SYSTEMIC SAFETY INVESTIGATIONS FOR AEROSPACE MRO'S

Ph.M. van Meer, Dr. A.A. Ghobbar, Prof. Dr. J.A. Stoop
Delft University of Technology, Faculty of Aerospace Engineering

Keywords: *Safety, Human Factors, Complexity, Classification*

Abstract

Whenever something goes wrong, one of the most important things to do is draw lessons from the event to prevent something similar from happening again. To this end, aerospace Maintenance Repair and Overhaul (MRO) organizations gather data on all operational glitches that have, or could have compromised operational safety.

The problem many organizations face is the fact that in spite of substantial amounts of data, the amount of safety-occurrences is not being reduced. This means that either the data does not contain all necessary information, or the information is not extracted from the data. Classification systems such as HFACS-ME and MEDA are frequently used to provide structure to safety data. These frameworks, based on the epidemiological model of accident causation are designed to record the event-chains as well as factors contributing to these events. However, they are not specifically designed to capture the higher level systemic factors which have been suggested in more recent accident causation theory. This view suggests that accidents, incidents or unsafe behavior do not occur as an abnormality, but rather as a consequence of normal human behavior within the system. Therefore, understanding these high level systemic factors will provide a means of increasing resilience against unsafe events.

This paper presents the results of a case study performed at the maintenance division of a major European airline. The research objective was to determine how the abovementioned concepts from modern accident causation theory can contribute to the safety performance of this particular MRO. In more detail, the goal was to maximize the learning potential from available operational data, considering the practical constraints on available resources. Furthermore, the existing operational structure for gathering and classifying data was considered a given for this research; recommendations should not conflict with the present situation, but rather be complementary.

To achieve the objectives the MEDA-based data was subjected to several statistical analysis methods specific to dealing with categorical, nominal variables, to extract trend information and isolate potential risk areas. To validate the results from statistical analysis, frequent sanity checks were performed in cooperation with high and low level safety management staff.
A specific case was selected from the found high-risk areas and was subsequently investigated using three systemic investigation methods available from recent literature. The methods were rated on aspects such as ease of use, investigative depth and definition of intervention strategies.

1 Introduction

Today's large aerospace maintenance organizations face minor incidents on a daily basis. Human interaction with a complex technological system that is under constant economic pressure is more than susceptible to erroneous actions that have or might have influenced safety [1]. Reporting of this type of occurrences is an important part of the Safety Management Systems (SMS), which are being

proclaimed by regulatory authorities, safety experts and industry leaders to be the future of safety management in aviation [2], [3], [4]. Many of the reported occurrences however, will not be found to have posed sufficient risk to justify the allocation of resources associated with a thorough investigation [4]. These incidents however can provide crucial insight into the capability of the system to adapt to variable process in- and outputs [5].

Epidemiological accident causation models, on which many of the occurrence taxonomies being used in the industry today are based, have recently received considerable criticism for their oversimplification of system functioning and their lack of having a systems approach [6], [7], [8]. Epidemiological models are considered to be capable of explaining what happened and how it happened in a clear and structured manner, yet the answer to why things happened, lacks. The epidemiological model and its derivate tools such as MEDA and HFACS-ME however hold a dominant place in both the minds of employees as well as in the organizational structure of many aerospace companies. As managing change in large organizations is a science in its own right, an attempt to introduce new notions and concepts in accident causation would be better of being evolutionary, rather than revolutionary.

This paper will present the proposed theoretical approach to determine how a complex socio-technical system can improve intervention strategies based on safety occurrence reports. To validate this approach, a comparative case study of three distinct views on the systemic model of accident causation will be described. The three approaches will then be evaluated using quantitative criteria.

2 Systemic approach

When an occurrence can be related to an erroneous human action, this is often referred to as ‘human error’. The ‘systemic view’ on human error considers occurrences ‘part of the process’, thus also regards human error as a normality rather than an abnormality. Dekker describes this as the ‘new view’ on human error: human error is considered a symptom of deeper problems rather than being the cause of an occurrence as was in the ‘old view’. The true learning potential of an occurrence thus does not lie in determining in what way people have acted unreliably, but lies in understanding why people did what they did [9].
Primary focus should therefore not be on determining how accidents occur i.e. how people make mistakes, but on how people function within and cope with the system.

Systemic models aim to describe the performance of the system as a whole, rather than of cause-effect mechanisms or epidemiological factors. Rather than structurally decomposing the system, the systemic approach favors a functional decomposition. This functional decomposition is essential to express the notion that accidents are considered emergent properties of the system, i.e. elements of a decomposed system can create synergistic properties for the whole of the system, which are invisible when focusing on the elements. Only when the system is viewed as a whole, will these properties emerge [10]. Safety is thus considered a naturally occurring result of the processes and functions of the system and is susceptible to change; it is a characteristic of how well the system, its functions and processes perform [11].

Going even further, Hollnagel argues that system safety is in fact increased by human adaptability and cognition, as people are capable of overcoming system design flaws and functional glitches and by their understanding of the system can detect when things threaten to go wrong [7]. In this context people are considered to make decisions on a daily basis on how to manage or adapt to variation, so that the output of their process remains within bounds. This decision is in essence finding the balance in trading off efficiency and thoroughness (Efficiency-thoroughness trade-off or ETTO), where efficiency means achieving the desired result with a minimum of resources, while thoroughness relates to achieving the desired goals regardless of resource limits [12]. This approach stands in shrill contrast with safety

management approaches that are capable of designating for example 'complacency' a 'root cause' to an occurrence.
Improving a system such that it (and anyone in the system) is capable of coping with adverse variability, without these setbacks resulting in accidents is the core concept of the developing science of 'resilience engineering'.

Westrum defines resilience as one of three characteristics of a system [13]:

- resilience is the ability to prevent something bad from happening,
- or the ability to prevent something bad from becoming worse,
- or the ability to recover from something bad once it has happened.

So if resilience of a system is the capability to adapt to variable process in- and outputs, then an occurrence marks the boundaries of the adaptiveness built into the system's design.
Occurrences therefore can provide valuable information about where the margins of resilience are eroding [5]. The detailed registration of occurrences and their contributing factors, combined with the scale of operations at large MRO's provides sufficient basis for statistical trend analysis. Statistical methods, such as correspondence analysis, statistically combine variables of the individual occurrences, sketching highlights of the system's safety boundaries. These points of interest can thus provide an interesting starting point for a new, systemic investigation to ascertain how the system has reached these boundaries.

This investigation should not only provide a description of the physical reality, but also incorporate the socio-technical context and operating environment [14]. An expression of these aspects is Stoop's DCP-diagram, (depicted in figure 2). Using this diagram as reference, the investigation should encompass design and control of the system as well as how it functions in practice on all possible levels of aggregation [14].

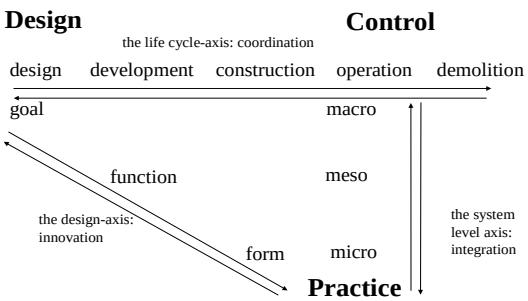

Figure 1. DCP-diagram [14]

Introducing the systemic investigation as an addition to the available safety management tools available in many aerospace organizations, might provide an answer to some of the aforementioned criticism on epidemiological analysis.
Based on the hierarchical representation of accident levels suggested by Leveson [6], the analytic cycle proposed in this paper is presented in figure 2.
The bottom-left side of the figure presents the frequently used analytic structure, where occurrences are analyzed and categorized according to the chain of events and contributing factors. Statistical software can be used to analyze the resulting data, in turn providing safety staff with information on safety trends.
Using these trends as the starting point of a systemic investigation might provide a better explanation of why these trends are occurring within the organization, in turn leading to potentially more effective intervention strategies.

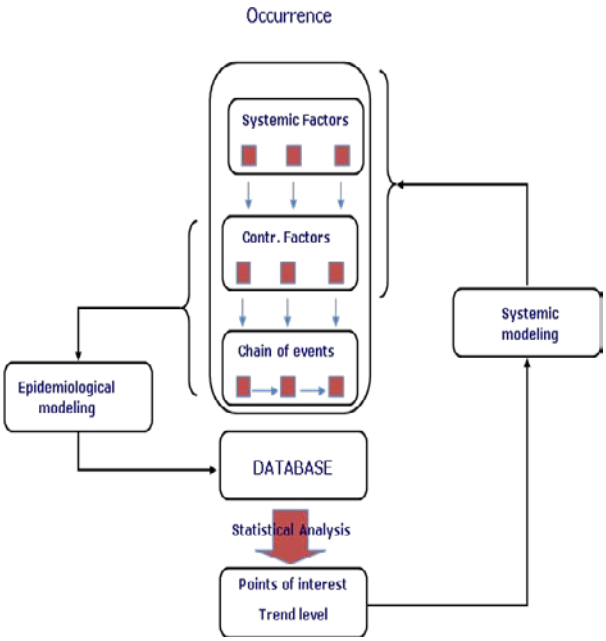

Figure 2. Proposed analytic cycle

The systemic approaches that have been included in this research are the Forensic analysis approach, the Systems Theoretic Accident Model and Processes (STAMP) and the Functional Resonance Accident Model (FRAM). All three these methods start with a functional decomposition of the system, yet the approach suggested to be taken thereafter differs.
Hollnagel's FRAM-approach is oriented towards identifying variability in functions and how interaction of this variability can (positively or negatively) affect the system [7].
The STAMP-model, developed by Leveson is strongly focused on systems and control theory: accidents are considered to be caused by a lack of constraints or control thereof [15].
The forensic engineering approach is based on an evidence-based analysis of the primary system functions. Furthermore, there is a focus on solution design spaces [14], [16].

3 Evaluation of the systemic approaches

To assess the different investigation methods, criteria from earlier comparative studies of investigation methodologies were used. A quantitative approach was favoured in order to facilitate determination of the most suitable method. Two comparative studies were used as basis for this assessment [17], [18].

The three-point scale used by Benner provides a clear view of whether the methodology satisfies, does not satisfy, or could satisfy the criterion after additional effort (2 points, 0 points and 1 point respectively). The qualitative criteria used by Sklet can easily be rated on this scale.

An overview of the criteria used by Benner and Sklet are given in tables 1 and 2 respectively.

Table 1. Quantitative evaluation criteria [17]

Criteria	Requirements
Encouragement	The methodology must encourage harmonious participation. Does the methodology promote harmony by encouraging parties to participate in investigations and have their views heard, minimize conflict by disclosing gaps in the investigation and efficiently but harmoniously control the presentation of individual views with appropriate technical disciplining techniques during the investigation?
Independence	The methodology must produce blameless output: Does the investigation methodology identify the full scope of the accident including the role of management, supervisors and employees in a way that explains the effects and the interdependence of these roles without imputing blame, fault or guilt?
Initiatives	The methodology must

	support personal initiatives. Does the methodology provide for positive descriptions of accidents that show convincingly what is needed to achieve adequate control of risks in a specific workplace, in a way that promotes informed and valid individual initiatives, without unnecessarily conveying blame, fault, or guilt?
Discovery	The methodology must support timely discovery process. Is the investigation methodology able to discover safety and health problems when applied to these problem areas? Does the methodology enable timely discovery or must discovery be delayed until credibility of sample sizes and causality requirements are met?
Competence	The methodology must increase employee competence. Does the investigation methodology provide direct inputs that will increase the competence and safety effectiveness of personnel through training in the detection, diagnosis, control and amelioration of risks? Are outputs resulting from the application of this investigative technology being used in training demonstrable safety effectiveness?
Standards	The methodology must show definitive corrections: does the methodology provide a timely comprehensive, credible and persuasive basis for establishing or reviewing efficacy of safety and health standards? Does it document accidents in a way that countermeasure options can systematically defined, evaluated and selected, avoiding personal opinions and judgments during multiple reviews?
Enforcement	The methodology must show expectations and behavioral norms. Does the investigation methodology support the required enforcement program by providing information perceptions of duties under a standard, its practicality, and its effects on risk levels by (a) defining the degree of compliance or nature of compliance problems and (b) showing the role of a standard in a specific accident in a way that objective observers can trust and rely on?
States	The methodology must encourage States to take responsibility. Does the investigation method encourage States to fulfill their occupational safety and health mandates by providing them practical ways to produce consistent, reliable accident reports, pretested for completeness, validity, and logic before they are submitted, this multiplying the effectiveness of their contributions.
Accuracy	The methodology must

	help the test accuracy of outputs. Does the methodology describe each accident in a way that can be technically "truth-tested" for completeness, validity, logic and relevance during the investigation, to assure the quality of the information in each case?
Closed loop	The methodology must be compatible with "pre-investigations" (or safety analyses) of potential accidents. Is the methodology compatible with the pre-investigation or analysis methodologies so those predictions can be used during investigations, so expected vs. actual performance of tasks and controls can be measured or validated by investigations and so the results can be linked routinely to work flow design improvements?

Table 2. Qualitative evaluation criteria [18]

Criteria	Description	Evaluation option
Graphical method	Does the investigation methodology use a predetermined graphical description of the accident sequence?	Yes/no
Focus on safety barriers	Does the investigation methodology include an analysis of how safety barriers influenced the accident?	Yes/no
Level of scope	Is the investigation methodology capable of involving all levels of aggregation in the socio-technical system (work/technology, staff, management, company, regulators, government)?	Level 1 through 6
Primary	Can the investigation methodology be used as a stand-alone investigation (primary), or is it to be used as supplement to other methods (secondary)?	Yes/no
Education	What level of education does the investigation method require before it can be properly used?	Expert/ Specialist/Novice

Not all criteria mentioned by Benner and Sklet are relevant for the evaluation given in this paper. The criteria that are abandoned are briefly discussed.

Competence: the method's audience is primarily Airline Alpha's safety community. In a later stage, the investigation findings could be (re-)formulated to address a larger audience. A distinction is made between output of the investigation method and the input for follow-up actions (see also the 'initiatives' criterion). Therefore it is not considered a requirement that the investigation method directly affects the competence of personnel.
Enforcement: identifying enforcement and compliance problems is not considered a relevant criterion based on the current views of accident causation.
States: this criterion is driven by the federal background of the done research and is not considered relevant for this research.

Focus on safety barriers: Although an often used and important concept, safety barriers are not considered to be an essential criterion in assessing validity of an investigation method.
Primary: this criterion is not relevant as all three investigation methods are primary methods.

In contrast with the earlier comparative studies mentioned, this study is aimed at assessing the investigation methodologies with respect to analyzing a trend, i.e. a group of incidents. This study is thus different in scope and the criteria used for analysis of the methods should also compensate for this.

The first important criterion is how well the investigation methodology is capable of dealing with multiple related, but possibly different incidents or accidents. The optimal situation is where a methodology identifies and focuses on commonalities and patterns between the incidents, without complicating the investigation or having to shift the investigation to a higher level of aggregation. On the other end of the spectrum, it is conceivable that a method requires either registering all differences between the incidents (in which case there are in fact multiple individual accident investigations linked together instead of trend analysis) or the methodology only uses the common denominator of all incidents, in which case detail of the investigation might be compromised. This notion is referred to as the 'multiplicity' criterion.

Once a trend analysis has been done, the results should be able to be used to verify whether new occurrences fit within the scope of the investigation and whether or not the conclusions of the trend investigation are also applicable for the single case. This is similar to the 'closed loop' criterion, differing de facto only in the direction of the loop; where the closed loop is used to identify whether the method is compatible with 'pre-investigations', for the trend analysis it must also be compatible with 'post-investigations'. This criterion is the 'backward loop' criterion.

Quantifying the qualitative criteria is done as presented in table 3, if otherwise the rating is applied as discussed earlier.

Table 3. Quantifying qualitative criteria

Criterion	2 points	1 point	0 points
Graphical method	The methodology is strongly oriented to graphical methods	The methodology uses graphical methods, but is not strongly focused towards them	The method makes no use of graphical methods
Level of scope	The methodology is capable of encompassing all levels of the socio-technical system	The methodology can encompass all levels, if an additional effort is made to do so	The methodology is not capable of encompassing all levels of the socio-technical system
Education	The methodology can be used with little or no additional training	The methodology can only be used after basic training in its use	The methodology can only be used after significant training and practice in its use

The criteria were rated after review of theory, and after application of the approach to a case study. The case study was performed at the maintenance division of a major European airline and involved an observed trend consisting of approximately 50 incidents from the period 2008-2009.

4 Results

The results of the comparison can be seen in table 4 and are graphically represented in the radar plot of figure 3.

Table 4. Rating results

	Forensic analysis	STAMP	FRAM
Encouragement	2	2	2
Independence	2	2	2
Initiatives	2	1	0
Discovery	1	2	2
Standards	2	1	1
Accuracy	2	2	2
Closed loop	2	1	1
Graphical method	2	2	2
Level of scope	2	2	1
Education	2	1	1
Multiplicity	2	1	1
Backward loop	2	1	1
Total score	23	18	16

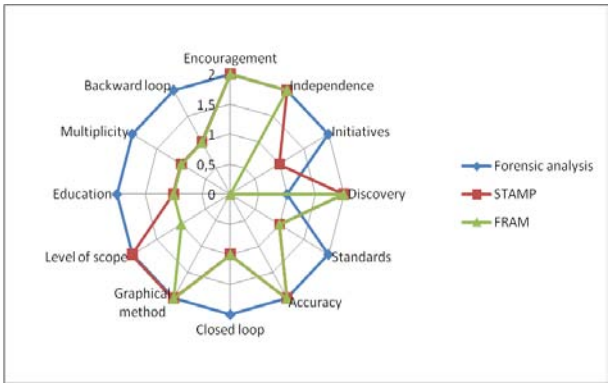

Figure 3. Radar plot of the resulting scores

From the results, it can be concluded that the forensic analysis performs best for the trend analysis. Especially for the criteria ´Multiplicity´ and ´Backward loop´, the criteria essential to trend analysis, FRAM and STAMP are outperformed by the forensic analysis approach.
The main weak point of the forensic analysis method is found to be the 'discovery' criterion. Any hypothesis is required to be backed up with sufficient factual data. This would incur that the method is strongest when applied in hindsight, i.e. when all events have already unfolded. Although it is not impossible to derive predictive findings from the application of forensic analysis, the structure of the method is not optimized to do so.
The dynamic systems model and control diagram as described in the STAMP methods were found to be particularly effective in giving insights in the functioning of the system without having to be proven based on gathered evidence; these representations were found to have predictive qualities that are lacking in the forensic analysis method.
It is therefore proposed to complement the forensic analysis method with a control diagram and a system dynamics model as proposed in STAMP.
The FRAM approach proved effective in visualizing functional interactions, however, as complexity of the interactions increased, the model became cumbersome to work with. Furthermore, due to FRAM's rather abstract nature, the solution design space also remains on an abstract level.

Referring back to the DCP-diagram in figure 1, it seems whereas the forensic analysis approach attempts to cover all three axes, STAMP has a primary focus on the vertical axis (control structure), while FRAM stays focused on the diagonal axis (functional level).

4 Conclusion

The scientific world is starting to take a new approach to safety and how it is affected by human behavior. This does not mean that previous approaches are without merit. On the contrary, the epidemiological model and the tools based on it are proving very useful in structuring occurrence reports. Even though epidemiological taxonomies remain on a superficial level of events and circumstances, they can provide an outline of the organization's safety structure and where it is exceeding its boundaries. Complementing this with a systemic investigation that explains why people

have behaved in the way they did, or why the events were possible to unfold without anyone reacting adequately to stop them can create an understanding crucial to learning from occurrences.
It was shown that systemic investigations along preset yet flexible guidelines help determine why the socio-technical system no longer functions as intended and whether or not the system should be reinforced or redesigned.

References

[1] Reason, J. and Hobbs, A. *Managing maintenance error.* Ashgate Publishing Ltd., 2003.
[2] European Commission. Directive 2003/42/EC, Official Journal of the EC, 2003.
[3] Stolzer et al. *Safety management systems in aviation.* Ashgate Publishing Ltd., 2008.
[4] Hudson, P. Safety reporting in aviation: safety management and safety culture in interaction, 2010, Yet to be published.
[5] Woods, D.D. and Cook, R.I. Incidents - Markers of resilience or brittleness?, *Resilience engineering: concepts and precepts.* Ashgate Publishing Ltd., 2006, pp. 70-76.
[6] Leveson, N.G. *Systems safety wngineering: back to the future*, Massachusetts Institute of Technology - Aeronautics and Astronautics, 2002.
[7] Hollnagel, E. *Barriers and accident prevention,* Ashgate Publishing Ltd., 2004.
[8] Dekker, S. *The field guide to understanding human error,* Ashgate Publishing Ltd., 2006.
[9] Dekker, S. *Just culture: balancing safety and accountability,* Ashgate Publishing Ltd., 2007.
[10] Hollnagel et al. *Resilience engineering: concepts and precepts,* Ashgate Publishing Ltd., 2006,
[11] Hollnagel, E. and Woods, D.D. Resilience engineering precepts, *Resilience engineering: concepts and precepts* , Ashgate Publishing Ltd., 2006, pp. 347-358.
[12] Hollnagel, E. *The ETTO-principle: why things that go right sometimes go wrong,* Ashgate Publishing Ltd., 2009.
[13] Westrum, R. A. Typology of Resilience Situations, *Resilience engineering: concepts and precepts,* Ashgate Publishing Ltd., 2006, pp. 55-65.
[14] Stoop, J.A. *Reader forensic engineering, AE4231,* Delft University Press, 2009.
[15] Leveson et al. *A systems theoretic approach to safety engineering,* Massachusetts Institute of Technology, Aeronautics and Astronautics Dept., 2003.
[16] ESReDA. *Guidelines for safety investigations of accidents,* ESReDA, 2009.
[17] Benner, L. Rating accident models and investigation methodologies, *Journal of safety research*, pp. 105-126, 1985
[18] Sklet, S. Comparison of some selected methods for accident investigation. *Journal of hazardous materials*, pp. 29-37, 2004

Contact Author Email Address

A.A.Ghobbar@TUDelft.nl

Copyright Statement