The Wrong Question to the Right People. A Critical View of Severity Classification Methods in ATM Experimental Projects

Alberto Pasquini[1], Simone Pozzi[1,2], and Luca Save[1,3]

[1] Deep Blue srl, Rome, Italy
[2] Sapienza University of Rome, Department of Psychology of Social and Developmental Processes, Rome, Italy
[3] University of Siena, Media and Communication Department, Siena, Italy
{alberto.pasquini,simone.pozzi,luca.save}@dblue.it

Abstract. The knowledge of operational experts plays a fundamental role in performing safety assessments in safety critical organizations. The complexity and socio-technical nature of such systems produce hazardous situations which require a thorough understanding of concrete operational scenarios and cannot be anticipated by simply analyzing single failures of specific functions. This paper addresses some limitations regarding state-of-the-art safety assessment techniques, with special reference to the use of severity classes associated to specific outcomes (e.g. accident, incident, no safety effect, etc.). Such classes tend to assume a linear link between single hazards considered in isolation and specified consequences for safety, thus neglecting the intrinsic complexity of the systems under analysis and reducing the opportunities for an effective involvement of operational experts. An alternative approach is proposed to overcome these limitations, by allowing operational people to prioritize the severity of hazards observed in concrete operational scenarios and by involving them in the definition of the possible means of mitigation.

1 Introduction

Every time a new system is introduced or an existing system is significantly modified, a safety assessment must be performed to identify if potential new risks are introduced as a result of the innovation. Safety assessments, especially in complex socio-technical domains such as Air Traffic Management (ATM), always require some kind of involvement of people with operational experience (e.g. controllers and pilots) whose knowledge is deemed essential for an adequate understanding and evaluation of risk. However most of the standard safety assessment techniques adopt as a central strategy the use of a safety matrix, aimed at classifying each hazard in terms of its expected severity and acceptable frequency. This method is generally intended to rely on *expert judgment* for an appropriate evaluation of the severity of hazards, but operational experts tend to experience difficulties when working at such a task. The assessment of severity is normally based on so-called *severity classification schemes* which identify a set of *severity classes*. Each class is associated to a different severity level and to a specific outcome (e.g. accident, incident, no safety effect, etc.).

M.D. Harrison and M.-A. Sujan (Eds.): SAFECOMP 2008, LNCS 5219, pp. 387–400, 2008.

In this paper we elaborate on our experience with severity classification schemes in the ATM domain to discuss the reasons why operational experts cannot easily use these schemes: (i) Hazards are typically identified as failures to a single function of the system, without considering the potential interactions of such function with other parts of the system, (ii) The severity of each hazard is assessed by considering its potential "final" effect, assuming only a linear chain of events and infringed barriers, thus neglecting the non linear dynamics of incidental scenarios.

2 Accident Models and Limits of Probabilistic Risk Assessment

In recent years a number of theoretical contributions have investigated the complex nature of accidents in socio-technical and safety critical systems like nuclear power plants, chemical industry and transportation systems. These contributions pointed out the limits of accident models based on linear sequences of events and cause-consequence configurations. The seminal studies of Charles Perrow [1] revealed that accidents can be seen as due too unexpected combinations or aggregations of events, named *complex interactions*. More recently Reason's Swiss Cheese Model [2, 3] has been considered successful in representing accidents as the result of combined failures at different levels in an organization, including unsafe acts by front-line operators and latent conditions such as weakened barriers [4, 5] and defences. Finally, other models like FRAM (Functional Resonance Accident Model) [6] or STAMP (System-Theoretic Accident Model and Processes) [4] highlighted the emergent nature of failures, which are often the result of dysfunctional interactions between different parts of the system, rather than simple malfunctions of specific components.

Compared to these theoretical advancements, there has been no comparable development in state of the art risk assessment techniques. Most of these techniques are based on a PRA approach (probabilistic risk assessment), i.e. they adopt as a central concept the well known definition *Risk=Severity x Frequency*. In such definition both the *severity* and the *frequency* are referred to the potential negative effects of the hazards which can be experienced by a certain system. Thus, in a typical safety assessment, hazards are defined as failures of one or more functions to be mitigated by reducing the frequency of their occurrence and/or the severity of their effects. The overall *level of risk* achieved by the system is the result of the aggregation of the risks identified for each specific hazard.

While this approach is theoretically appropriate for close or simple systems essentially made of hardware components, the application to complex socio-technical systems is problematic, as it relies on a linear representation of hazardous events which is inconsistent with the complex accident models mentioned above.

2.1 Assumed Linear Link Between Hazards and Their Effects

No matter which is the specific graphical notation adopted, PRA typically relies on models of accidents and incidents based on linear chains of events, representing the notion that the preceding event or condition must be present for the subsequent event to occur, i.e. if event X had not occurred than the following event Y would not have occurred [4]. These event-based models make it difficult to incorporate non-linear relationships (e.g. feedback between system components). A linear link between an

identified hazard for a specific function (e.g. a technical failure or a human error) and a *final effect* of the hazard itself (e.g. a minor incident, or a severe accident) oversimplifies the relationship between a single component failure and its possible negative safety effects on a system level. I.e. it disregards the well-known notion that a failure to a single function can never be considered as the sole cause of a negative effect for the safety of a system. As the *final effect* is used as a criterion to assess the severity of a specific hazard, this considerably influences the final results of the assessment. As argued by Leveson [4], this approach -which was appropriate in process industry design (e.g. nuclear power plants)- is largely insufficient in other kind of systems in which emergent configurations of different kind or resources (humans, mechanical, procedural) are essential elements of both the correct and unsafe functioning of an organization [6].

2.2 Initiating Events in the Chain Assumed to Be Mutually Exclusive

A well known limitation of event based models (e.g. Fault Tree Analysis) is that *basic events* are usually assumed to be mutually exclusive. While this assumption simplifies the mathematics in a PRA, it may not match the reality. Leveson explained how seemingly independent failures may have common systemic causes that result in coincident failures [4]. For instance in the Bophal accident, what might have appeared as an unlikely coincidence of failures was engendered by common design and management decisions.

This methodological limitation of PRA is strictly related to the one mentioned before. Assuming the *basic events* as mutually exclusive in the determination of an accident considerably simplifies the task of modelling cause-effect configurations, thus making simpler the numerical definition of the risk associated to each failure at component level. However it can hide critical interactions between different functions or components, which are essential for identifying the appropriate mitigation means.

2.3 Functional Failures and Dysfunctional Interactions

Traditional PRAs focus on *functional failures*, i.e. on the non-performance or inability of specific components to perform their intended functions. However, the more complex safety critical systems have become, the more accidents have been determined by *dysfunctional interactions* [4]. Dysfunctional interactions happen when system elements perform as it is expected (i.e. as specified by requirements) but still the overall system behavior results to be unsafe. Accidents happen not only because a pilot deviates from a specified procedure or because a hardware component does not perform as in its specifications. Accidents may be engendered by a critical interaction among different components (electromechanical, digital, human). If the safety assessment is exclusively focussed on functions and component failures, very little insight is produced in order to mitigate the hazardous situations deriving from dysfunctional interactions.

3 Safety Assessment Methodology in Air Traffic Management

A prerequisite for performing a safety assessment based on a PRA approach is that of identifying a relationship between a set of identified failures for each specific function and a set of possible consequences. In the Air Traffic Management world, this is

typically accomplished by filling in Functional Hazards Assessment (FHA) tables and by elaborating them with cause-effects propagation models, such as Fault Tree Analysis (FTA) and Event Tree Analysis (ETA).

3.1 The Assessment of Severity

SAM (Safety Assessment Methodology) [7] is the standard method for safety assessment in ATM promoted by EUROCONTROL. It is made up of three main phases: Functional Hazard Assessment (FHA), Preliminary System Safety (PSSA), System Safety Assessment (SSA) (see central column of Figure 1). The phases are in parallel with the lifecycle of the system under assessment (see left column in Figure 1). This paper is mainly focused on the first phase of the SAM, i.e. the FHA.

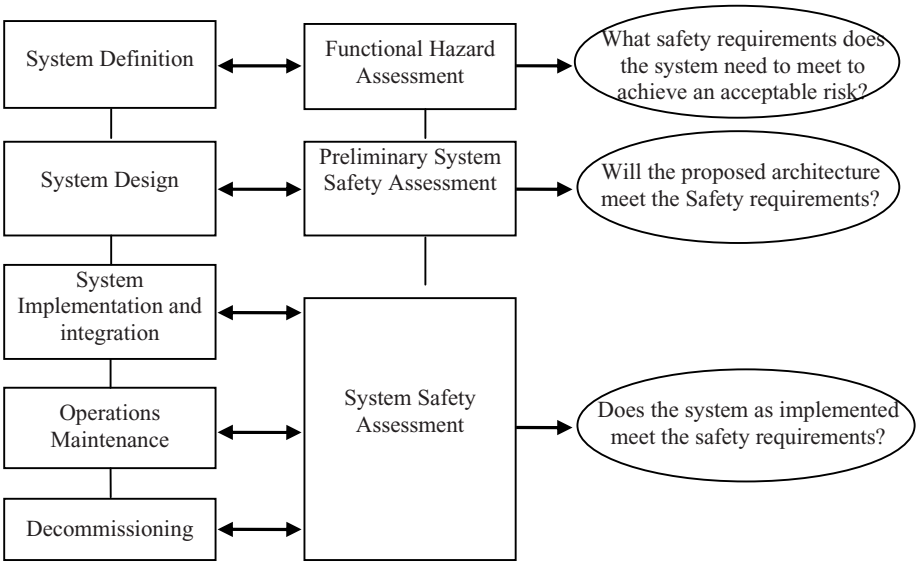


Fig. 1. The SAM Methodology

The main goal of an FHA is specifying a set of safety objectives. These are defined by following five sub-phases:

1. Identify all potential hazards associated with the system
2. Identify hazard effects on operations, including the effect on aircraft operations
3. Assess the severity of each hazard effect
4. Specify Safety Objectives, i.e. determine the maximum frequency of a hazards' occurrence
5. Assess the overall foreseen risk associated with introducing the change or new system.

Operational experts involvement is quite easily achieved in the first step (identify hazards), while the second phase (identify hazard effects) is more difficult and the

third phase (assess severity) can become extremely challenging. Operational experts (typically air traffic controller and/or pilots) are supposed to identify, in collaboration with technical and safety experts, the effect of each hazard identified in phase 1. Effects are then included in textual format in a specific column of the FHA table. Subsequently the experts are required to classify each of these effects in terms of severity, by using the SAM Severity Classification Scheme. The scheme identifies 5 different Severity Classes (SC), from the most severe to the least sever:

- SC1 Accidents [most severe]
- SC2 Serious Incidents
- SC3 Major incidents
- SC4 Significant incidents
- SC5 No Immediate Effect on Safety [least severe]

In principle the same hazard can have more than one effect, based on contextual conditions. A typical example is the differentiation of the effects of the same hazard, based on traffic conditions (e.g. low vs high traffic) or weather conditions. However the FHA table should identify a specific SC for each effect, without any particular attention if two effects are produced by the same source hazard. It is to be noted that the SCs are the same adopted in the EUROCONTROL requirements on safety occurrence reporting (ESARR 2 [8]), i.e. they are used by national service providers to classify real occurrences experienced in operational air traffic control centres.

3.2 Problems with the Use of Severity Classes

According to our experience (see case studies in section 4), the severity classification scheme is not easily applied in the Safety Assessment. While SCs are fit for purpose when reporting and classifying real occurrences, they are very difficult to use when assessing the safety of "pre-operational" systems. The most problematic aspect is the assumed linear link between a specific hazard and its possible effects. A specific failure – be it a technical failure or a human error – can never be considered as the sole cause of an accident. For an accident to occur, a hazard very often combines with several other hazards and contextual conditions. However, when adopting the functional approach which is typical of PRA, hazards generally does correspond to specific failures. We deal with single failures that could at the same time cause an accident (SC1), different kinds of incidents (SC2, SC3 and SC4) or even no immediate effect on safety (SC5). This is a commonly well recognised point, as demonstrated by the emphasis safety management systems place on near-miss events collection and analysis [9]. A near-miss usually shares the same causal factors with real incidents, where mostly contextual (sometimes even fortuitous) factors determine the different outcomes (i.e. no or very limited damage in near miss). Deriving consequences from each single hazard considered in isolation neglects the above reasoning on incident dynamics. In theory, any hazard can result in a serious incident or in an accident, depending on the way it interacts with other system weaknesses.

Operational experts are normally able to provide very detailed accounts on critical situations and can give valuable insights on the possible consequences of failures or dysfunctional interactions in concrete operational contexts. However, when faced with the task of classifying a single failure in terms of the 5 SCs, they manifest

uncertainties, expressed with sentences like "it depends on…it depends how…". If forced to make a choice, they generally tend to produce classifications that reflect certain assumptions about the contextual situation, or they provide a rationale justifying their answer. Neither assumption nor the rationale will be considered in the following of the assessment. The resulting classification will instead directly influence the setting of safety objectives and the definition of safety requirements.

Another spontaneous strategy is that of ranking the severity with respect to other hazards. In so doing, experts do not take into account the SC labels (accident, serious incident, etc), but rather reason on a priority ordering. The SAM guidance does acknowledge that the same hazard can have different effects and then different severities. It is specified that SCs should be assigned to the hazard effects, rather than to the hazard itself. The most commonly used method consists of identifying - based on expert judgment - the worst credible effect of each hazard, then in setting safety objectives taking into account only that effect[1].

The worst credible effect in the given environment of operation should determine the severity class leading to setting of the Safety Objective, using expert judgement. It means that somehow the probability of the hazard leading to certain effect (Pe) has been taken into account when deciding the worst credible severity of the hazard effect [10]. Even if the severity classification should not be influenced by considerations on the acceptable frequency, according to this quote one could claim that the decision on what is the worst credible effect is instead linked to considerations on the actual hazard frequency[2]. There is an implicit recommendation not to select a too much severe effect, unless its occurrence is not considered reasonably frequent. Ignoring such an implicit recommendation is likely to produce overambitious safety objectives.

4 Asking the Wrong Question to the Right People

The case studies presented in this section are both pertaining to safety assessment experiences in ATM related projects. The first case is the development of an FHA aimed at assessing the improvement of a Short Term Conflict Alert (STCA) in an European military ATC unit. The second case concerns the overall assessment process of an Airborne Separation Assurance System concept (ASAS), in the context of the European Program "Mediterranean Free Flight". These case studies provide evidence of some of the methodological limitations described in previous sections. They also document our attempts to overcome such limitations and propose an alternative approach.

Our approach is inspired by authors like Erik Hollnagel and Nancy Leveson and by their recent efforts to propose methods more in line with state of the art accident models (e.g. FRAM and STAMP). Our main strategy is that of extensively rely on the domain knowledge of expert operational personnel (e.g. controllers and pilots). The method we suggest exploits a scenario-based approach [11, 12]. The use of scenarios

[1] Note that the SAM guidance material proposes 4 different methods for setting safety objectives: Quantitative Method, Prescriptive Method, Criticality Method and Qualitative Method. For the sake of simplicity in this context we only refer to the last one.

[2] On the contrary, a rigorous application of the methods requires that the severity of hazard effects is assessed before. The acceptable frequencies are only established afterwards, based on a Severity x Frequency matrix.

is essential to place hazards and their possible consequences in concrete operational situations.

4.1 Case Study 1: Assessment of a New STCA for a Military Unit

Case study 1 concerns a safety assessment made in September 2006 for the introduction of an improved Short Term Conflict Alert (STCA) to be installed in a European military ATC unit. STCA is a system that assists the controller in maintaining separation between aircraft, by generating on the controller's display an alert of a potential infringement of standard separation minima. The military unit under analysis was already equipped with a modern ATC system including an STCA. However, the specific needs of the military environment (military formation flights, aerobatic maneuvers, etc.) created a large number of nuisance alert. The safety assessment was mainly focused on the safety impact of new technical solutions.

Essential part of the safety assessment was an FHA workshop, based on a number of brainstorming sessions attended by 10 people, including safety, technical and operational experts (i.e. military controllers and pilots). Main objectives were: (1) Identifying the most relevant hazards, (2) Understanding their effects on the ATM system, (3) Assessing the severity of their effects, (4) Identifying possible mitigation means.

The workshop profited from a scenario-based approach, consistently with what already made in the framework of other studies [12]. As an input to the hazard identification phase, a set of seven military related scenarios were identified, in collaboration with a controller and a technical expert. The scenarios were textual descriptions of typical operational situations representative of the military environment under analysis (an example is in Figure 2). Additional cells in the table provided information on the expected behavior of current STCA and on the technical solutions included in the improved STCA, in order to manage the specific situation.

The scenarios served to provide a description of the new system, from an operational point of view, to controllers who were not particularly familiar with STCA functioning. A second purpose was to support hazard identification brainstorming, by providing a concrete operational context. This purpose, in particular, was an attempt to integrate the functional approach, which requires starting from single functions of the system and thinking about their possible failures. The functional approach was actually maintained. However the scenarios complemented the functional perspective, as technical failures and human errors were imagined in concrete situations, allowing engineers and operational experts to derive also more complex hazards, like combination of different hazards or dysfunctional interactions.

The output of the brainstorming sessions was a list of 27 hazards, including a description of possible operational consequences and effects on safety, which were included in a typical FHA table [10, Appendix A, pp. A4-A5]. Example of hazards were: "Duplicate Mode A", "Lost Wingman", "Incorrect military formation detection", "Incorrect SSR code list input", "Controller not aware of STCA suppressed for specific aircraft", etc. In the FHA table, hazards were grouped to keep a reference to the scenario in which they were identified. According to the established method, the hazards identification phase was followed by the assessment of hazard severity and by the discussion about possible mitigation means. At the end all the results were included in the FHA table.

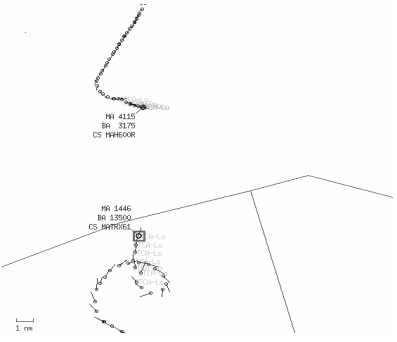
OS 2 - AREA TO AIRWAY	
Description	Traffic manoeuvring inside a military area next to a civil airway (ATS routes) with lateral or vertical manoeuvres.
Operational implications	Short reaction time for controllers to react if A/C penetrates civilian airspace. High speed manoeuvring, high ROC/ROD and steep turns versus steady flight profile. Aerobatics being performed both by singletons and by formation flights. Need for ATCOs to input BFL (Block Flight Levels).
STCA implications	Nuisance alerts are generated inside formations. Nuisance alerts due to excessive prediction times and high speed manoeuvring. BFL to be taken into account at the CWP Linear (any) prediction less accurate for the military traffic. If aerobatics are performed in formation, split tracks can occur.
	
Technical solution adopted in the new STCA	Creation of buffer zones around aerobatic areas using wider parameters as the Aircraft approaches the boundaries of the area. Use of BFL as in the current system. Dynamic activation/de-activation of STCA regions (improved FUA Level3).

Fig. 2. Example of a scenario template used during an FHA brainstorming session

4.1.1 The Decision to Give Up with Severity Classes

While in the hazard identification phase, the workshop attendees were very active in generating ideas and in providing descriptions of the possible consequences of the hazards, much more difficulties were experienced when the experts when confronted with the Severity Classification Scheme. First of all, it turned out to be difficult to identify the specific effect on safety of each hazard. Then experts stated that none of the hazards would have been the sole cause of an accident, but nearly all of them could potentially play a role in determining an accident. In addition, the categories *serious incident*, *major accidents*, *significant incidents* or *no immediate effect on safety* were considered difficult or impossible to apply. Even the safety indicators provided in the scheme (e.g. Effects on air navigation services, Exposure and Recovery) were not considered helpful, as the associated descriptions of possible hazards effects are obviously expressed in general and abstract terms: e.g "partial inability to provide or maintain a safe service" or "hazard may persist for a substantial period of

time". For example defining what is a "substantial period of time" will totally depend on subjective evaluations of the specific operational circumstances experienced and will not necessary imply the risk of producing a serious accident.

The limited time available for the workshop (one day and half in total) and the feeling of being stuck with hazard classification resulted in a spontaneous solution directly proposed by some of the attendees. While both operational and technical experts were not able to classify hazards in terms of the SCs, they had no difficulties in distinguishing between *high severity* and *low severity* hazards. Furthermore they remarked the importance of establishing a priority between hazards with an immediate need for a mitigation and hazards that could have been analyzed later. In other words, they proposed to shortlist a number of candidates for the following phases. A further distinction was made between hazards the mitigation of which was considered easier and hazards requiring further study. Even though this solution could appear not rigorous in methodological terms, it highlights the strong link perceived between hazards safety effects and the phase of designing mitigation actions.

4.2 Case Study 2: Assessment of ASAS Spacing Concepts in MFF

MFF was a large project of six years duration recently concluded, sponsored by the European Union under the TenT Programme. MFF was co-ordinated by ENAV - the Italian Air Traffic Control service provider - and involved several air traffic service providers, especially from the Mediterranean area, and EUROCONTROL. The scope of MFF was to define, test and validate operational concepts and procedures for more efficient use of airspace through the delegation of some tasks related to separation assurance, relying on concepts like Free Routes, ASAS Spacing & Separation, Free Flight. It focused on the application of those procedures in the particular geographical context of the Mediterranean area. The new operational concepts and related procedures were defined in the early phases of the project [13] and their fitness-for-purpose was evaluated through a set of validation exercises, with an iterative process of concept refinement and validation. This included several cycles of Model Based Simulations, three sets of Real Time Simulations (RTS), three Safety Cases, and an extensive set of Flight Trials (FT). Cockpit simulations were used in support of both RTS and FT.

The research issue we faced in this project was mainly due to the experimental nature of procedures and applications to be assessed. The introduction of the ASAS procedures profoundly changes parts of the existing ATM system, including changes in hazardous conditions and safety issues. Given the novelty of ASAS applications, there was no previous experience of them, nor any existing system with similar characteristics. The safety assessment process was then developed to face two complementary constraints.

1. Controllers needed to be familiarized with the new procedures and applications, so that they could contribute to the safety assessment as experts.
2. No experience was available on the system behavior, so a variety of simulation exercises was set up, in order to identify potential hazardous conditions hard to anticipate in the design phase. These simulation exercises could not replicate the complexity of a real system, but still some system elements could be put in place and observed while working together.

The integration of Real Time Simulations with the safety assessment process seemed a sound solution for both of the above problems [for more information on the MFF safety assessment process and on the use of safety scenarios in RTS, please refer to 11, 12]). The key aspect of such integration was the injection of a limited sample of hazards in the simulation through the implementation of safety scenarios. The major difficulty was to reconstruct a realistic situation where the procedure and the related hazard could be analyzed from a systemic point of view, preserving all contextual factors that shape controller's behavior. Safety scenarios were then used to avoid the assessment of hazards in isolation, so that credible situations could be presented to controllers. The safety scenarios included events such as system failures, pilots and controller errors, and other operational problems (see Table 1 below for an example of scenario story board).

Table 1. Story board and actions for a safety scenarios

Hazard Identification Code: SA2		**Airspace Sector: EW**
Time		**Events**
9.48	Accomplice Pilot Action	AZA123 asks to descend to FL290 for technical reasons
9.50	If Accomplice Pilot Action	IBE3674 and AFR432 are cleared to self-separate while crossing
9.50	then Possible Event	AZA123 interfere with IBE3674 and AFR432 (self-separation on-going)

They provided at least two immediate benefits. First, they gave experts an opportunity to reason about what did not work when the system failed, thus supporting the safety analysts in clarifying some aspects of the hazards. Second the safety analysts had the opportunity to learn through the direct observation of the controller behaviour during the exercises and to obtain information directly on a series of dedicated events.

However, if we get back to the main line of reasoning of this paper, what was observed during the RTS could not be considered satisfactory as far as the severity assessment was concerned. Although the RTS context allowed making post-hoc observations of the events and not just guessing the possible effects of hazards, using the 5 severity classes resulted to be problematic. Firstly the most severe one (SC1) - corresponding to an accident- is simply not simulated in the RTS environment. The closest the simulation can get to an accident is when two aircraft pass one through the other, which in the simulated world results in no damage to any of the two. The two aircraft simply keep flying on their track after "the collision". More important, the rating on the other 4 levels was very difficult even when adopting the two basic criteria indicated to rate real occurrences, namely (i) percentage of separation infringement and (ii) whether the controller had detected the loss of separation[3]. The two criteria encounter the same drawback we mentioned in the previous section, that is they both

[3] The two criteria are indicated in the ESARR2 [8] and ESARR4 [10] Severity Classification Schemes from which the SAM Scheme has been derived.

address the severity of the end result of an event, which is often the product of highly specific contextual factors. In other words, safety analysts could not simply observe the RTS event and then rate the severity on the basis of the separation infringed and of the controller detection, as this would have implied rating *the factors that had produced the event in the specific RTS setting* rather than assessing *the severity of a single hazard*. Again, we tried to partially overcome this limitation by profiting of the controllers' expertise. Two workshop sessions were organised after the end of two major simulations, with the objective of reviewing the information gathered on the hazards. Hazards were presented together with the safety scenarios, so that experts could reason about the single hazards not in isolation, but bearing in mind a more realistic situation, that is in interaction with the other system elements. As in the previous case study with the STCA, experts needed to reason about concrete cases in order to draw meaningful estimate on the severity. In the MFF case, scenarios (which had been in a sense validated in the RTS) provided these concrete cases.

The lesson we draw from the MFF case is that the severity rating encountered difficulties in its application even in a case where it could be applied as a post-event classification (i.e. assessing events that were implemented in a simulation). In our opinion, these difficulties stem from the nature itself of the assessment, that is from the fact that experts are asked to assess the severity of an event as representative of a hazard, whilst experts question this very link between hazard and event. They find it hard to trace a linear link between the hazard and the event, and need to draw their estimates from more complex situations, or better said from more realistic situations.

5 An Alternative Approach to Safety Assessment

In previous sections we have presented some issues we faced in the safety assessment process, in particular those due to the severity classification scheme. In this section we would like to draw some tentative lessons learnt from the above discussion.

5.1 Assess Hazardous Situations Rather Than Single Hazards

A direct and simple link between a specific hazard and a given effect is a rare case in complex socio-technical systems. It is usually a complex configuration that jeopardizes the system defenses. However, it is almost impossible to predefine in formal terms these configurations. They can be somehow anticipated only by means of a thorough operational knowledge. Thus technical failures or human errors are better understood only if analyzed in the context of concrete operational scenarios either describing past events or envisaging future situations.

The traditional functional approach, i.e. consider individually all system functions and imagining their possible failures, is an essential starting point of all safety assessments. Nevertheless it should be always complemented by the analysis of the same events in the context of wider *hazardous situations*, which are better handled and understood by operational experts. Such an integrated approach presents at least two main advantages:

1. It gives more opportunity to identify not only the simple functional failures, but also those dysfunctional interactions which generally represent a more insidious threat for the safety of a complex system.
2. It allows the assessors to work jointly on three different aspects of a traditional safety assessment, i.e. hazards, effects and severity. The distinction between hazards and effects does not make sense from an operational point of view. What is seen as the causal factor in a certain context can be easily perceived as the consequence in a different one. With respect to the assessment of severity, critical scenarios (i.e. hazardous situations) appear as the only meaningful context to express a motivated judgment.

5.2 Prioritize Hazards Rather Than Classify Severity

A hidden assumption of functional approach methods is the need to perform an *exhaustive* assessment of all possible hazards. A corollary of such assumption is that analyzing *all* the single functions of a system and identifying *all* their potential failures will ensure that a complete assessment of risks has been performed. Nevertheless the identification of all potential hazards is far from being a viable solution for a variety of reasons.

First of all, socio-technical systems like air traffic management systems are too complex for a detailed identification of all system functions. Secondly, hazards do not derive only from failures of single functions but also from dysfunctional interactions among perfectly working functions. These cannot be identified by analyzing each function separately. In addition, due to their emerging nature, they are anyhow difficult to anticipate in pre-operational phases. Last (but not least) the time available for a safety assessment is generally limited in real situation, so an implicit prioritization is always made.

Based on these considerations, a detailed classification of each hazard in terms of the 5 SCs appears less important than a careful prioritization of what has been identified. The list of hazards can be never considered exhaustive and there is generally no time available to cover all hazards with a specific safety objective. Thus, it is of paramount importance that the most urgent hazards to be mitigated are identified, no matter which is their rating on the Risk Classification Scheme. In analogy with what has been described in Case Study 1, a subset of hazards can be classified as urgent, to make sure that fundamental design decisions are not made before these have been adequately considered. The remaining hazards - at least those which have been considered as relevant - should also be recorded, at least to make sure that they are not forgotten in following design stages.

5.3 Consider Safety Objectives and Mitigation Means Jointly

A sharp separation between safety assessment and design processes does not appear realistic. From the one end, ensuring that safety is independent from production pressures is an important requisite for the credibility of safety targets. In addition, the well known phenomenon of *risk homeostasis* [14] should be always prevented, in order to ensure that safety improvements are not automatically converted in production benefits. On the other end, looking after safety also means thinking about alternative

design solutions, by considering measures on either the technical, the procedural and the training side. The same safety target can be achieved with different design solutions and with considerable variations in terms of cost and availability. Thus practical considerations suggest maintaining an adequate communication flow between safety and design at all stages of safety assessment. Separation and independence is more a requirement for different organizational functions, rather than a prescribed working method.

The need to consider jointly safety objectives and mitigation means is in contrast with traditional FHA, as the FHA is supposed to reason only in terms of abstract functions, without any speculation on how a specific function will be implemented. As for the analysis of hazards and for the identification of severity classes, the approach suggested in this paper goes in a different direction. In our opinion, if pilots and controllers' experience is essential for identifying the possible hazards effects on the system, it is hardly understandable why their expert knowledge should not be used to assess the safety benefits of various design solutions. This implies that mitigation means are considered also at the FHA level, to make sure that operational experts can actually contribute to the definition of safety objectives.

6 Conclusions

In this paper we move from the discussion of what appears to us as a fallacy in current state-of-the-art safety assessment, that is the severity assessment seems to blatantly contradict last-generation safety theories. The line of reasoning is then developed by showing the impact of such fallacy in two case studies. We also present some practical solutions we devised to mitigate the issue. We are well aware that such solutions are mostly *ad hoc* adaptations, far from representing "the solution" to the point we raised.

In our opinion the key tension we encountered in the safety assessment process is between analytical techniques and a more holistic vision. On one side, we need analytical techniques to pinpoint safety threats. On the other, these analyses "tears the system apart" and tends to overlook the fact that in reality the system elements will work together. To address the actual functioning of the system we then need more holistic techniques, to "reassemble" what we have separated for clarity's sake. Our proposal is to ground this holistic view in narrative scenarios, to show system interactions as they happen in the everyday functioning. Future research should address the tension between the two polarities – analytical *versus* holistic – and devise solutions to integrate the two perspectives. At the present moment we see the two polarities as representing a contradictory tension we have to deal with, most likely by reflecting on their complementarities rather than opting for one of the two.

Acknowledgements. The authors would like to express gratitude to the Eurocontrol SPIN Task Force representatives who promoted and supported the FHA study regarding the STCA. Special thanks are due to the ATCC Semmerzake team for hosting the FHA workshop and actively contributing to it. We would also like to thank all the colleagues of the MFF project for the fruitful collaboration on the activity. The MFF project was partially funded by the EU under the TEN-T program. The authors gratefully acknowledge the support provided to this work by the EU project "ReSIST: Resilience for Survivability in IST".

References

1. Perrow, C.: Normal Accidents: Living with High-Risk Technologies. Basic Books, 2nd edn. Princeton University Press, Princeton (1984)
2. Reason, J.T.: Human error. Cambridge University Press, Cambridge (1990)
3. Reason, J.T.: Managing the risks of organizational accidents. Ashgate Publishing Limited, Hampshire (1997)
4. Leveson, N.G.: A New Accident Model for Engineering Safer Systems. Safety Science 42(4), 237–270 (2004)
5. Leveson, N.G.: Safeware. System safety and computers. Addison Wesley Publishing Company, Reading (1995)
6. Hollnagel, E.: Barriers and accident prevention. Ashgate, Hampshire (2004)
7. EUROCONTROL, Air Navigation System Safety Assessment Methodology (SAM) (2006)
8. EUROCONTROL, ESARR 2 - EUROCONTROL Safety Regulatory Requirement. Reporting and Assessment of Safety Occurrences in ATM (2000)
9. Van der Shaaf, T.W., Lucas, D.A., Hale, A.R.: Near miss reporting as a safety tool. Butterworth-Heinemann, Oxford (1991)
10. EUROCONTROL, Air Navigation System Safety Assessment Methodology (SAM) (2004)
11. Pasquini, A., Pozzi, S., McAuley, G.: Eliciting Information for Safety Assessment. Safety Science (in press)
12. Pasquini, A., Pozzi, S.: Evaluation of Air Traffic Management Procedures - Safety Assessment in an Experimental Environment. Reliability Engineering & System Safety 89(1), 105–117 (2005)
13. Mediterranean Free Flight, MFF Operational Concepts & Requirements (2001)
14. Wilde, G.J.S.: Target Risk. Dealing with the Danger of Death, Disease and Damage in Everyday Decisions. PDE Publications, Toronto, Canada (1994)