, before such statements of safety performance can be argued, the definition of safety should be clearly understood. The term *safety*, though universally used, is used without questioning the exact definition and can hence be defined in varying contexts (Hollnagel, 2014). Hollnagel (2014) dictates a generic definition of safety as "…the system property or quality that is necessary and sufficient to ensure that the number of events that could be harmful to workers, the public, or the environment is acceptably low" (p. 1-2). Specific to aviation, ICAO (2013a) defines safety as "the state in which risks associated with aviation activities, related to, or in direct support of the operation of aircraft, are reduced and controlled to an acceptable level" (p. 16). Risks, in this context, comprise the probability and severity of outcomes from hazards or situations (ICAO, 2013b). Both definitions of safety, despite their difference in application, are subjective in nature. One's perception of an *acceptable* level of safety can differ greatly from another one's perspective. Elaborating on ICAO's definition of safety, the performance of safety must be clear as well. In this regard, ICAO (2013a) cites safety performance as "a state or a service provider's safety achievement as defined by its safety performance targets and safety performance indicators" (p. 16). In conclusion, safety is defined in terms of exposure to risk that is controlled to an acceptable level, while the performance of safety is expressed by means of specific targets and indicators set by the service provider, indicating a certain freedom in defining a state's or organisation's actual level of safety.

2.4.1. Introduction to safety performance measurement and monitoring

The demonstration of safety performance is generally applied on an aggregate state level, by establishing a State Safety Programme (SSP), or on a service provider's individual level, by means of a Safety Management System (SMS). Under the SSP, aviation service providers are required to implement such SMS to manage and improve safety on an organisational or individual level. Each state sets out requirements for the establishment of safety management systems so that the essential safety hazards are identified and risk management controls are implemented. Safety information can thereby be more effectively exchanged between other states or service providers. Also, each state should reach an agreement with service providers' SMS safety performance indicators and target and alert levels and should be periodically reviewed against the suitability of its safety performance. In this respect, the state has an oversight function by monitoring the (effective) implementation of safety management systems of service providers for maintaining and continuously improving their safety performance.

Following up on the definition and measurement of safety performance, an acceptable level of safety performance (ALoSP) is subsequently defined by a minimum threshold of the same safety performance targets and indicators. Indicators are used to demonstrate, most often in a quantitative manner, a state's of service provider's safety performance, which usually consist of high consequence indicators (e.g. rates of accidents). To indicate whether the recorded performance is acceptable or not, certain target and alert levels are established to quantify the acceptable and unacceptable regions of the respective safety performance. An alert level is normally determined based on the average performance of the preceding year(s) as multiple alert levels are defined by the standard deviation from this average. Setting alert levels allows premature detection of deviations of the targeted safety performance and breaching those levels demands finding out the causes and determining consecutive course of action to correct the deviation. More straightforward is the target level of safety performance. As the SSP and SMS both aim to continuously improve safety performance, the target level is usually set at a lower figure than the preceding year, or rather monitoring period, in terms of percentages (e.g. 5% lower than the preceding year) (ICAO, 2013a) (See Figure 2).

Even though the acceptable level of safety performance of a service provider or state can be expressed by means of performance targets and indicators, the true meaning of acceptable safety performance is contextual and is determined by the state in an SSP. ICAO (2013b) cites that "the State's ALoSP criteria may vary depending on the specific context of each State's aviation system and the maturity of its safety oversight system" (p. 78). Consequently, the essence of safety performance measurement and monitoring is aimed at the reduction of severe safety events and the constant improvement of safety performance.

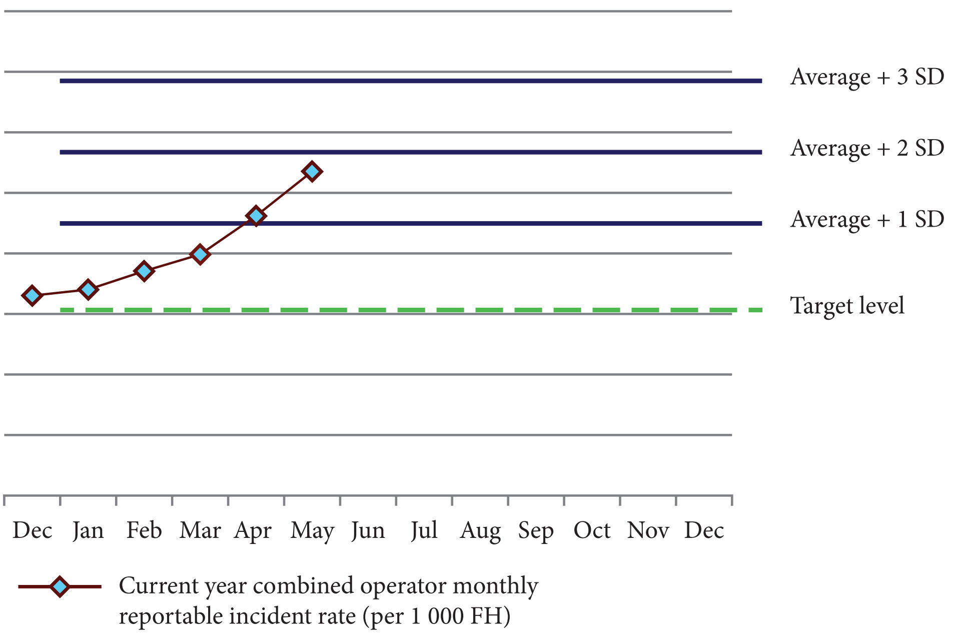

Figure 2. Example safety performance target level settings.
Source: ICAO. (2013). *Annex 19 Safety Management*. First Edition. Montréal (Quebec): International Civil Aviation Organization. Note. SD is Standard Deviation

Safety performance is furthermore often associated with safety culture as an integral component of the safety management system (Roelen & Klompstra, 2012). Defining safety culture has been proven to be a difficult task as there is no universal definition of the term (Wise, et al., 2010; Stranks, 1994; Davies, et al., 2004). Reiman & Pietikäinen (2012) define safety culture[2] as "the ability and willingness of the organization to understand safety, hazards and means of preventing them, as well as ability and willingness to act safely, prevent hazards from actualising and promote safety". Air Navigation Service Providers (ANSPs) are required, in this respect, to measure its safety management effectiveness as a safety (key) performance indicator (SPI or KPI), which is subsequently measured through the level of implementation from its management objectives, including safety culture. Another safety performance indicator required for ANSPs is the measurement of just culture (European Commission, 2011; European Commission, 2010b). Just culture, also known as reporting culture, is defined by the European Commission (2014) as "a culture in which front-line operators or other persons are not punished for actions, omissions or decisions taken by them that are commensurate with their experience and training, but in which gross negligence, wilful violations and destructive acts are not

---

[2] Due to underlying complexity of defining safety culture, the understanding of it shall be confined to this single definition.

tolerated" (p. 8), in which personnel are encouraged to report such information for improvements of the organisation's safety (performance). It is therefore an integral part for the identification and hence the prevention of hazards through the reporting of safety related information. Considering that reporting safety related information is part of the willingness to improve safety and thus the cultural behaviour of personnel, just culture is an essential component of the more comprehensive safety culture (Reason, 1997; European Commission, 2014).

2.4.2. *Safety performance data collection*

Safety (performance) cannot be effectively measured and acted upon without the essential collection of safety related data (Hollnagel, 2014). ICAO (2013a) cites that safety data "may include accidents and incidents, events, non-conformance or deviations and hazard reports" (p. 34). As safety investigations are reactive by their very nature they can provide information for the improvement of the aviation (safety) system, since causes and hazards are identified and determined following an occurrence. On a proactive perspective, hazards are continuously sought, recorded, analysed and assessed against safety risks as essential part of an SMS. This process of identifying hazards enables the service provider to address and mitigate the safety risks that may lead to higher consequence occurrences before such events occur. Known as the safety risk management (SRM) in an SMS, set safety performance targets are achieved through the management of safety related risks. (Figure 3).

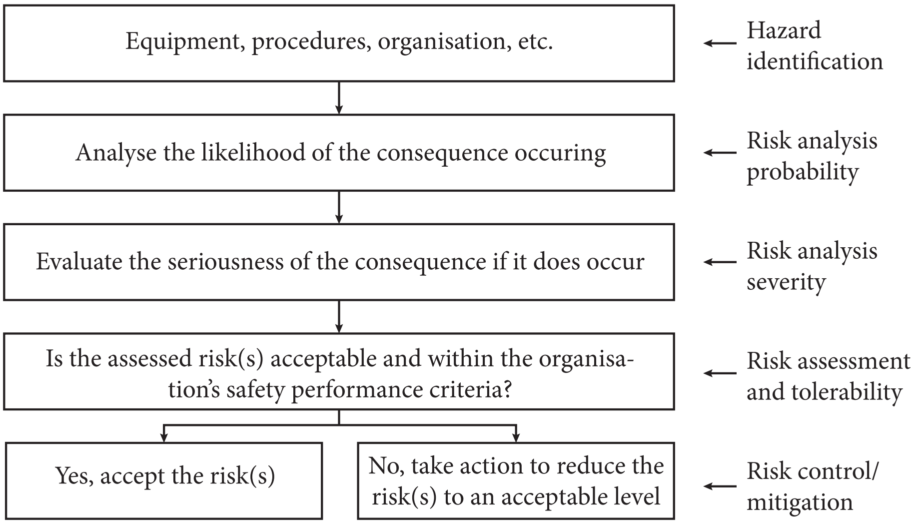

Figure 3. Safety risk management table for a Safety Management System.

Source. ICAO. (2013). *Doc 9859, Safety Management Manual (SMM)*. Third Edition. Montréal (Quebec): International Civil Aviation Organization.

As part of the safety risk management, ICAO (2013b) classifies three types of methodologies for the identification of hazards: reactive, proactive and predictive. Reactive hazard identification uses past occurrence data, such as incidents and accidents that are obtained by the investigation of such occurrences to identify hazards apparent in those occurrences. The proactive method uses information of the current situation to identify hazards through processes such as audits and reporting systems. And lastly, the predictive method uses data gathering to identify possible future hazards and therefore to mitigate the adverse outcomes consequent to those hazards.

To obtain information about hazards and thus to measure its safety performance and consecutive risk mitigation proactively, service providers have set up voluntary and mandatory reporting systems. Furthermore, safety performance measurement and monitoring may be derived from safety studies, safety reviews, safety surveys, audits and internal investigations. Safety performance indicators are consequently determined and selected based on the safety performance measurement and monitoring of these information sources (ICAO, 2013b).

2.4.3. Safety performance indicator types

Safety performance indicators provide a state or service provider with information about the effectiveness of its safety management (system) and the indication of emerging safety risks. It is also used to monitor any known safety risks and as deciding factor for corrective actions (ICAO, 2013b). For realising these specific purposes of safety performance indicators, there are several ways of classifying them. Safety performance indicators are most often classified as either leading or lagging indicators (SM ICG, 2013). Lagging indicators focus reactively on occurrences after they have happened and can hence measure the past safety performance of the respective system that is being monitored. Corrective actions for improvements in safety are therefore only executed after the information of occurred events is available (Hopkins, 2009; O'Connor, et al., 2011; Reiman & Pietikäinen, 2012). Consequently, these indicators use negative outcomes of events to demonstrate a system's safety performance in terms of rates and trends (SM ICG, 2013). Though measuring occurrences based on their outcome leads to the issue known as the 'regulator paradox', which implies that when safety increases, less data is present to measure it. Safety performance in aviation on a regional and global scale is most often depicted in terms of accident or fatality rates in annual safety reports of multiple sources (ICAO, 2013c; O'Connor, et al., 2011; Flightglobal, 2016; ICAO, 2015a), hence global accident rates are the primary indicators in aviation (ICAO, 2013c). As mentioned before, aviation accident rates are at an all-time low and as these safety events are declining in number, less data becomes available to measure it. Outcome-based metrics thereby do not have sufficient amount of data to be an effective indicator of safety performance and decreases in reliability (Hollnagel, 2014). Leading indicators, on the other hand, focus actively by identifying areas that should be improved and can be used to anticipate occurrences before they actually happen (Wreathall, 2009; O'Connor, et al., 2011). These leading indicators can therefore provide premature information to measure current, and proactively improve and predict future, safety performance of a system. The Dutch National institute for Public Health and the Environment (2012) indicates that leading indicators could include safety related data such as of reported deviations to loss events, inspections, personnel culture surveys, and planned/action carried out. Considering both types of indicators, the use of lagging indicators can demonstrate the validity of leading indicators as they provide quantitative proof by, for instance, occurrence measurement of accidents (Wreathall, 2009; O'Connor, et al., 2011; Hopkins, 2009; Bellamy & Sol, 2012).

ICAO (2013) distinguishes safety performance indicators for the individual (safety management system) level with high or low consequence indicators. High consequence indicators, also known as reactive indicators, use metrics of severe occurrences such as accidents and serious incidents. In contrast, low consequence indicators, also known as proactive and predictive indicators, use metrics of incidents and activities. ICAO furthermore states that these performance indicators "are essentially data trending charts that track occurrences in terms of event rates", in which the high consequence indicators should be implemented prior to the lower consequence indicators in a safety management system.

EASA (2014a) has classified common safety performance indicators for aggregate state level SSPs through the Network of Analysts Safety in terms of two tiers. Tier 1 focusses on accidents per operation type (i.e. commercial air transport large aeroplanes, commercial large helicopters, commercial light helicopters, other commercial fixed wing and private flying). And the second tier focusses on the three severity occurrence classifications (i.e. accident, serious incident and incident) and its associated occurrence type. This tier is aimed at the identification of precursor occurrences, namely incidents. Specific safety performance indicators are selected for major accident categories, such as controlled flight into terrain (CFIT) or runway excursions (REs). These accident categories can subsequently be identified using (precursor) indicators related to those events. For example, a descent below Minimum Safe Altitude (MSA) is related to CFIT accidents and loss of control on ground is related to REs. Data of both tiers ought to be normalised in terms of rates per a set number of flights (e.g. accident rate per million flights).

The Safety Management International Collaboration Group (SM ICG) (2014) proposed a safety performance measurement system that focusses on the aviation safety system as a whole. This system comprises indicators distributed over three tiers, namely, safety outcome indicators (tier 1), service providers' performance indicators (tier 2) and regulator performance indicators (tier 3) (Figure 4).

	A: Indicator Usage	**B: Outcome Indicators**	**C: Process Indicators**	**D: Inter-tier Correlations**
Tier 1 Integrated civil aviation system	Public information/ long term trending, identification of significant risk areas	(1) Accident rates, Incident rates, Fatalities (etc.) (2) Breakdown of event rates for significant risk areas	∑ Safety Management capability (effectiveness of): • Identifying common cause hazards • Effectiveness of regulatory risk controls	N/A
Tier 2 Service provider performance	Risk mitigation by service provider and regulator "most wanted issues" (SMS/SSP)	Per category of service provider: • Outcome related to significant risk areas	SMS performance: • SRM/compliance with regulatory specifications • Ability to identify unique cause threats • Effectiveness of risk control actions	Influence of service provider activities on safety outcomes
Tier 3 Regulator performance (activities)	Safety Risk Management by regulator (SSP)	• Activities and initiatives to adress specific risk areas • Effectiveness of risk controls (correlation with service provider behaviors and aggregate outcomes) • Effectiveness of risk control application (Oversight system performance - design assurance and performance assurance)	Safety risk management capability: • Ability to identify common cause threats • Ability to develop risk control • Ability to manage risk across organizational boundaries. Resource allocation: • Efficiency of safety risk control (cost/benefit)	Influence of regulator activities on service provider behaviors Influence of regulator activities on safety outcomes.

Figure 4. Safety performance measurement matrix.
Source: SM ICG. (2013). *Measuring Safety Performance Guidelines for Service Providers.*

In the second left column, the main usage for the respective tier is indicated. That is, the regulator performance indicators used for its SSP, service provider's performance indicators are used for risk mitigation and safety priorities (e.g. CFIT, LOC-I, runway incursions, etc.) and the integrated civil aviation system uses its performance indicators for public information, trends and significant risk areas identification. Performance indicators in the next two columns are, in addition, divided over outcome and process indicators. As the name suggests, outcome indicators look at the results of occurrences and the safety management processes. And process indicators look at the functioning of these safety management processes. A key aspect of this system is the interrelationships between all three tiers in which measures by indicators can be validated through correlation of the tier above (i.e. applicable to tier 2 and 3), as shown in the rightmost column. The first tier focusses on the event rates of high consequence occurrences (i.e. accidents and serious incidents), but also on the common causes that are precursors of these severe events. Measurement of tier 1 is therefore considered a reactive one. Subsequently, the second tier focusses on the assessment of safety risks to prevent the events from tier 1. It therefore measures activities or events other than high consequence occurrence which are related to the performance of SMSs and safety priorities. At last, tier 3 focusses on the safety risk management of the respective regulator. Indicators of this tier should measure the effectiveness of the regulator's actions to address the issues recognised through performance measurements of tier 1 and 2.

2.5. Safety event's controllability as performance indicator

Chapter 1 discussed the direction and objective of this study. In particular, it stressed the use of a new type of taxonomy based on the controllability by the involved person(s) of a developing safety event. This taxonomy will be used and exploited for multiple purposes throughout the course of this study. That is, verifying the performance and effect of the taxonomy on a larger scale by what has been achieved before by Karanikas (2015) and consequently attempting to realise the use of the taxonomy as a safety performance indicator in the aviation industry, but also creating priority criteria for the selection of incident investigations. Precisely because of these purposes, the literature review of the study performed by Karanikas must be comprehensively addressed and fully self-contained in its own right (i.e. without external perspectives); it is after all the mainspring of this study.

Karanikas (2015) noted that safety events' outcomes "dominate the contemporary accident rates computation in the scope of measuring safety performance" (p. 183). These events are considered and classified, by severity classification, without the implication of the events' potential and human intervention in attempt to control the developing safety event. It is for these observations that the study focussed on the development and introduction of a new safety performance measurement that implicates the human involvement and its effectiveness in the act to mitigate safety events before such severity outcome classification is considered. Specifically, the taxonomy is utilised at a large aviation organisation with aircraft accident reports from the years 2000 to 2010, resulting in a total of 808 reports analysed.

2.5.1. Working environment of the study

The organisation in question comprises of three different acting sections (i.e. Flight operations, support and operations, and flight training) that all manage various aircraft types. In addition, aircraft acquired before 1985 are considered second generation aircraft and aircraft acquired from 1985 are considered third generation aircraft. It should be noted that each section operates their aircraft in multiple bases (e.g. Flight operations operates in seven bases). This organisation used accident rates with respect to the severity classification as safety performance indicator and determined the contributing factors (e.g. crew acts) percentage prior to the respective study of Karanikas. Safety events are divided by means of a severity classification designated by A (catastrophic damage/injury) through D (minor damage/injury), or if an event does not fit the respective descriptions it is designated as an incident (See Table 1 on the right hand page). This classification scheme incorporates both the injury types and total costs for defining the severity classification of a safety event.

Table 1: Severity accident classification

Implications	Accident class			
	D	**C**	**B**	**A**
Total cost in Euros (Including accident damages on equipment, environment and third parties, as well the costs for accident response, accident investigation, medical care, etc.)	≤ 20.000,00€	≤ 20.000,00€ ≤ 500.000,00€	≤ 500.000,00€ ≤ 2.000.000,00€	≤ 2.000.000,00€ Or total damage of aircraft – infrastructure involved
Injuries	Minor injuries causing either the need for part time working, or limitation in the tasks assigned to, or change of tasks, or medical care beyond first aids provision	1. Minor injury causing absence for more than 1 day after the accident day. 2. Minor injury causing the need for permanent change of job specialty.	3. Major injury or permanent partial disability caused. 4. Hospital admission requirement for more than 3 employees.	Fatal injury or permanent total disability caused

Source: Karanikas, N. (2015). *An introduction of accidents' classification based on their outcome control.* Safety Science, Issue 72, pp. 182-189.

2.5.2. *The controllability taxonomy*

Karanikas furthermore noted for the development of controllability in a new taxonomy "that there is actually no symmetry between causes and consequences since the same error may have completely different effects in a different context" (p. 184). Hence the focus of this taxonomy lies within the control attempt or intervention of the end-user that contributed to the end-state of the occurrence. If an event developed without the involvement of an end-user, the end-state could be considered as being uncontrolled. He concluded that such events indicate areas that may require special attention. These considerations ultimately resulted in the new classification scheme as shown in Table 2 on the right hand page.

The resulting classification scheme comprises three control classifications (hereinafter referred as "accident [control] classes") as shown above in the first column: controlled, uncontrolled and neutral. If an occurrence implicated the attempt of the end-user to control the outcome, it is defined as a "controlled" event. In contrast, if no attempt has been made to control the outcome of the occurrence by the end-user, it is defined as an "uncontrolled" event. With respect to the controlled state of an occurrence as shown in the second column, it could either be positively controlled (i.e. the outcome was not worsened by the end-user) or negatively controlled (i.e. the outcome was worsened by the end-user). Certain consequences of controlled events may be the result of standard procedures or standard reaction by the end-user. The outcome of such an event is to be expected from the arising situation and is hence defined as a "neutral" event.

2.5.3. *Analysis of accident data*

The study focussed on unveiling the various distributions and relations by means of frequency analyses and Chi-square tests respectively. The frequency analysis looks at the distributions of the accident control classes as depicted in the first column of table 2. This distribution functions as the "safety performance indicator of accidents' progress control attempt". In addition, a frequency analysis is performed for the distribution of the positive and negative control attempts of the controlled accident class to function "as an indicator of human intervention effectiveness". And finally, the third frequency analysis incorporates the distribution of the severity classification of table 1 of only the controlled accident class. This is because safety performance of an organisation is mainly focussed on controlled events in order to demonstrate

Table 2. Published accident classification according to control attempt and effectiveness of user intervention.

Accident control classification	User reaction classification	Examples
Controlled: The user attempted to control the accident march	Positive: User's actions did not worsen the outcome; the accident outcome was managed successfully; no errors or violations were noticed during the control attempt	• Aviation: Safe landing after engine flame out • Ground transportation: Successful fire extinguishing following vehicle's malfunction
	Negative: User's actions following the safety event initiation resulted in adverse outcomes due to human errors or violations	• Aviation: Incorrect technique to recover the aircraft from unstable state • Marine industry: Ineffective ship manoeuvring to avoid adverse sea conditions
Uncontrolled: Safety event's consequences were developed without control; there had been no intervention until the time the outcomes were noticed.	None	• Aviation: Important impacts on the engine compressor blades due to Foreign Object Damages observed during After Flight Inspection • Chemical industry: Blood quality problems due to exposure to hazardous chemicals were identified after periodical medical checks
Neutral: Inevitable application of normal procedures; standard reactions to identified problem	As expected by prescribed procedures	• Aviation: The tire blew out and the pilot stopped the aircraft • Ground transportation: The driver stopped the vehicle due to engine oil light illumination

Source: Karanikas, N. (2015). *An introduction of accidents' classification based on their outcome control.* Safety Science, Issue 72, pp. 182-189.

the performance of safety. Considered in the foregoing paragraphs, the contributing factors of safety events are analysed, through Chi-square tests, against the accident control classes to reveal differences in the contributing factors. Further analysis has been performed to reveal any relations with the frequency analyses' considered variables. These variables are known as the dependent variables and consist of the accident control classes (i.e. controlled, uncontrolled or neutral), distribution of user attempt in controlled events (i.e. positive or negative) and event severity distribution of the controlled accident class (i.e. A, B, C or D). These dependent variables are analysed against the independent variables, which are categorised as temporal and systemic or local related in this particular research. Temporal variables include the date of the event (i.e. day, month and year) and system or local related variables include the aircraft type and generation, the accountable base, base size and the section related to the base.

2.5.4. Results and conclusions from the study

By following the same order as denoted in the methodology in the former paragraph, the results of the frequency analysis shall be discussed first. The distribution of the three accident classes of the new taxonomy was 43.3% of controlled accidents, 43.7% of uncontrolled accidents and 13% of neutral accidents. While leaving out the neutral accident class it can be concluded that the controlled and uncontrolled accident classes are evenly distributed. Karanikas furthermore stated that the outcomes of accidents "are half depended on 'luck' and half caused by its employees' attempts to control the situations on the accident scene" (p. 188) and therewith indicating that areas comprising 'luck' (i.e. uncontrolled accidents) require attention to address the uncontrollability. Elaborating on the controlled accident class, 87.2% resulted in a positively alleviated end-state of the respective accident, which indicates the effective intervention capabilities of the respective organisation's employees. The analysis of the severity classification in regards to the same controlled accident class resulted in 4.3% class A, 1.4% class B, 64.1% class C and 30.2% class D. When considering the distribution of severity classes without the controlled accident class, it is noted that class C decreased by 10% and class D increased by 9%. This observation indicated that the organisation's current severity classification distribution computation may lead to misleading conclusions as the C-class and D-class accidents respectively weigh in lower and higher without considering the controllability. Subsequent Chi-square analysis demonstrated that the contributing accident factors are significantly different according to the accident control classes and mostly affected the uncontrolled accident class. There are multiple variables that affected the controllability attempt by the involved person(s); these are the aircraft types and generations, the accountable base and base size and the related section. Notable mentions of these results are the generation of aircraft, in which second generation aircraft are more involved in controlled accidents, and the base size, in which larger sized bases (i.e. with more than 500 employees) were more involved in uncontrolled accidents. The effectivity of intervention of controlled accidents is most notably demonstrated by the aircraft generations, accountable base and the base size. Second generation aircraft were more effective in controlling the accident with a higher rate of positive outcomes. Also, larger sized bases presented a higher rate of positive outcomes.

The results of the study indicated that the new classification can demonstrate whether the outcomes of accidents are dependent on luck or the organisation's employees attempt to alleviate the outcome severity. It may also demonstrate the human effectiveness in attempt to alleviate such events and its contributing (affecting) factors. Therefore, areas that are uncontrolled, thus indicating luck, need additional attention to address the controllability. Karanikas concluded that this new classification could be used before the severity of accidents is considered to illustrate a "more realistic method" for measuring an organisation's safety performance. Safety performance could then be measured with the distribution of controlled and uncontrolled accident classes and the effectiveness of controlled accidents by human intervention. Targets of safety could subsequently be set to increase the proportion of controlled accidents and positive outcomes of the controlled accidents. This classification can indicate whether areas are controlled, including its effectiveness, or uncontrolled and can thereby identify where attention is needed. As a result, organisations "may drive their initiatives and efforts in a more effective way under the resource constraints inevitably present in every business" (p. 188).

2.6. Concluding remarks

Considering the review literature hereinbefore, it seems that safety is often expressed in terms of outcome based data of safety events. Accident and incident rates dominate safety statistics at both the individual as aggregate state level. Nevertheless, more proactive measures are implemented to identify hazards, through risk management, before they develop into more severe accidents. If such accidents do occur, its severity drives the allocation of resources of the assigned safety investigation authority. On the whole, selecting whether or not an investigation should be conducted depends on the statutory requirements. However, selection may be based on elements such as social and public impact, the severity of the occurrence in terms of fatality count, or the physical size of the aircraft. Investigative resources are subsequently allocated according the aforementioned selection criteria. As accidents and serious incidents are required to be supported by an inquiry, priority of investigative resources tends to converge towards these types of occurrences. Though less severe and more frequent incidents are often considered to provide better preventative measures of severe accidents, yet are not prioritised as such. In addition, Karanikas' study centralised the human controllability as a performance indicator and stated that it could be utilised as a more realistic method of measuring an organisation's safety performance before the current industry severity classification is considered.

Next steps

This chapter discussed the current practices of safety investigations, safety performance measurement and monitoring and the means provided of this study to address severity based prioritisation of investigations and outcome-oriented safety performance measurement. This taxonomy shall be tested and analysed for its functionality, effectivity and ability to be used as a new performance indicator for a (seemingly) better representation of safety performance, but more importantly, it will be exploited to base priority criteria on the human controllability in safety events before instituting an investigation. The next chapter will therefore outline the methodology to utilise the taxonomy on safety investigation reports for the acquisition of representative data.

3

Chapter 3

Methodology

Contents

3.1. Research sample

3.2. Framework development

3.3. Quality assurance

3.4. Data analysis methods

3.5. Summary

Methodology

“Perfection is not attainable, but if we chase perfection we can catch excellence.”

- Vince Lombardi, 1913 - 1970.

From the preceding chapter, it became clear that the outcomes of safety events drive the allocation of resources for conducting safety investigations. This outcome-centric perspective is also most often utilised to determine a state’s or service provider’s safety performance. Taking this into consideration, this chapter will describe the methods used for the application of the new controllability taxonomy as introduced by Karanikas (2015), which will be exploited to evaluate its effectiveness as a new safety performance indicator and to generate investigation priority criteria in attempt to provide an alternate approach to today’s investigation selection practices. The methodology performed in this study is, additionally, explorative in nature as hypotheses determined several associated factors that may unveil possible relations affecting the controllability and resource allocation in order to test the taxonomy and outcome bias respectively. Chapter 3 therefore focusses on exploring mainly the main research question: What factors are associated with the controllability of safety events in aviation? On the whole, a sound framework has been created that included controllability, outcome-bias and associated factors which were used to record data from safety investigation reports.

Before any reports were analysed, a substantiated determination of the research sample is expected (Section 3.1). The complete research sample ought to be analysed on its associated factors, outcome-bias and controllability. A framework is therefore developed and described in Section 3.2. As occurrence outcome severity and the controllability taxonomy may be subjectively interpreted, its application consistency and reliability has been tested to ensure consistent results (Section 3.3). The analysis methods for obtaining statistical results of the analysis are subsequently discussed in Section 3.4. In the last section (3.5), the methods employed are summarised in preparation for the presentation of its results in the following chapter.

3.1. Research sample

Multiple methods may be considered appropriate for obtaining the answer to the main research question; however, the preceding research of Karanikas (2015) was the main drive of this particular study. In that regard, using similar research methods was paramount for the relatability of both studies. Instead of seeking alternate methods for answering the research question(s), the study focussed on the analysis of safety investigation reports.

As the main and only source for this study is safety investigation reports, the investigation authorities that publically publish these reports are the only sources considered for selection. The European Commission (2010a) regulates that safety investigation reports of accidents and serious incidents should be made public within the shortest time possible. This does not entail, however, that an absence of publically available incident reports exists, quite the contrary. As a result, all three types of occurrences are assessed in the analysis to identify bias towards the outcome and/or its severity these occurrences.

In order to obtain a varied perspective on how safety investigations are conducted and reported, five safety investigation authorities were selected based on the authority's locality, nationality, scope and the availability of safety investigation reports. This variety among safety investigation authorities was considered to be an important factor to ensure that differences in investigation practices (e.g. accident outcome bias, extent of investigations), if present, were presented in the results. In addition, the department at which the study is performed has conducted multiple studies that employed similar selection criteria for analysing safety investigation reports. As such, this particular study analysed the complete research sample of 297 investigation reports to enable the department the interlinking of data among the various studies. These reports were selected on obtaining an even distribution along a time period from 1990 till 2014 as to reveal differences over time. The research sample of the five safety investigation authorities, including the distribution of safety events' classification, is presented in Table 3.

Table 3. Details of selected investigation authorities and research sample as distributed by accident class and safety investigation authorities.

	Safety investigation authorities					
	AAIB UK	ATSB	DSB	TSBC	NTSB	**Total**
Country	United Kingdom	Australia	The Netherlands	Canada	United States of America	N.A.[1]
Continent	Europe	Australia	Europe	North America	North America	N.A.
Research sample						
Accidents	39	18	23	44	72	196
Serious incidents	14	7	34	5	0	60
Incidents	7	14	11	1	1	41
Total	60	39	65	60	73	297

[1] N.A. stands for Not Applicable
Note. AAIB UK is Air Accident Investigation Branch UK; ATSB is Air Transport Safety Board (Australia); DSB is Dutch Safety Board; TSBC is Transportation Safety Board Canada; NTSB is National Transportation Safety Board (USA).

3.2. Framework development

As discussed in the previous chapter, the preceding research applied the controllability taxonomy on a large sample of safety investigation reports. As the main purpose of this study is to evaluate this taxonomy, its application was naturally replicated to demonstrate similar results. Nonetheless, this application is thought of being more diverse both in scope and aim of use that subsequently requires a sound, well-argued framework. This framework's sole intend was to facilitate the storage of data in which the associated factors to the controllability of safety events and bias towards outcomes these events were explored. The explorative nature of the study gave the researcher absolute freedom of choice to what end these factors were recorded. However, the more variables considered the less safety investigation reports could be analysed (due to the limited time scheme). Only variables that either contributed to the research objectives or were found significant in the literature were explored.

Firstly, it was considered to record ordinary data (e.g. temporal data, location of occurrence, flight characteristics) as discussed in Section 2.3.2 that is standard to safety investigation reports (Section 3.2.1). Another essential part of the study aims to explore potential safety investigation authority's bias towards adverse outcomes of safety events, which was determined to be recorded by the length of its associated investigation report (Section 3.2.2). As may be expected, the effectiveness of human action in response to the initiation of safety events is critical in the controllability taxonomy. Such information was considered to be recorded by job experience to those involved in the event. Human effectiveness' impairment, in contrast, was considered to be recorded through fatigue and its associated factors (e.g. time on duty) (Section 3.2.3).

In short, the independent variables considered for the analysis of the controllability taxonomy and outcome-bias are decided with the only intend to explore what factors might be associated with the controllability taxonomy. All concerned factors incorporated in the analysis are listed below and thereafter discussed:

- Location of occurrence
- Origin of operator
- Temporal factors
- Aircraft specifics
- Flight characteristics
- Occurrence details
- Human (experience) factors
- Fatigue
- Safety investigation reports details

3.2.1. Ordinary data

Location and origin

As global accident rates in aviation continue to decline, differences between the regional categorisation become more apparent (ICAO, 2015a; EASA, 2014b). The regions Africa and Pan America, specifically, score the highest accident rates as shown in Table 4.

As noted in the table, ICAO (2015a) employs the Regional Aviation Safety Groups (RASGs) categorisation for defining the global regions. It comprises of five regions: Africa (AFI), Asia Pacific (APAC), Europe (EUR), Middle East (MID) and Pan America (PA). These regions differ slightly from the definition of continents and are therefore recorded alongside the determination of the respective continent. Apart from the type of regions, its application is used for both the location of the occurrence and origin of the operator (i.e. nationality of controller). In the event that a safety event occurs above international waters, it was recorded as such.

Table 4. Accident statistics in aviation of 2014 as classified per RASG region.

Region	Estimated departures	Accidents	Accident rate	Fatal accidents	Fatalities
Africa	0,7	6	**8,6**	1	118
Asia Pacific	10,2	18	**1,8**	3	449
Europe	8,9	26	**2,9**	1	298
Middle East	3,0	7	**2,3**	2	39
Pan America	9,9	41	**4,1**	0	0
World	33	98	**3,0**	7	904

Source: ICAO. (2015a). *ICAO Safety Report 2015 Edition.* Montréal(Québec): International Civil Aviation Organization.
[1] Estimated departures expressed in millions
[2] Accident rate is expressed per one million departures

Effectiveness of communication between personnel is paramount in the conduct of a flight. Cultural differences, as indicated through country or region of origin, are often cited in degraded communicational effectiveness in the cockpit (Engle, 2000; Hutchins, et al., 2002). It is therefore considered a potential factor in the effectiveness of control in the analysis.

Temporal Factors

As discussed in section 2.4.1, safety performance indicators essentially use safety data to identify trends. In order to test the controllability taxonomy as a safety performance indicator, identification of variations over time is considered a mandatory factor. As such, the year, season and month in which an event occurs is recorded in the analysis. The time at which such an event occurs is recorded as well in order to reveal whether the time of day or lighting conditions changes human effectiveness.

Aircraft and flight characteristics

Karanikas (2015) revealed that the generation (i.e. derived from age) and type of aircraft seemingly affected the controllability attempt and effectivity in controlled attempts. Airbus (2015) additionally states that an aircraft's generation[1] affects the accident rate significantly, in which older generations have significantly higher accident rates. Hence, the age of aircraft was obtained as the specific age at the time of the accident (i.e. the age in years between date of construction and date of occurrence). The distribution between types of aircraft was established through the differences in propulsion type: jet, propeller and rotorcraft aircraft (EASA, 2015a).

---

[1] Third and fourth (current) generation aircraft are manufactured from 1980 and 1988 respectively. Source: Airbus, 2015. *Commercial Aviation Accidents 1958-2014 A statistical Analysis.* Blagnac: Airbus S.A.S..

In order to reveal outcome-bias, the physical size of an aircraft may contribute to an event's level of attention (ICAO, 2015b; Sarsfield, et al., 2000). ICAO's (2013d) Accident/Incident Data Reporting (ADREP) taxonomy, which categorises gathered accident data, is employed for multiple variables in the framework in order to obtain consistency in classifications. The physical size of the aircraft is incorporated by the standard mass classification of aircraft as presented below:

- 0 - 2250 kg
- 2.251 - 5.700 kg
- 5.701 - 27.000 kg
- 27.001 - 270.000 kg
- >270.000 kg

Aside from the type of aircraft, the nature of operation was considered different from each other. These differences in operation are considered to change factors such as skill and experience levels of pilots, workload and procedures (FAA, 2016). Flight phases are often discussed as being critical in aviation safety, since almost 90% of all aviation accidents occur during the take-off and climb phases and descent to landing phases (Airbus, 2015). Again, the ADREP taxonomy (categorisation) is utilised for the type of operation and flight phases for the recording of variables in the framework (ICAO, 2013e).

Length of safety investigation reports

In order state the presence of outcome-bias towards the severity of a safety event, the length of the associated investigation reports was considered. The length was determined to be extracted from the word count of the respective reports. As discussed in Section 2.3, safety investigation reports generally follow a certain format of five distinctive sections. The word count from all of these sections is separately recorded in addition to the total word count of the reports. In specific to issued safety recommendations, the number of recommendations is relative to the nature of the events and ultimately contributes to the improvement of safety in aviation. From issued safety recommendations that were addressed to EASA in the year 2014, it seemed that certain safety investigation authorities are more reluctant than others to issue these recommendations. For example, the British AAIB had addressed thirteen safety recommendations to EASA, while the Dutch safety board addressed five recommendations and the Canadian transportation safety board only one (EASA, 2015b). Safety recommendations contribute to the improvements in safety by the publishing of safety issues at hand. It is for this fact that the quantity of recommendations a safety investigation authority publishes is thought to resemble an authority's urgency (e.g. due to severity or human error) to improve or rectify the safety issues. Safety recommendations were hence incorporated in the analysis.

3.2.2. Occurrence and severity classification

The regulatory classification of safety events (See section 2.1) was considered essential in the conduct of the analysis and research as a whole. Namely in revealing the relation between the severity of occurrences with respect to the extent of its associated safety investigation reports. This classification is therefore recorded in the analysis. However, this classification is solely based on the outcome and its severity of a safety event. A categorisation of occurrences, regardless of its classification, was therefore needed to extrapolate the effects of occurrences' severity and reveal whether the outcome (i.e. by the classification) or the type of outcome (i.e. by the categorisation) affected the results. Certain types of occurrences are considered critical as their quantity far exceed the other remaining types. Both Airbus (2015) and ICAO (2015a) define the "high risk accident occurrence categories" in which runway safety events, loss of control-inflight (LOC-I) and controlled into flight terrain (CFIT) combined represented more than half of all accidents in 2014. Once again, ICAO's ADREP taxonomy (categorisation) is employed to record the variables for the occurrence categories in the framework (ICAO, 2013f).

It has been hypothesised that more severe occurrences gain more attention (see section 1.1), from sources such as media and safety authorities, causing investigative resources to be allocated accordingly (Sarsfield, et al., 2000; ICAO, 2005; Greenwell, 2003). Needless to say, analysing the attention towards an event's severity leaves the study with only one classification, namely the statutory occurrence classification (i.e. accidents, serious incidents or incidents, see Section 2.1.2). In addition, this type of classification does not gradually incorporate the level of severity of an event. For example, events that involved 100 fatalities are considered accidents. However, when an aircraft only sustained substantial damage, it will be classified as an accident nonetheless. Differences among these two examples can therefore not be retrieved with just the current classification. In order to establish whether more severe occurrences gain more attention, a severity classification is a mandatory element.

In the study conducted by Karanikas (2015), a severity classification has been used to distinguish safety events in terms of its implicated costs (which naturally incorporate accumulated damages to aircraft, equipment and environment, etc.) and injuries to personnel (See table 1 on page 32). This classification separates four levels of severity as denoted by four descriptors (i.e. A, B, C and D). ICAO (2013b) provides a similar severity classification that is aimed for a service provider's management of risk. Nevertheless, it was regarded as being practical for the establishment of a sound severity classification in this particular study. ICAO's severity classification has six factors that may be considered for the description of severity. These are the safety of aircraft, physical injury, potential revenue loss (i.e. quantified in currency), asset damage (i.e. quantified in currency), environmental damage and corporate reputation damage. It also comprises of five distinctive severity levels, namely insignificant, minor, moderate, major and catastrophic. Taking into account the nature of this study and the availability of such information in investigation reports, potential revenue loss and corporate reputation damage has been disregarded for the establishment of the severity classification. Asset damage is defined by a service provider for a predetermined figure. However, there are five levels of damages in addition to the losses expressed in currency, namely no damage, minor, substantial, major and catastrophic damage. The ultimate severity classification employed in defining the level of severity is shown on the right in Table 5.

Table 5. Severity classification of aviation safety events

Severity level	Descriptor	Safety of aircraft	Damage level	Injury level
A	Catastrophic	Aircraft/hull loss	**Catastrophic damage**	Multiple fatalities
B	Major	Complete failure of significant/major aircraft systems or results in emergency application of flight operations procedures	**Major damage**	Single fatality
C	Moderate	Partial loss of significant/major aircraft systems or results in abnormal flight operations procedure application	**Substantial damage**	Serious injury
D	Minor	Degrades or affects normal aircraft operational procedures or performance	**Minor damage**	Minor injury
E	Insignificant	No significance to aircraft-related operational safety	**No damage**	No injury

Note. This classification is adapted from "*Doc 9859, Safety Management Manual (SMM)*. Third Edition." by ICAO, 2013b.

Even though table 3 comprises of all necessary data for analysing safety reports, certain definitions must be defined to prevent misconceptions in the assignment of severity. A serious injury is defined by ICAO (2010, p. 22) as "an injury which is sustained by a person in an accident and which:

a. requires hospitalization for more than 48 hours, commencing within seven days from the date the injury was received; or
b. results in a fracture of any bone (except simple fractures of fingers, toes or nose); or
c. involves lacerations which cause severe haemorrhage, nerve, muscle or tendon damage; or
d. involves injury to any internal organ; or
e. involves second or third degree burns, or any burns affecting more than 5 per cent of the body surface; or
f. Involves verified exposure to infectious substances or injurious radiation."

In this regard, all injuries not complying with the definition of serious injury must be considered minor. When considering aircraft damage levels, catastrophic damage involves (hull) loss of aircraft which is determined by whether the aircraft's state is beyond economical repair and/or completely destroyed. ICAO unfavourably does not define major damage, but is however used in safety risk management (ICAO, 2013b). Substantial damage is defined by damages or structural failures that "adversely affected the structural strength, performance or flight characteristics of the aircraft and would normally require major repair or replacement of the affected component" (ICAO, 2013g, p. 2; ICAO, 2010, p. 20). With respect to this definition, damages less severe than substantial damage is considered minor. Damages that fall outside the confinement of substantial damage (i.e. substantial damage not affecting structural strength, performance or flight characteristics), but are not considered as hull loss were used as major damage[2].

3.2.3. Human performance and fatigue

It has been discussed that human performance is most often concerned in the causation of, or contribution to, (aviation) safety events (See Section 1.1). A person's effectiveness of its control attempt is therefore sought in the performance

---

[2] With hindsight, it should be noted that safety investigation reports (often) defined an aircraft's damage severity and, as a result, did not degrade the purpose of this classification.

of corrective actions after an occurrence's initiation. These actions are, however, commonly impaired by fatigue as it is often a causing factor of safety events (Wise, et al., 2010). The FAA (2007) cites fatigue as "a condition characterized by increased discomfort with lessened capacity for work, reduced efficiency of accomplishment, loss of power or capacity to respond to stimulation, and is usually accompanied by a feeling of weariness and tiredness" (p. 2). It hence decreases a person's ability to perform tasks due to reduced level of alertness, attention and/or concentration (Wise, et al., 2010). In addition, a fatigued person is considered to be more prone to make mistakes at the least favourable moments, which is contributed by sleep deprivation, shift work and long work days (ECA, 2012). The European Transport Safety Council (ETSC) (2013) states that fatigue is contributed by a person's circadian rhythm, sleep and workload. Predictable changes in temperature and light, that are caused by the rotation of the earth, can be found in daily human behaviour; known as the circadian rhythm. This circadian rhythm indicates the (human) body when it is time to sleep, such as the setting of the sun and drop in body temperature. However, obvious cues such as changes in daylight do not regulate the pattern of the circadian rhythm, but it is the biological clock of humans that predict and control the circadian rhythm (Wickens, 2009). Considering the sleep component, the quality and amount of sleep, time awake and sleep inertia are essential to the level of fatigue. Finally, the workload encompasses the total duty time and the (work-) load related to a specific task (ETSC, 2013).

In a study to demonstrate the personal perceived fatigue of pilots in both short and long haul flights, Bourgeois-Bougrine et al. (2003) showed that fatigue is indeed related to sleep deprivation, workload and disturbances in circadian rhythm. Flight duty times affected the perceived level of fatigue in which the fatigue increases with the increase in flight duty time. In addition, a pilot's age seemed to affect his or her ability to resist fatigue as identified by his or her actions, communication in foreign languages and interruptions during activities, in which an age of 34 years or younger demonstrated a better resistance to fatigue than ages of 35 years and up. Consequences of the perceived fatigue included the reduced levels of attention and social communication, but also making minor mistakes (e.g. math and interpretation). A similar study has been conducted by the European Cockpit Association (ECA) (2012) that drew almost identical conclusions. Disruptive work schedules, night and long (flight) duties were considered key components that caused fatigue in the cockpit.

By no means was fatigue considered as only or main factor contributing the human performance in the cockpit. Nevertheless, scientific literature and statistics agree upon the fact that fatigue is a serious issue in aviation. It is for these facts that fatigue is recorded as variable in the analysis in order to test whether fatigue does indeed affect human behaviour and therewith the controllability in safety events. Derived from the literature above, signs of fatigue are extracted from the respective person's age, rest period before duty, amount of sufficient sleep before duty (i.e. last sleep period), time on duty till time of occurrence, but also from the time of day. Daytime was expressed in the four designators: Afternoon (i.e. 12:00-17:59), Evening (i.e. 18:00-23:59), Night (i.e. 00:00-05:59) and Morning (i.e. 06:00-11:59). As the definition of fatigue is widely used in aviation as contributing factor in accidents, it was found that fatigue is indeed often cited in safety investigation reports whether or not it affected the pilot's performance in the event. Therefore, when fatigue was considered as contributing factor it was recorded, as a separate variable, by "YES" in the framework. If fatigue was disregarded as contributing factor or not even considered as probable factor in the reports, it was denoted by "NO".

Apart from human performance impairment due to fatigue, Bazargan & Guzhva (2011) revealed that pilot experience (in general aviation) does impact the involvement in the causation of accidents by pilot error, but also the likely involvement in fatal accidents as a whole. It seemed that more experienced pilots are less involved in accident causation, but more likely involved in fatal accidents. Pilot experience, or controller experience in general, was therefore considered in the analysis to reveal the effect it may have on his or her performance. This was recorded through pilot, or flight engineer (FE), flight hour experience for type rating and total flight hours in which each "controller" is recorded in order of function. In practice, most often the composition of functions was captain and first officer as controller 1 and controller 2 respectively. The third and fourth controllers were only recorded with the presence of a second officer and/or flight engineer.

3.3. Quality assurance

In order to test the validity and reliability of the actual analysis, four raters (with similar background and specialisation to each other and the researcher) have been requested to take part in an inter-rater reliability test. By means of this test, the consistency of outcomes among the five raters[3] was obtained for the individual application of the

---

[3] The researcher took part in the assessment

severity classification and controllability taxonomy on four randomly selected safety investigation reports from the research sample. All of the remaining recordable data of the actual analysis were not included, as this was considered to be (significantly) less influenced by rater bias or interpretation. In a scheduled group session of about 60 minutes, all raters have been briefed (by the researcher) on the definition and use of both classifications. A concise presentation ended with the elaboration of several sample cases per type of classification and controllability. Hereafter, the raters were requested to perform the analysis without mutual consultation in pursuit of absolute individual judgement.

The validity and reliability among the raters were obtained by means of the Intraclass Correlation Coefficient (ICC). Values obtained through the ICC measurement range from 0, indicating absolute disagreement, to 1, indicating absolute agreement (Shrout & Fleiss, 1979; Stemler & Tsai, 2008). The test's measurement did not only record the controllability and severity, but also the contrast between the outcomes of the human intervention for controlled events (i.e. positive and negative outcomes). As such, the severity classification scored an ICC value of 0.749, the controllability 0.361 and human intervention outcome 0.933. On one hand, the severity classification indicated a fair agreement and thus deemed acceptable for the research. On the other hand, the controllability scored unacceptably low while the human intervention scored an almost perfect agreement. When considering the analysis data of the raters, it seemed that the interpretation between "Neutral" and "Positively controlled" classifications may have affected the rater agreement. It should be noted that this affected Karanikas' (2015) research as well. In addition, variability was found between "Positively controlled" and "Negatively controlled" and between "Uncontrolled" and "Negatively controlled". In hindsight, the low quantity of analysed reports was considered a degrading factor whereby no conclusions would be justified and no hypotheses could be confidently stated. After subsequent interviews with the raters it seemed that their judgement may have been affected by another factor: their unfamiliarity with analysing investigation reports.

Following the results of the first test, a subsequent one focussed on addressing the detected issues. This included the hypothesis of the overlapping classification of neutral and positively controlled safety events, but also the additional attention to the analysis methods unfamiliar to the raters. Since the very essence of the test is to measure the reliability among the raters' judgement of the classifications and not the raters' reading skills, certain adjustments to the execution of the test were made. Only the information deemed necessary to apply the taxonomy is used in the test to focus on the actual judgement of the raters. In this test the raters were requested to analyse twenty randomly selected investigation reports[4] from the research sample whilst only recording the controllability taxonomy. In this instance, the controllability scored an ICC value of 0.917 and the human intervention outcome 0.845, which indicated a substantial agreement and deemed satisfactory to assure reliable and consistent application of the taxonomy. However, the test confirmed that the interpretation of 'Neutral" and "Positively controlled" classifications varies as this affected the agreement in more than 80% of the cases that contained a disagreement among the raters.

Overlapping interpretation

Both the 'Positively controlled' and 'Neutral' accident control classifications are considered as control actions in attempt to alleviate adverse outcomes after the initiation of such an event. However, 'Neutral' control is restricted by the extent of control, regardless of an event's controller-influenced outcome. 'Neutral' is hence applicable if the control attempt is confined within the definition of the "inevitable application of normal procedures; standard reactions to [the] identified problem" (Karanikas, 2015). Merriam-Webster's (2016) online dictionary defines reactions as "the ability to act and move quickly in order to avoid sudden danger". Moreover, the term "inevitable application" is thought of the differentiating definition of 'Controlled' and 'Neutral' attempts, which will help improve the consistency in designating these classifications. Taking this in consideration, the following criteria are defined for the 'Neutral' control classification.

A control attempt is indicated as "Neutral" when:

- The controller's action is considered a *reaction*; as identified through the sudden (expected) action in order to avoid the perceived problem;
- The controller's action is considered an *inevitable application of normal procedures*; as identified by the taken action(s) of the controller that are normal procedure and, due to the controller's acquired training, experience or common knowledge, (beforehand) expected to be performed.

---

[4] Preliminary analysis ensured sufficient cases with both neutral and positively controlled designations were included.

All controller applications of other types of procedures (e.g. non-normal or emergency procedures) were not appointed as ‘Neutral’, regardless of its reactionary nature or inevitable application. Actions that could not be appropriately assigned to as the inevitable application of normal procedures and were not considered a standard reaction, were subsequently designated as “Controlled”. That is, application of normal procedures did not rule out the possible “Controlled” classification.

3.4. Data analysis methods

It can be derived from the literature hereinbefore that the focus of the analysis is to test the association between the dependent and independent variables. That is, are the independent variables associated with the outcomes of the dependent variables? The complete list of the recorded and analysed dependent and independent variables is shown in Appendix I. The dependent variables accident control class and control attempt outcome were first separately analysed by means of a frequency analysis. This analysis was utilised for the accident control classes as a safety performance indicator of the control attempts in safety events. Similarly, the contrast between the outcomes of control attempt was exploited to indicate the human effectiveness of the implicated control attempt.

In order to evaluate the differences among the dependent and independent variables, the Chi-square test was employed. Known as the “test of independence”, two categorical variables can be tested if differences are due to chance or due to one of the variables that are being tested (Berenson, et al., 2012). To state this dependence, the significance level α (for all statistical tests) was set to 0.05. All independent variables were analysed with respect to the dependent variables, including the extent of safety investigation reports. However, only the potential independent variables that could plausibly affect the extent of safety investigation reports were the only variables analysed for this dependent variable (See Appendix I). In situations where the requirements of the Chi-square test were not met (i.e. small sample sizes), the Fisher’s Exact test was employed to correctly compute the significance of the respective data (Fisher, 1922).

The recorded data consisted of both numerical and nominal data, in which the numerical data was tested for normality with the Kolmogorov-Smirnov Test. In particular, the test was used to determine the nature of the data: parametric (i.e. normal distribution) or nonparametric (i.e. distribution-free) (Massey, 1951). As such, it was found that none of the variables demonstrated a normal distribution[5].

The classification of data from the independent variables was determined on the median for numerical data[6] and equal frequency distribution for nominal data. In order to enable the execution of Chi-square tests, besides the distribution of amount of words to indicate the extent of safety investigation reports, groups of nominal values were made for each numerical data set based on its distributed nature of frequencies and/or median. All nominal data, including the extent of safety investigation reports, were classified according an even frequency distribution among two to seven classes, depending on the distribution of frequencies and nature of the data (see Appendix II for the classifications of independent variables). Data with a recorded frequency less than five were not incorporated in the analysis to ensure validity of the statistics.

Another analysis was utilised to evaluate the differences between numerical variables (i.e. between an independent variable and the length of safety investigation reports [7]). Spearman's rho (Spearman, 1904) was employed to measure the tendency between (nonparametric) numerical variables when increasing or decreasing. This correlation indicates the influential tendency of one variable to increase or decrease when the other increases or decreases. The Spearman's rho values range from -1 to 1, in which a value of zero indicates absolute independence (i.e. no tendency to change). Values 1 and -1 indicate perfect correspondence and perfect inversed correspondence respectively. Finally, certain statistical results and nature of corresponding variables were found interesting to explore by setting layers in the Chi-square tests. The relation between variables could indicate statistical significance, however, the association of these variables could be affected by another independent variable. In this regard, the significant results were further explored for the effects of a third independent variable, namely the controlling variable. On the whole, all statistical results were computed with statistical software *SPSS Statistics* version 22.

---

[5] Significance for all variables were calculated less than 0.05 ($P<0.05$).

[6] All numerical data was found distribution-free, i.e. nonparametric.

[7] Note that the terms dependent and independent variables are explicitly used for the variables of the study, not the statistical analysis.

3.5. Summary

The main purpose of Chapter 3 was to describe, but more importantly, substantiate the methods chosen for the analysis of safety investigation reports. It was in particular aimed at providing the information for answering the main research question; finding the associated factors to the controllability in safety events. As the controllability taxonomy was re-evaluated for this study, with respect to Karanikas' (2015) research, equivalent methods were used. Previous studies at the research department with similar research data and criteria contributed to the selected research sample of 297 safety investigation reports. Factors that could affect the controllability in safety events or severity based prioritisation (i.e. outcome bias) of investigations respectively, were thoroughly explored in this chapter. However, before these factors were recorded in the analysis, an inter-rater reliability test was employed for the controllability taxonomy and severity classification to obtain whether the collection of data is deemed consistent and reliable. It was ultimately concluded that the application of both the controllability taxonomy and severity classification was found consistent among the raters, however, a significant overlapping of two definitions in the controllability taxonomy required a sound demarcation. Methods used for the analysis of statistical data consisted of a frequency analysis, Chi-square tests and Spearman's Rho analyses. Goodness of fit tests were also performed to indicate the distribution type of numerical data that was utilised for classification purposes of the specific data set. All tests were performed with the significance level set at 0.05.