Reliability Engineering and System Safety, Volume 183, March 2019, pp. 298-310
DOI:10.1016/j.ress.2018.11.024

A new systems engineering structured assurance methodology for complex systems

G. P. Farnell[a], A. J. Saddington[a,1], L. J. Lacey[a]

[a]*Centre for Defence Engineering, Cranfield University, Defence Academy of the United Kingdom, Shrivenham, Swindon, SN6 8LA, UK*

Abstract

As technology advances, systems behaviour becomes more difficult to predict and control, resulting in a lack of systems assurance across the supply chain. Here we describe a structured approach to address the assurance of complex systems developed and operated within highly regulated environments. The new approach is based on a methodology that can address both new and legacy systems, and influences system intervention. We propose an enterprise approach by observing the importance of all organisational contributions to a safe working system throughout the intended project life cycle. This research was catalysed by the need to address the certification of the F-35B stealth fighter for UK operations from 2012 onwards. We offer a pragmatic strategy to achieve systems control by adopting a holistic approach to systems engineering while promoting the development of an enabling environment that can determine system threats and enable appropriate controls. We propose a systematic coordination process to minimise the potential for 'organisational drift'. This holistic approach to systems engineering and assurance is defined as '*systems engineering structured assurance*'. The methodology provides a confidence assessment for a particular product or system while remaining agnostic to regulatory constraints. The diligent completion of the methodology increases systems confidence and informs the regulatory environment.

Keywords: Structured assurance, systems, engineering, safety, methodology, complex

1. Introduction and background

The technological advancement of man-made systems such as aviation, exploration and power-generation enterprises, continues to challenge design engineers and operating teams attempting to understand, predict and control system behaviour. This degree of uncertainty reflects a lack of recognised system assurance across the supply chain, where the supply chain includes the designer, builder, and user communities, including support specialists. The fundamental problem is similar in nature to the International Council on Systems Engineering (INCOSE) grand challenges. Two of the grand challenges are relevant to this paper [1].

- "*System complexity and associated risk is appreciated, characterized and managed*".
- "*Systems engineering provides the analytical framework for designing and predicting the behavior for trusted, resilient systems*".

The need to conduct first-time certification of the F-35B stealth fighter for operations within the US and UK legislative environments led to the development of a systems engineering methodology to ensure confidence in the F-35B system against

Email address: a.j.saddington@cranfield.ac.uk (A. J. Saddington)
[1]Corresponding author.

a background of a programme suffering significant technical and operational challenges. The 'systems engineering structured assurance' (SESA) methodology describes the overall approach used to achieve the enabling conditions and engineering processes required during systems development, certification and operations. This approach allows system threats to be understood and the appropriate controls to be applied, and proposes a systematic coordination strategy to minimise the potential for organisational or systems-related 'drift'. The US defense systems engineering executive cited shortfalls in system engineering expertise and practice throughout the systems engineering communities that support acquisition of complex programmes [2]. The process enhancements described in this paper should help to bridge these gaps by providing a "leadership methodology" to facilitate the design and the operation of complex systems. The methodology links the system's 'builder', 'buyer' and 'user' communities enabling the behaviours of equipment, people and organisations to be better understood and constrained.

The F-35 Joint Strike Fighter design formally began during 2001 with testing beginning in 2006 along with a Director, Operational Test and Evaluation (DOT&E) audit. During 2009, Farnell ordered the first three F-35B air systems into production for the UK government, with the first planned to emerge from the production line during July 2012. As the Type Airworthiness Authority (TAA) for the F-35B, he was accountable for certification of the air system to operate safely, with a UK pilot, operating within UK legislation, and in accordance with the emerging UK Military Aviation Authority regulations (MAA formed April 2010). At this time a number of separate collaborative arrangements were developed, with varying levels of formality, to fully understand the F-35B air system. These included: the JSF Joint Program Office for financing and managing programme delivery; the Joint Operational Test Team for participation in whole system testing and gaining access to data; the UK Defence Science and Technology Laboratories to access subject matter experts (SMEs), to further develop SME knowledge, and to elevate the UK expertise and understanding in a deliberate and dedicated manner; NavAir to leverage system understanding, participate in system analysis and gaining access to data; DOT&E for exchange of information and understanding to develop an enhanced insight and healthy critique of the programme, thereby informing an enhanced intelligent customer status. These collaborative arrangements were ground breaking, enabling access to technical data as well as being able to influence the way the UK air systems would be constrained and employed; this approach has been characterised as creating the 'enabling environment'. The insights and data obtained from the collective collaborative arrangements were used to develop a comprehensive picture of both performance and safety [3]. In 2011, working with immature regulations anchored in legacy systems, the MAA informed the TAA that the UK's F-35B aircraft would not receive a 'Release to Service' as the design was immature and the design standards were not recognised. At this time work began on a parallel research programme to the certification effort, to identify how a complex system could potentially be understood, given a failure to design in accordance with the required standards [3]. A comparison was, therefore, conducted of the US air system design versus UK standards to develop a gap analysis.

Embracing all systems, at varying modification standards (hardware and software), required a level of detailed analysis in preparation of making an argument for a safety case. It became clear that, during a build inspection, the air system's out-turn build standard was different to the intended design. In addition, air vehicle testing, observed in collaboration with DOT&E, identified that out-turn performance was at variance to that predicted and expected. The realisation that viewing safety in isolation would not provide a complete assurance picture was ascertained during the period 2010-2012. A methodology was developed, therefore, to assess performance (versus expected) and safety (versus expectation from the required US and UK standards). The assurance methodology (subsequently known as SESA) was born as a means of understanding, characterising

and assuring both system performance and safety. This methodology enabled a complete system assurance to be conducted to fully understand the risk exposure, at out-turn build standard, as well as characterising the strategy to contain the risks. SESA enabled provision of the assurance to operate the three UK air systems under a personally-authorised military flight test certificate between July 2012 and November 2017. Application of SESA prevailed, driving a culture of capturing data from US and UK air systems and addressing critical system fixes for both the US and the UK. SESA was an unpopular approach within some programme delivery communities as it clearly articulated system shortfalls. However, the targeted data gathering, system understanding and judicious application of constraint, alongside continued data gathering whilst under military test permit, resulted in the MAA granting the UK F-35B a release to service in November 2017.

2. Understanding systems behaviour

Degani [4] studied human-automation interactions in which design, procedures, management and training contributed to systems failure. Understanding such complex systems required an awareness of the problems with human-automation interactions, but the original design must also be fit for purpose [4]. In this context, a complex system can be characterised by emergent properties. A system component may have a particular functionality but, unlike hierarchical systems, this is not recognisable as a sub-function of the global functionality. Instead, the behaviour of several connected components display side effects that contribute to the global functionality. Each behaviour has a side-effect and the sum of the side-effects produces the out-turn functionality [5]. Hence, the global functionality of a system with 'emergent properties' is the sum of all 'side effects', of all emergent properties and functionalities. The systems described in this paper can, therefore, be regarded as displaying 'emergent functionality'.

The analysis of accidents in enterprises such as commercial aircraft, general transport systems, power generation process controls, and medical devices, revealed that both the design and management often contribute to system failure [6]. The study, jointly conducted by the National Aeronautics and Space Administration (NASA) and the Department of Transport, also concluded that the four causal factors studied (design, procedures, management and training) were interdependent. The study found that the medical sector, in particular, has continued to face the challenges of inadequate design in addition to the continual struggle to create the appropriate conditions for a positive safety culture [6].

A Government Accountability Office report highlighted the continuing struggle experienced by the Food and Drug Administration (FDA) when recalling faulty and potentially faulty medical devices within the US medical community [7]. Even medical devices with FDA approval are often recalled due to associated serious health problems [8]. However, approximately 50% of device recalls fail, leaving hazardous devices in circulation and use [7]. The main causes of medical device failure include design errors (>25%) and user errors (>50%) with management and procedural errors making up the remainder [9]. The inability to track devices in the field is an additional, second-order problem.

The space shuttle Challenger disintegrated during its tenth mission in 1986 and the Columbia disaster in 2003 illustrated how the previous paradigm of 'organisational drift' was repeated. Organisational drift occurs when the operational performance of a system varies from the design intention without the full understanding of the managing organisation [10]. The Presidential Commission investigation found an array of failures that could be traced to the socio-economic strain resulting from both the cost growth of the NASA shuttle programme and the budgetary pressures imposed by US authorities in response to this [11]. Dekker [12] argues that the production pressures at NASA became institutionalised and it was business as usual to negotiate and re-negotiate levels of risk, causing the organisation and its decision makers to drift incrementally into a completely new

socio-technical construct. Apollo 11, Ariane 5 and the Mars Polar Lander suffered problems due to system failures in which every component worked as planned but the overall systems failed in a surprising and unexpected manner [13]. This further illustrates a failure to understand the complexity associated with entire systems or indeed systems of systems.

2.1. Assurance in complex systems

Understanding the potential behaviour of complex systems presents a transformational challenge making it difficult to determine the true risks of operation, particularly when new technologies or new ways of working are introduced [12]. A new system can present challenges associated with V-model testing[2]. Galin [14] estimated that approximately 70% of errors embedded in safety-critical software are introduced during the requirement specification and architectural design phases. The rapid increase in avionic software content and the implication of design debt[3], shows that approximately 80% of all errors are discovered late during the V-model, i.e. during system integration testing or later [16]. The rework effort required to rectify problems identified this late in the life cycle can be 300 to 1000 times the cost of the original development [14]. Many residual errors are unlikely to become evident until system development is reasonably mature, and some residual errors may not be discovered until there has been some operational experience [14]. These circumstances are exacerbated when there is an incomplete understanding of the system, and such a situation can be fuelled by security and intellectual property concerns, which limit access to product or project data.

The challenges described above can be addressed by developing an assurance partnership in which an appropriate environment is provided for the designer, builder and user participants during the requirement-setting process. This partnership thus creates a single-team, whole-system arrangement. Creating an appropriate enabling environment to provide a coherent regime of V-model testing and assurance would increase the likelihood of success.

Another challenge, described as 'practical drift', is observed when an organisation operates a system in a completely different way to that originally intended, i.e. beyond the stated performance and safety parameters. For example, new users may adapt system operating procedures and adopting an approach to business that evolves over time. This operational performance may in turn reduce the effectiveness of the original performance and protection arrangements [10]. A whole-life methodology that provides appropriate conditions for the continuing recognition of whole-system and whole-environment awareness would be particularly valuable from the user perspective. This was the catalyst for the initiation of the SESA concept and methodology [3], which is developed in Section 3.

2.2. Shortfalls in knowledge and practice

The performance of complex systems is difficult to predict with appropriate levels of confidence [17]. Escalation in software-involved systems, the globalisation of software provision and a lack of consensus concerning the predictability of software mean that traditional methods to predict system performance and safety have become less relevant [18]. When combined with the continuation of error-inducing practices at the requirement-design stage, and the discovery of errors during

[2]The 'V'-model is a term applied to a range of models, from a conceptual model designed to produce a simplified understanding of the complexity associated with systems development, to more detailed, rigorous development life cycle models and project management models.

[3]Technical debt, also known as design debt, "...represents the cost of the accumulated amount of rework that will be necessary to correct and/or recover from the deviation between: the current design of the system, versus a design that is minimally complex yet sufficiently complete to ensure correctness and consistency for timely delivery." [15].

later testing stages, this presents a challenge to system certification and in particular the certification of software-involved systems [16]. Leveson [18] recommends that a thorough understanding of system threats must be prioritised and that such work should be initiated during early process design. This requires a collaborative relationship between the designer/builder, the regulator and the user communities at the outset of a project, which may not be practical under competitive commercial practices. The collaborative relationship would require suitable skills and co-operation to facilitate an appropriate understanding of the system.

The acquisition of the F-35B by the UK authorities provided several challenges given the inconsistencies between US and UK socio-technical policies and in the arrangements necessary to address variations in technical standards; one example being the UK regulator expecting a system to be designed to Defence Standard (DEF STAN) 00-970, whereas the US were designing to Military Handbook MIL-HDBK-516B. The ongoing challenges for the development and test programme for the F-35 are explained in detail by the DOT&E [19], which states that there are nine reliability measures for the programme. These are represented by three for each variant: mean flight hours between critical failure (MFHBCF); mean flight hours between removal (MFHBR); and mean flight hours between maintenance event (MFHBME) [19, p. 60]. Additionally, the mean time to repair (MTTR) and provision of technical data were inadequate. These challenges were difficult to overcome when combined with the need for retrospective access to detailed technical and test data that addressed numerous first-time technologies.

The original reliability measures (MFHBCF, MFHBR and MFHBME) set for the F-35 programme struggled to reach the required targets [19]. The following system shortfalls represented the major contributors towards potential system critical failures: electrical power system; power and thermal management system; integrated air vehicle architecture (which includes the Integrated Core Processing System and the cockpit displays including the helmet mounted display system); access doors and covers; landing gear; oxygen system; stabilisers; lift fan system; crew escape; and flight control system [19, p. 63]. The threat associated with these system-critical failures presented a challenge to the progression of operational tactics and training due to the resulting low reliability and availability at that time [19, p. 63], and the consequential impact on MTTR as a result of the limited technical data. A methodology was required, therefore, to embrace not just the engineering but the management of equipment, people and process. The UK had to develop an enhanced level of technical understanding in order to orchestrate a safe system of work for pilots, maintainers and data analysts. The methodology to achieve this involved an enabling environment (original equipment manufacturers and US Government agencies) and a framework of thinking that went beyond just safety to include both safety and performance together. Ball [20] has illustrated that these elements do not separate well when designing and operating complex systems.

Figure 1 summarises the challenges facing the designers and operators of complex systems, where the lessons identified from accident causation, V-model testing and practical drift have been integrated into a representative project life cycle. The challenges endure throughout the project so an appropriate enabling environment should ideally be developed early in the project life cycle and effort is required to maintain this environment throughout. System 'assessment' should be heavy during concept and design/development, reducing during manufacture and after certification, with a light touch through in-service life.

First, a common view of performance and the necessary degree of safety protection is required, implying that the objectives and philosophies of the contributing individuals should be aligned within organisations and across the relevant organisational boundaries. The arrangements required to underwrite the achievement of common objectives may be termed enabling arrangements. Achieving such an enabling arrangement during project initiation is an idealistic approach, whereas retrospective application, although feasible, would be sub-optimal. Second, the necessity to fully understand threats and the associated risks

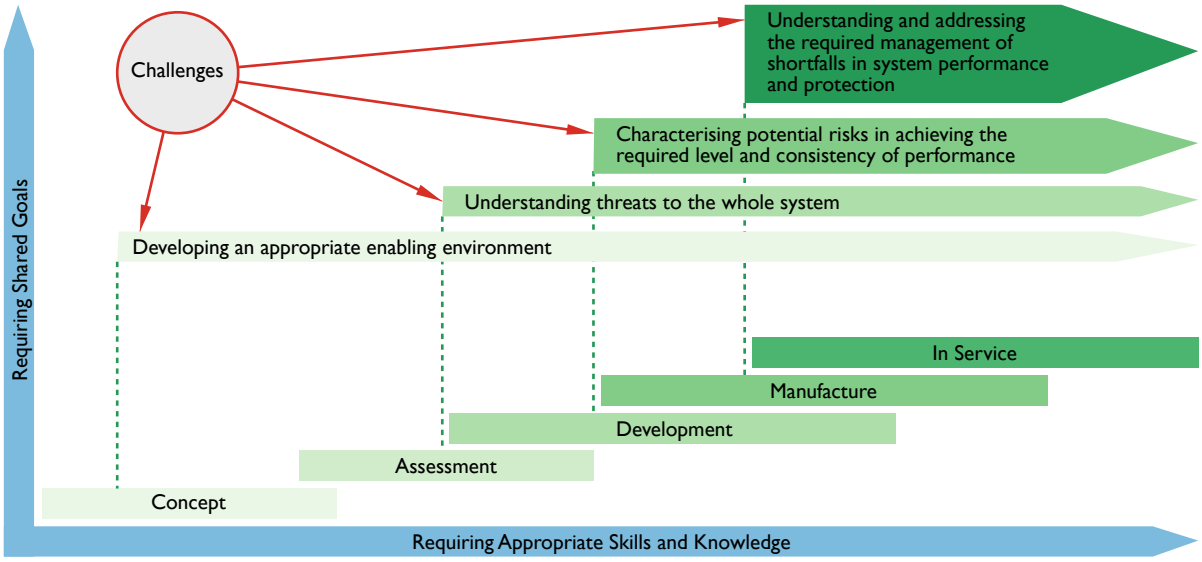

Figure 1: Challenges that must be addressed to understand complex systems

within the specific context of the user environment are based on the ideas of Perrow [21], Hollnagel et al. [22] and Dekker [23]. Third, the grand challenges set by INCOSE [1] place further emphasis on whole-life management of the system as built, as well as raising the level of ambition for a systems engineering approach to improve the degree of system trustworthiness. The latter would require a unified approach combining properties such as reliability, safety and security. Feiler et al. [24] argued that such a unified approach, embracing the properties of performance and protection, had yet to be achieved and hence proposed "an improvement strategy for an integrate-then-build practice" to increase confidence with the verification and validation. More recently a complementary concept, known as RAMSSHEEP[4], has been developed. Whilst RAMS[5] was devised in the 1970s, RAMSSHEEP was created as a risk-driven maintenance concept [25]. Acknowledging that RAMSSHEEP post-dates the SESA methodology, which was devised in 2012 to underwrite the safety case for the UK F-35 programme, it could nevertheless form a useful sub-set within the SESA methodology, being very similar to the probabilistic techniques used in the SESA hazard analysis. A whole-life methodology that provides appropriate conditions for the continuing recognition of whole-system and whole-environment awareness would be particularly valuable from the user perspective. This was the catalyst for the initiation of the SESA concept and methodology [3], which is developed in Section 3.

3. Developing an enabling environment

In responding to the current challenges facing designers and operators, Figure 2 draws on the new 'Military Airworthiness Regime' proposed by Haddon-Cave [26], which includes the MOD 'Four Pillars of Airworthiness'[6]. An assurance strategy anchored around four pillars ('competence'; 'design standards'; 'independence'; 'assurance management systems'), known as the 4P assurance strategy, incorporates the skills of the individual and the system's design standards: the whole system

[4]RAMSSHEEP: reliability, availability, maintainability, safety, security, health, environmental, economics and politics.

[5]RAMS: reliability, availability, maintainability and safety.

[6]The MOD Four Pillars of airworthiness are: "use of competent people"; "use of recognised standards"; "independent assessment"; "safety management system".

(enterprise) may include many systems and many design standards. The 4P assurance strategy embraces the constituent parts of the system, the necessary competences, and the appropriate technical design standards. These standards are supported by an assurance and environmental management process that can be enhanced through alliances with a likeminded partner. Alliancing provides ready access to subject matter experts and training, as well as sharing opportunities to allow individuals and organisations to learn from experience. This 4P assurance strategy provides a clear communication framework to determine the activities necessary for a solid foundation underpinning either a specific system safety case or a more comprehensive system assurance case.

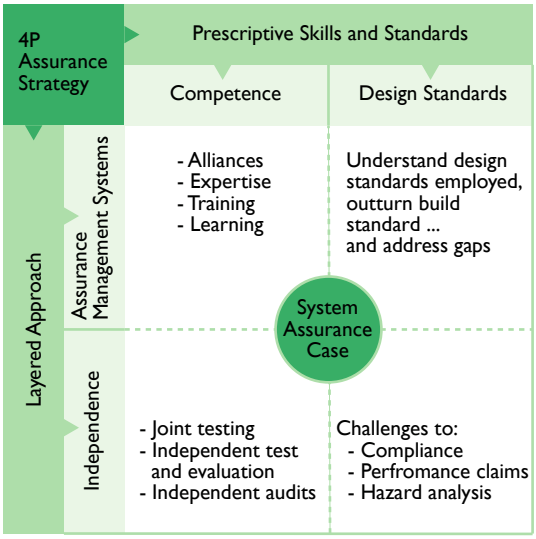

Figure 2: The 4P assurance strategy covering the constituent parts of the system, the necessary competences, and appropriate technical design standards

3.1. Understanding threats to the whole system

A further development of the 4P assurance strategy is the idea of a partnership between the customer and designer communities, whereby a deliberate strategy is pursued to develop the necessary systems engineering evidence by developing the assurance strategy alongside the system being designed and tested (Figure 3). Such a partnership would demand collaborative work, in which system requirements are shared, incrementally developed, and jointly reviewed by the user and designer communities [24]. System threats must be identified and understood, both before and during system development, and with focus on the rectification of missing or inappropriate control actions. This could be achieved, for example, by applying 'systems theoretic process analysis' [18], particularly where the early identification and mapping of control actions enables the emerging design to be configured correctly. Such assurance partnerships can help to address the demanding systems engineering challenges set by the INCOSE objectives. Developing a real understanding of any complex system is best achieved through a structured approach that defines the requirements, the impact of differing socio-technical systems, and the potential trade-off between the best outcome for safety and system performance. Creating the most appropriate conditions for the development and operation of a complex system depends on an appropriate enabling environment. Understanding the threats to the system from a whole-system perspective and addressing the shortfalls provides a means to control the system configuration. Acknowledging and addressing system shortfalls provides an opportunity for system decision-making during the design stages. Driving a consistent approach throughout the supply chain will require a collaborative approach among all stakeholders, with initial

conditions established by creating the optimal enabling environment. A 4P assurance strategy should provide the conditions necessary for success.

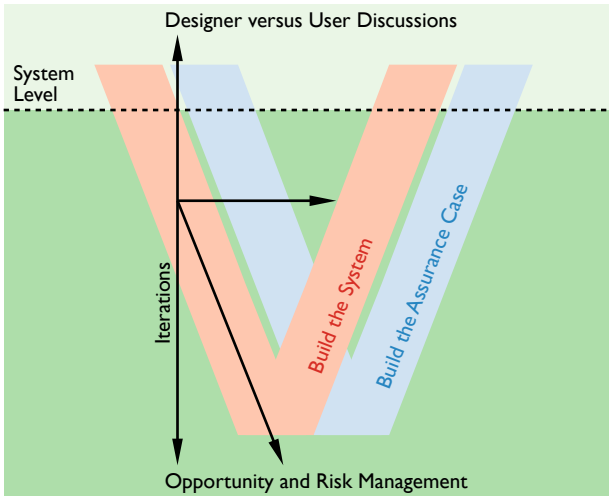

Figure 3: Understanding threats through assurance, adapted from Feiler et al. [24]

3.2. Characterising risk for the whole system

Proactive planning to identify and eradicate system shortfalls, as advocated by Galin [14] and Feiler et al. [24], is considered a sensible approach, and 'systems theoretic process analysis' offers the opportunity to rigorously examine system control actions. This would require a thorough understanding of the scope of the system undergoing development and certification, and agreement by all stakeholders. Having identified system shortfalls during early development, when the design architecture is unable to eradicate threats, a methodology characterising the safety and effectiveness risks would be helpful. A safety and effectiveness framework (Figure 4) can characterise the threats and the associated risks across the system susceptibility[7] and vulnerability[8] continuum, whereby the risks associated with an adverse event (and the resulting failure) can be mapped directly to the system. This approach can highlight where system failures could affect both safety and system performance, and can identify the corresponding system control relationships. From a supply-chain perspective, the joint involvement of the designer, builder and user communities minimises the many challenges arising during system development. Figure 4 also shows how events and actions across the system can be modelled to identify potential trade-space within the selected boundary of the required performance parameters, as well as identifying the requirement for additional safety features. The primary requirement is to identify opportunities to include design characteristics that will enhance performance and protection as well as optimise the balance or trade-space between these objectives. Positive contributors that improve performance and safety defences (e.g. margin, resilience, redundancy and adaptive behaviour) could be incorporated into the modelled activities to develop a comprehensive framework that can be used to brief all stakeholders on the system strategy. The framework, shown in Figure 4, also allows the coordination of safety and effectiveness strategies, and the detailed implementation and configuration of safety and effectiveness activities. Following the identification of shortfalls that affect performance and/or

[7] Susceptibility refers to "... the inability of a system to avoid being disrupted by one or more damage mechanisms in the pursuit of its mission" [20, p. 445].

[8] Vulnerability refers to "... the inability of a system to withstand the damage ... and its liability to serious damage or destruction" [20, p. 603].

safety, it is advantageous to apply a methodology that coordinates planned system interventions and determines the impact of such variation with respect to the management of the system baseline.

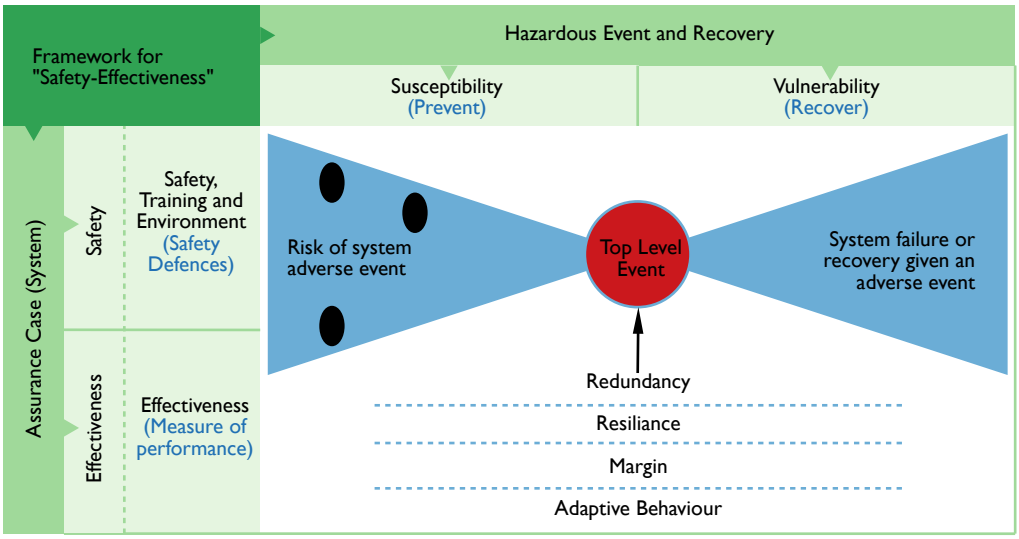

Figure 4: Safety and effectiveness framework, adapted from Ball [20]

3.3. Understanding and addressing management of system shortfalls

The design-to-operate framework, shown in Figure 5 provides the framework for managing, maintaining and planning the SESA interventions. It offers the opportunity to formally record system shortfalls identified during the development of the system assurance case, as well as the residual shortfalls from the emerging assurance themes. These themes include training, inspection, management and procedures (Figure 5, 'discretion' row) that could be delivered as part of the system by the designer/builder organisation (Figure 5, 'provided system performance' row). The shortfalls would be mapped to the operator (Figure 5, 'operator/maintainer' column) where additional mitigations might be identified and applied by the user communities. Similar to the techniques applied in the safety and effectiveness framework, the coordination of the system areas for improvement and associated options for improvement can be developed at the stakeholder strategy level. Additionally, the design-to-operate framework could also be applied at a more detailed implementation level, by fully engaging the supply chain.

Drawing all the frameworks (Figures 2–5) together into a single methodology has been defined as 'systems engineering structured assurance' (SESA) [3]. A new process definition characterised during this research has been captured and is reported herein. A process diagram describing the application of the complete SESA methodology is provided in Figure 6. From left to right, the SESA methodology can be tailored to support three chosen project scenarios: initiatiating a new project; joining a legacy project; and introducing significant changes to an existing project. System requirements and the corresponding system assurance would be developed, and would lead into a dedicated task to initiate system assurance alongside the design and manufacture of the system. The priority during the early phase of a new project is the creation of an appropriate enabling environment, where the intention is to encourage and facilitate close cooperation between the designer and operator communities.

As discussed in the introduction, the UK introduced F-35 certification activity as a legacy project, and an enabling environment was, therefore, developed retrospectively. This is highlighted in Figure 6, which shows how the SESA methodology was

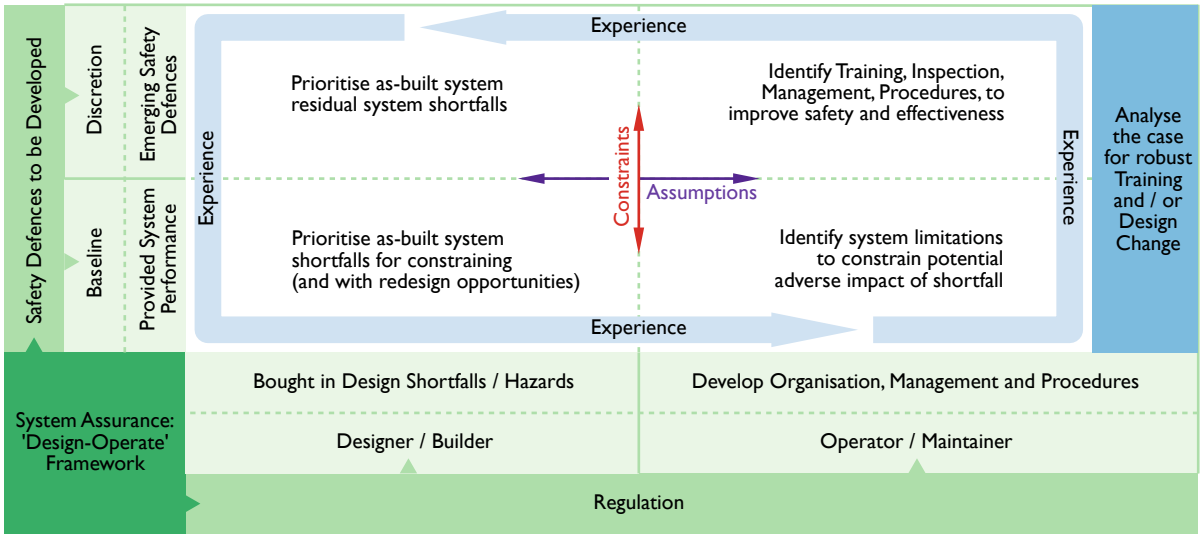

Figure 5: Design-to-operate framework

mapped to the F-35 assurance. Accordingly, residual system performance and safety shortfalls created a business-case decision between tolerating the shortfalls or introducing changes to rectify them. Introducing significant changes to an existing project (even when the SESA methodology has not previously been applied) will benefit from the application of the SESA methodology to all aspects of the change or intervention activity. The change could take the form of a modified design, additional or modified training, or additional or modified organisation and management procedures. The organisation and project objectives are developed sequentially once the enabling environment has created the appropriate conditions to proceed. System threats are understood and where possible addressed, and where necessary mitigations are developed against residual threats as well as anticipating additional threats that may arise during operations. SESA also embraces the design-to-operate framework in which system threats can be mapped against existing and planned mitigation measures. This provides excellent coordination and configuration control, and offers real benefits for the operation and through-life management of complex systems. The idea behind the SESA methodology is to use a systematic approach for understanding the system and to develop a confidence assessment for any of the three project scenarios identified above.

4. The application of SESA to new system programmes – a case study on the UK F-35B

During the development of a new system, challenges will create pressure from within the project and from outside the immediate system activity. The external pressures are likely to be socio-technical and socio-economic in nature, and are often associated with legislative and/or financial constraints. There have been significant innovations in the design of human-machine interfaces, new computing architectures and new software standards for systems operation, including safety-critical systems. The programming language ‘C’ was used by the designer/builder in the F-35 project to improve performance whilst reducing the cost of further development and ownership, and represents its first application in a safety-critical setting. However, the use of C in safety-critical applications attracted regulatory scrutiny during certification of the F-35B, given that the UK Ministry of Defence’s Military Airworthiness Authority regulations originally prescribed a different software approach. This represents an example of the impact associated with differing socio-technical policies. Furthermore, significant technical advancement has been achieved by the sheer scale of software content [16] encouraged by customer demands for more capability. This

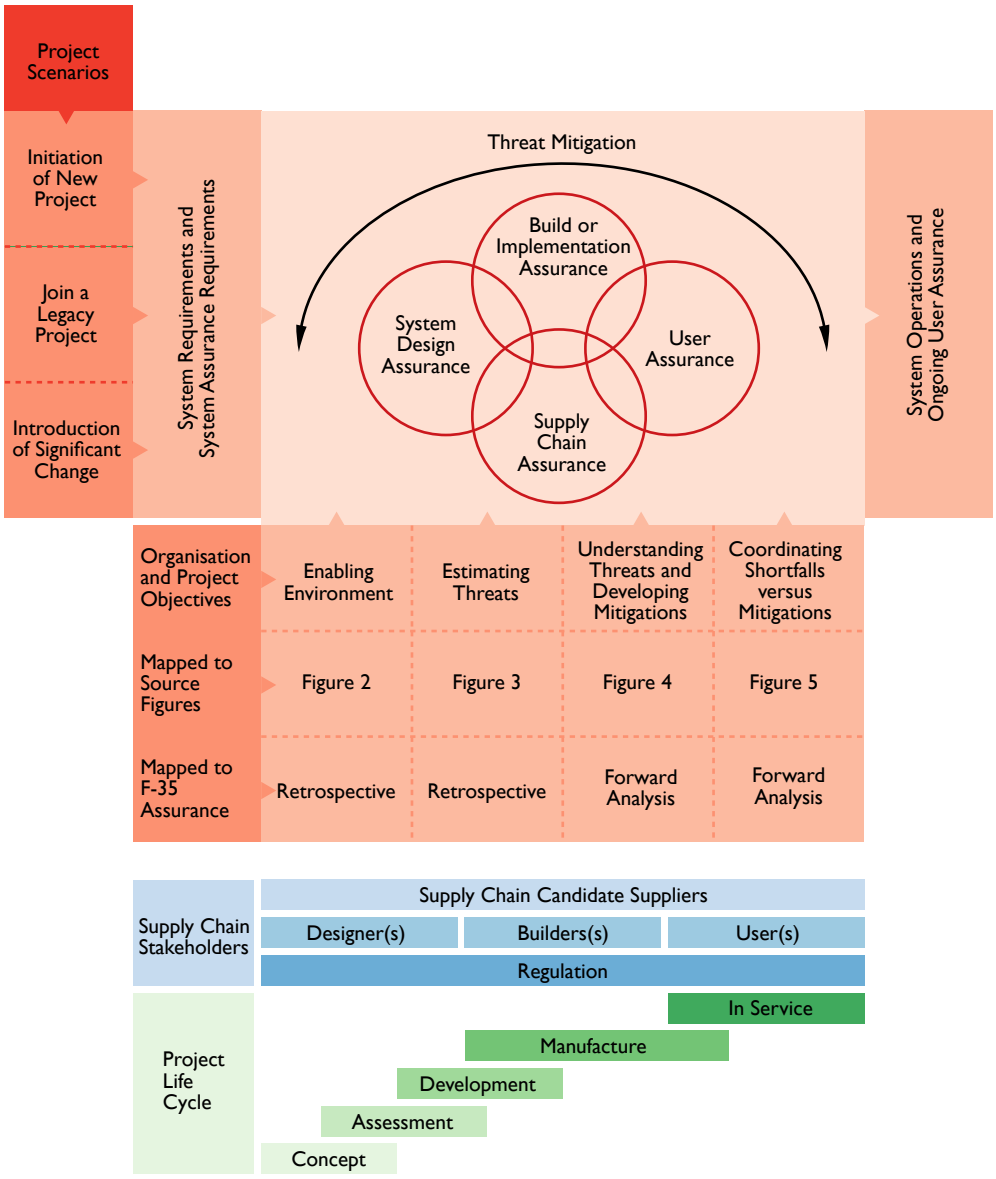

Figure 6: Generic systems engineering structured assurance (SESA) approach across the supply chain showing the mapping to the F-35 assurance

brings the challenge of further system problem reports during software testing, in turn highlighting pressures of time, cost and technical prioritisation. An optimised supply chain, which collaborates to address the socio-economic and socio-technical challenges, and encourages the optimisation of organisational relationships, is an idealistic goal for the enabling environment.

An assurance methodology was developed, and employed, to confidently identify the gaps in performance and safety standards specifically for the F-35B air system. This methodology became SESA and was applied and introduced during the manufacturing phase with the three UK air systems in build. This was, therefore, a retrospective application to understand the system performance, using data from testing US assets, and to apply constraints across the whole system: for example, constraining system performance limits, the way the pilot operates, the degree and type of pilot and maintainer training and the way the operating organisation would be configured. The SESA methodology was used as the only means of fully understanding the system and its shortfalls. It allowed system understanding to be demonstrated as well as the implementation

of appropriate control measures. The approach provided a mechanism for addressing the requirements of UK legislation in a way hitherto not considered, where out-turn performance and safety were considered side-by-side. Had the SESA approach been actively used during early design and test, whilst system changes could have been introduced earlier during the V-model, then SESA would have reduced the number of system shortfalls.

There is no means of absolutely proving that SESA, if used early during the design, would have changed the current outcome for the F-35. However, if the enterprise behaviours had been established from the outset during the early part of the 'V', rather than during manufacture, the opportunity to introduce change is more readily available. The question then is how would we know the problems that the air system will face during design? Several opportunities offer merit to address the identification of such problems. Using the problem of onboard oxygen generation as one example. Firstly, testing of the oxygen subsystems indicated several emergent problems (as did many other systems after retrospective investigation) that could have been addressed at a very early stage before incorporation into the air vehicle. Secondly, benchmarking against similar subsystem designs; a previous aircraft, using a similar configuration of subsystem components had suffered problems. Thirdly, multidisciplinary review teams of suitably qualified and experienced personnel (SQEP) would have identified the problems early; creating the right environment is a key stepping stone of SESA. The SQEP teams would have user and operator knowledge associated with first and second points described above. Retrospective application of SESA was demonstrated as a means to control the release and approval of system capabilities. It is preferable, however, to apply SESA to the full system life cycle, applying any system interventions or changes at the earliest possible stage and thereby reducing the cost of intervention. Today, we are beginning to replicate the SESA concept in early development using digital-twin models of systems and often the whole system where possible. This enables the system performance to be estimated through emulating the lower level system characteristics.

Figure 7 summarises the challenges and strategic effectors that influence the development of innovative and complex systems. Gaps in policy, knowledge and practice across the supply chain hinder attempts to optimise systems engineering both within organisation and across borders with collaborating entities [3]. The overlapping influence lobes in Figure 7 highlight the blurring between policies, regulation, and culture. This can lead to confusion and the need for interpretation between legislation, regulation and policies. Confusion over the interpretation of a statuary instrument in any system design or certification activity could generate adversarial interactions when seeking regulatory approval, and could potentially lead to myopic behaviour, organisational and inter-organisational 'stove-piping' and a poor teaming approach, i.e. the opposite of an appropriate enabling environment. Examples of policy gaps due to insufficient or absent guidance include drift, automation and resilience (Figure 7). Economic and regulatory inconsistency, both within and across borders, leads to an overwhelming dependency on the local interpretation of policies. The designer/builder was empowered by the US authorities to select the design standards for the F-35 and left the various international user communities to address any variations relative to extant socio-technical policies. Embracing the broader socio-technical construct is an important step towards understanding and addressing environmental concerns.

4.1. Developing an assurance methodology

To facilitate comprehensive analysis, systems should ideally be considered from a whole-system perspective, where stakeholders develop a healthy interest in the strategic context as highlighted above. A whole-system approach should include the physical system, supporting information and control systems, operators, maintainers, and operator/maintainer training

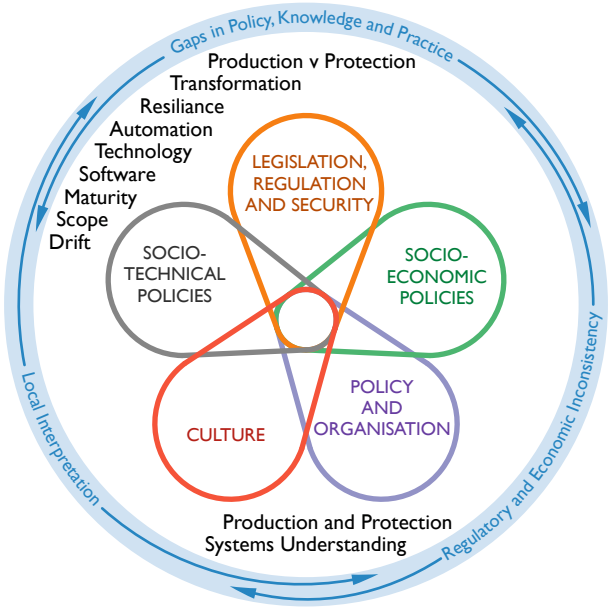

Figure 7: Strategic effectors and challenges

systems (Figure 8). The socio-technical and socio-economic context (and influencing mechanisms) should be investigated to understand the broader system of influence and interest. Figure 8 shows how to construct the hierarchical relationships among the influencing entities, and how to develop side-by-side models that show differences between the designer/builder and operator/maintainer environments. Understanding the socio-technical context helps to identify variations in policy and regulation, thus allowing rational assessment of the system's design implications prior to its release into the operator/maintainer environment. For example, examining the specific threats understood within the design versus operating environments for the F-35 facilitated the continuation of an appropriate enabling environment.

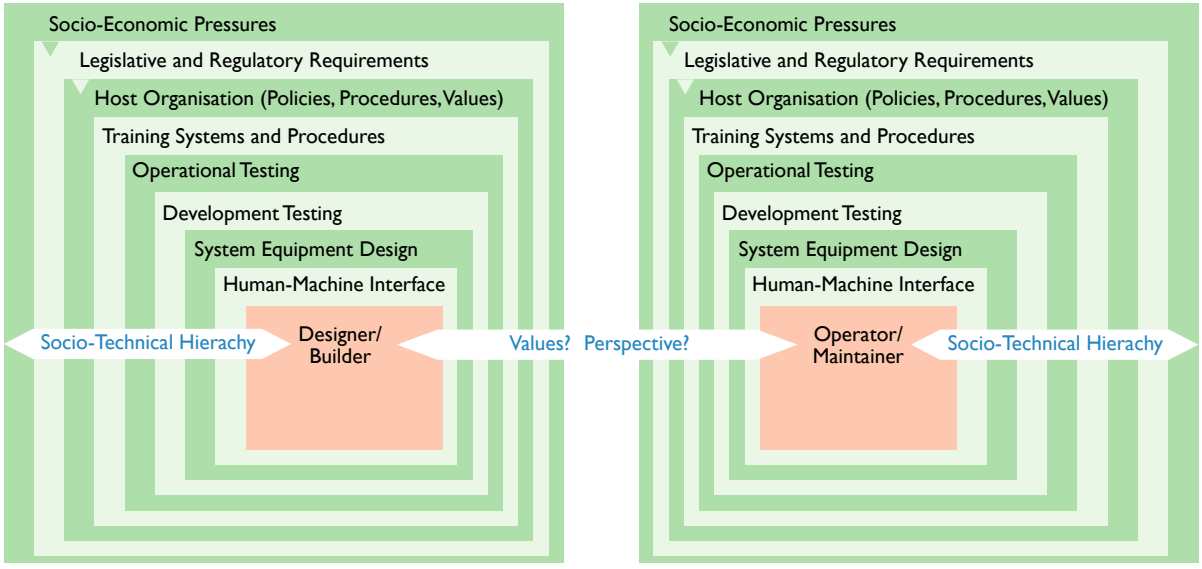

Figure 8: The extended assurance control structure

The designer/builder and operator/maintainer organisations may well have completely different socio-economic and socio-

technical systems, and the pressures, experiences and perspectives of the employees across the supply chain may vary considerably. A designer organisation can reduce the exposure to risk to as low as reasonably practicable, while addressing the tolerability of risk [27] and pursuing a thorough technical safety argument, by adopting the same objectives and similar values and perspectives as the operator/maintainer community. The builder-buyer-user concept (Figure 9) can be combined with the 4P assurance strategy (Figure 2) to create a harmonised collaborative approach. The enabling environment in Figure 9 would ideally include the supply chain and its supporting entities.

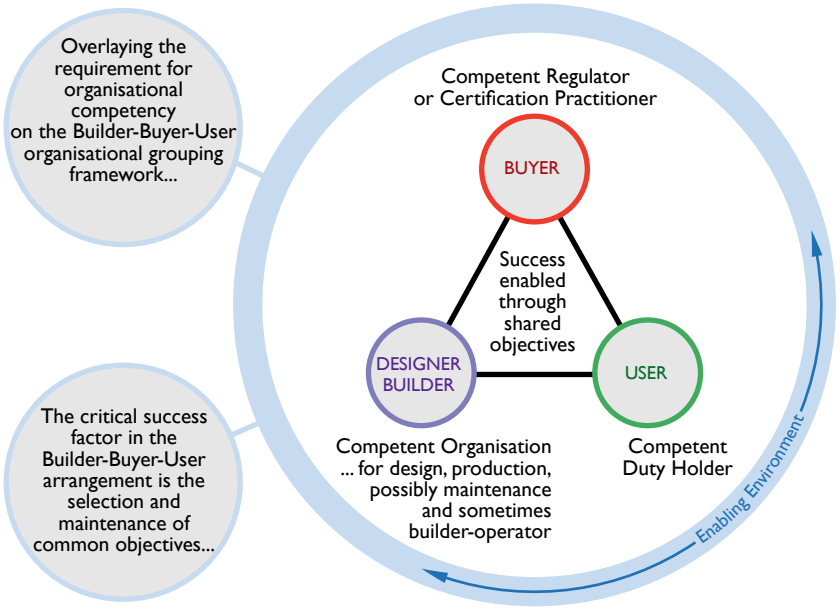

Figure 9: The builder-buyer-user collaborative approach

Despite the shortcomings and criticism of UK safety case management [28, 29], the retrospective development of a safety case and technical safety argument is valuable as a structured means to develop and collate system evidence, and this approach would equally apply to an assurance case. Furthermore, this assurance technique develops confidence through a phased approach, whereby capability is constrained below that offered by the developed system and then released incrementally as confidence improves. A phased approach was employed to incrementally approve the use of the F-35 capability in a measured and controlled way, thereby improving confidence. The employment of innovative features was limited until a stable baseline, and the risk associated with moving each step forward, was understood. The application of the SESA methodology fundamentally depends on the organisational enabling environment. Moreover, a hazard-seeking culture, as opposed to a culture of avoiding the challenges, provides a healthy posture, enabling the most appropriate mind-set to thoroughly identify and understand threats to the system. This approach is particularly important when there are high levels of uncertainty, a higher than usual level of software content, and organisational challenges associated with accessing data in the early stages of programme delivery. Beyond this point it becomes impossible to extricate as the cost of intervention escalates beyond reasonable cost-benefit calculations.

To better understand the potential impact of system problems and mitigation strategies, it is necessary to address the system problems while adopting a risk management approach that identifies all critical and catastrophic risks. A probabilistic risk assessment, albeit often a theoretical estimate during early project work, should be undertaken but considered as indicative, with the focus on understanding and addressing the threats leading to critical and catastrophic risks. UK legislation indicates

that 10^{-6} is the minimum risk-to-life requirement, and the only satisfactory way to achieve a high confidence outcome (given the unpredictability of systems) is to minimise the severity associated with all critical and catastrophic hazards [27]. This approach can only be achieved effectively by creating and maintaining an appropriate enabling environment.

Figure 10 is a whole-system representation of the safety and effectiveness framework (see Figure 4) where themes can be developed from either traditional or transformational system features that lead to key events. It shows a generic, yet systematic, approach to investigate innovative system design interventions and represents a simplified template of that actually employed to fully understand the risks associated with the intended F-35 new ways of operating and/or the application of new technology. These design interventions present both a threat to understanding, and a threat to estimating, emergent behaviour of individual systems and, importantly, the emergent behaviour of the whole system. It identifies transformational features that improve performance, whilst at the same time inducing threats to the sustainment of the system, and searches for missing control actions. Figure 10 further illustrates how potential safety-critical events such as collision, loss of control, the onset of fire and/or explosion and failure to generate the required electrical power, all flow from the more predictable system failures, as well as flowing from uncertainty associated with new technology. For example, an anomaly or failure in the software controlling life support oxygen generation within the F-35 air system could result in a loss of situational awareness through pilot hypoxia, leading to possible loss of control and loss of life. Moreover, recognising the criticality of oxygen generation early could lead to greater testing, thereby identifying performance issues, which translate into safety critical issues, when taking the system to its performance extremes. This is vitally important analysis given the necessity to determine the 'safe to operate' status of the air system needed to underwrite the certification effort. This analysis should never stop, as a continual review is required to support the 'safe-to-operate' assessment alongside the effectiveness (performance) assessment, as the air system continually evolves through incremental capability development in addition to the necessity to both guide and respond to the inevitable development in achieving mission objectives.

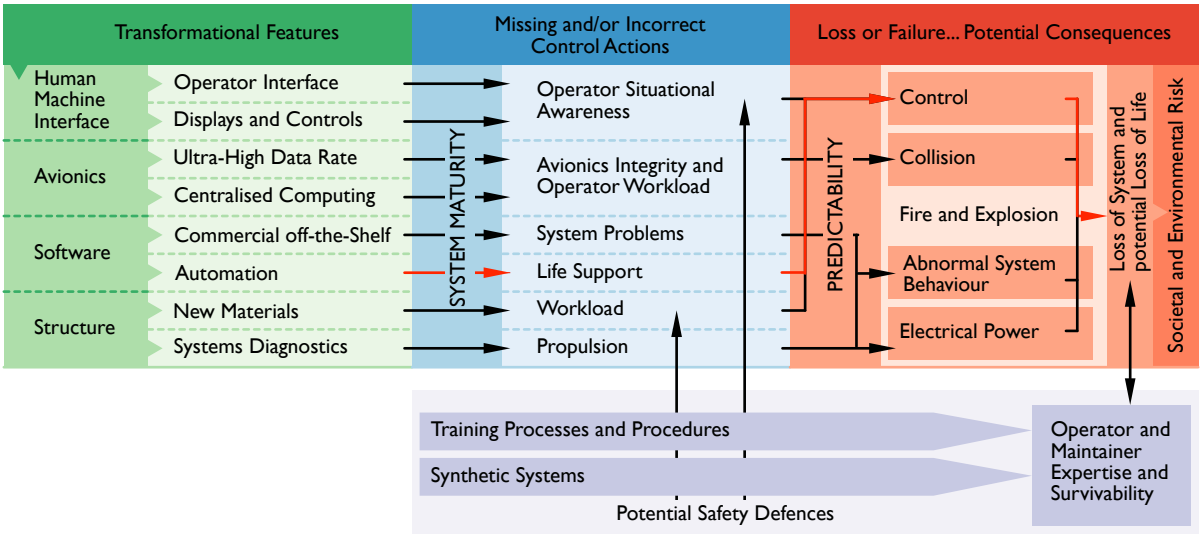

Figure 10: Transformational systems: whole-system threat analysis

The safety and effectiveness framework is able to provide a collective view of system shortfalls, whether considering safety or effectiveness, enabling consideration of these scenarios in both prevention or recovery phases of operation. Figure 11 illustrates High Risk Indicators (HRIs) 4, 8 and 11 against the 'Prevent' and 'Recover' operational phases. It should be

noted that, whist the HRIs implemented were based upon the US MIL-STD-882D, specific adaptations were made for the F-35 programme. Notably, there are differing alignments of the risk category (high to low) to the risk assessment value (1 to 15). These changes were made to embrace catastrophe where there was insufficient evidence to inform the probability of occurrence, such as insufficient testing and consequently insufficient evidence. Any event that could have a catastrophic outcome was important for the assurance case. The boundary of acceptability was identified as that point where it could be stated categorically that the likelihood of occurrence was improbable; for the UK F-35 programme this was HRI.11.

Using the F-35 as an example, Figure 11 illustrates how several intolerable (HRI.4) risks may be mitigated towards a more tolerable level of risk (HRI.8); these could be further reduced to a tolerable (improbable) status (HRI.11). HRI.11 is often, but not always, ALARP; pragmatically, the residual risk, and the level of intervention required to effect system change, may not always pass the risk versus benefit test. Thus, HRI.11 tended to be the final and acceptable status for many hazards. Each HRI.4 risk represents a catastrophic outcome, which in turn is potentially subject to further catastrophic outcomes during the recovery phase. The framework clearly illustrates where risks have been mitigated through the application of system constraints or operating limits (illustrated through an arrow to the right) and where the opportunity for further mitigation could be achieved through moving towards the adjacent box. An "(L)" represents where a system limitation would be applied to hold that level of HRI. This may be in addition to other measures such as additional (targeted) training. The asterisks, "*", link all contributors towards pilot workload challenges; these can be minor hardware (HW) or software (SW) issues or "target won't locate", for example and all tagged as HRI.8. The two-way arrows represent where either safe or effective operations could be compromised through incomplete or inappropriate training, leaving a degree of uncertainty concerning the effectiveness of any mitigation. This methodology identifies, therefore, where targeted training could be applied to achieve further certainty (or reduce uncertainty) as a means to achieve ALARP status.

The following will consider one example from each quadrant, commencing with 'Prevent-Safety'. Following high g and high-altitude flight testing, the pilot failed to get sufficient oxygen which was assessed as HRI.4. Follow-on altitude chamber testing reduced this to HRI.8 by limiting the maximum altitude at which the pilot can fly and further reduced it to HRI.11 by redesigning aspects of the oxygen system. Damage to the aircraft air-to-air refuelling nozzle ('Prevent-Effectiveness'), during nozzle engagement or fuel transfer, could lead to a large fuel discharge and ingestion into the engine inlet [19]. This was categorised as HRI.4, but can be reduced to HRI.8 by imposing an operational limitation, and further reduced to HRI.11 by a system fix, such as a redesign. Ejection OOE (out-of-envelope) represents a means of preserving pilot life and occurs in the recover phase (both the 'Recover-Safety' and 'Recover-Effectiveness' quadrants). At the time, ejection out-of-envelope was unproven and untested and so was categorised as HRI.4 but could be reduced to HRI.8 by pilot awareness and training not to eject within this envelope, and further reduced to HRI.11 by incorporating a system fix, which could be a redesign. It should be noted that ejection, back-up power and STOVL (short take-off, vertical landing) provide recognised means of effecting safe recovery. Furthermore, STOVL uniquely provides tactical effectiveness.

The fundamental measure of system reliability for the F-35 is the MFHBCF for each component or system. This was described by Gilmore [19, p. 60] together with the importance of measuring MFHBR and MFHBME by a reliability growth programme to achieve the required growth rate, as determined by the Duane Postulate. However, other outcomes have been experienced that are not necessarily (or obviously) related to a particular system where, during flight testing, unexpected emergent behaviours have occurred. Assessment and mitigation of the hazards described above represents a vital step in the journey towards identifying 'enterprise' improvements that will enable recovery of potential vulnerabilities towards a 'safe-

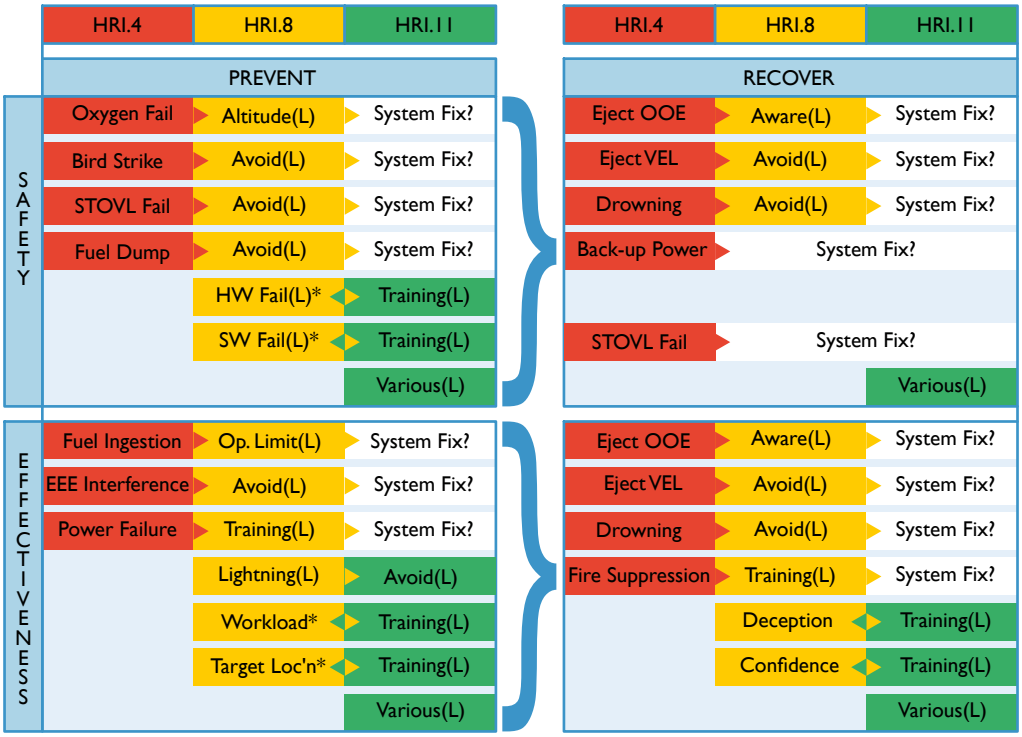

Figure 11: An example of translating system hazards into a foundation system assurance case (derived from Figure 4); EEE – electromagnetic environmental effects; OOE – out of [flight] envelope; VEL – velocity exceeds [survivable] limits

to-operate' condition. Given technical shortfalls in system design, there may be several options to mitigate adverse emergent behaviour. For example, applying modifications or limitations to certain systems to achieve known safe states, human factor interventions, and organisation and/or process changes. Requirements for long-term system fixes are also identified but these may introduce considerable cost.

The safety and effectiveness framework provides the basis for developing the system assurance case, for communicating the level of risk and rationale as well as providing the means for accurate and meaningful configuration control. Moreover, the framework enables discussion and development of rational, as well as prioritisation, for investment in system fixes requiring tough decision making when resources are scarce. Confident system assurance will be enabled through this simple and systematic approach to understanding the system and the associated system risks. The Prevent, Recover, Safety, Effectiveness quadrants offered by Figure 11, enable risks to be easily understood, explained and communicated at all levels of detail, minimising the potential for lack of clarity and promoting appropriate and informed discussion at all levels and across all disciplines. The approach described above, combined with analysing the outcomes with SQEP, has enabled the SESA approach to be successful in clearly expressing system operation limitations to levels that are predictable and where whole system operation delivers a more stable outcome.

4.2. The contribution of SESA across the supply chain

The critical success factor enabling the successful creation of the 4P enabling environment (Figure 2) is the development of shared objectives across the supply chain. In addition to these shared objectives, collaboration and appropriate behaviours (e.g. hazard seeking) with improved access to data across the supply chain leads to the development of positive skills and

more efficient management, operational and certification effort by all stakeholders, as demonstrated by the F-35 assurance case study [3]. The effort expended on the UK F-35 assurance would have been more effective had it been applied during the concept and design phases. The late introduction of this assurance methodology actually created the urgent need to seek a means of achieving system assurance through retrospective means. Ironically, this retrospective endeavour was the catalyst leading to the development of the SESA approach. Potential control actions that require attention need to be investigated, and a risk-management approach should then be introduced using the safety-effectiveness framework (Figure 4). This will identify trade-offs between shortfalls and the coordination of mitigation strategies to reduce or overcome the perceived risks. Furthermore, opportunities that trade safety and performance can be used for the strategic briefing of stakeholders seeking both understanding and consensus. Thereafter, the safety-effectiveness framework would serve as a detailed aid to balance whole-system safety and effectiveness. Having identified performance shortfalls, the design-to-operate framework (Figure 5) provides an opportunity to develop the relationship between the post-production system shortfalls and the full range of mitigations associated with introducing design modifications, discretionary training and/or organisational, management and procedural mitigation measures. The business cases for potential system design changes would be developed and discussed at both the working and strategic stakeholder levels by using the model to increase the level of whole-system understanding. A priority during the early phase of a new project is, therefore, to create an appropriate enabling environment, encouraging close cooperation between the designer and operator communities and focusing on the whole-system level. Figure 12 shows a SESA methodology maturity model with all the frameworks and models illustrated against a baseline project life cycle from concept to the in-service phase. The methodology has been set against a background of the 4P framework, which is intended to create the conditions for success by continually improving the contribution of the four pillars within the 4P assurance arrangements.

SESA was an original development and was introduced by the UK F-35B buying and certifying team. What SESA achieved was to identify during manufacture the key challenges that would affect certification, reliability, and military effectiveness. Although employed as a retrospective application to the UK F-35 assurance and certification process, SESA provided the supporting conditions that enabled initial assurance from July 2012 operating under military flight test procedures. This assurance arrangement successfully enabled a continuation of testing under restricted employment, developing operator and maintainer techniques, as well as responding to the incremental addition of F-35 capability until November 2017 when the UK F-35 was formally granted a Release to Service by the UK military regulator. Prior to the formal service release, there was insufficient confidence in the maturity of the air system and insufficient data to provide a certification assessment by the UK Regulator. The UK regulatory position is clear in that an aircraft must be compliant with DEF STAN 00-970 and its supporting standards; the F-35 programme was constructed on a conflicting basis, driven principally by the ambition to contain cost, giving Lockheed Martin (and its partners) the freedom to select appropriate technology and standards. This was at odds with traditional procurement programmes where customer direction is provided by mandating the standards to be employed.

In response to this, the UK team developed an *actual* versus US DOD versus UK MOD equivalence case, constructed retrospectively with the gaps in equivalence addressed critically for knowledge development and further assessment. SESA, in providing a methodology to constantly review both safety and effectiveness, enabled whole system assurance on a temporary basis whilst further system learning continued. Unfortunately, the UK regulator has no responsibilty for system capability or system effectiveness; the assumption is that if the specified design standard applies, the system will perform. During the assurance work the MFHBCF was well below expectations with the overall system availability suffering considerably. The technical data was immature, being either produced late or an emerging representation of the system in use. SESA was used

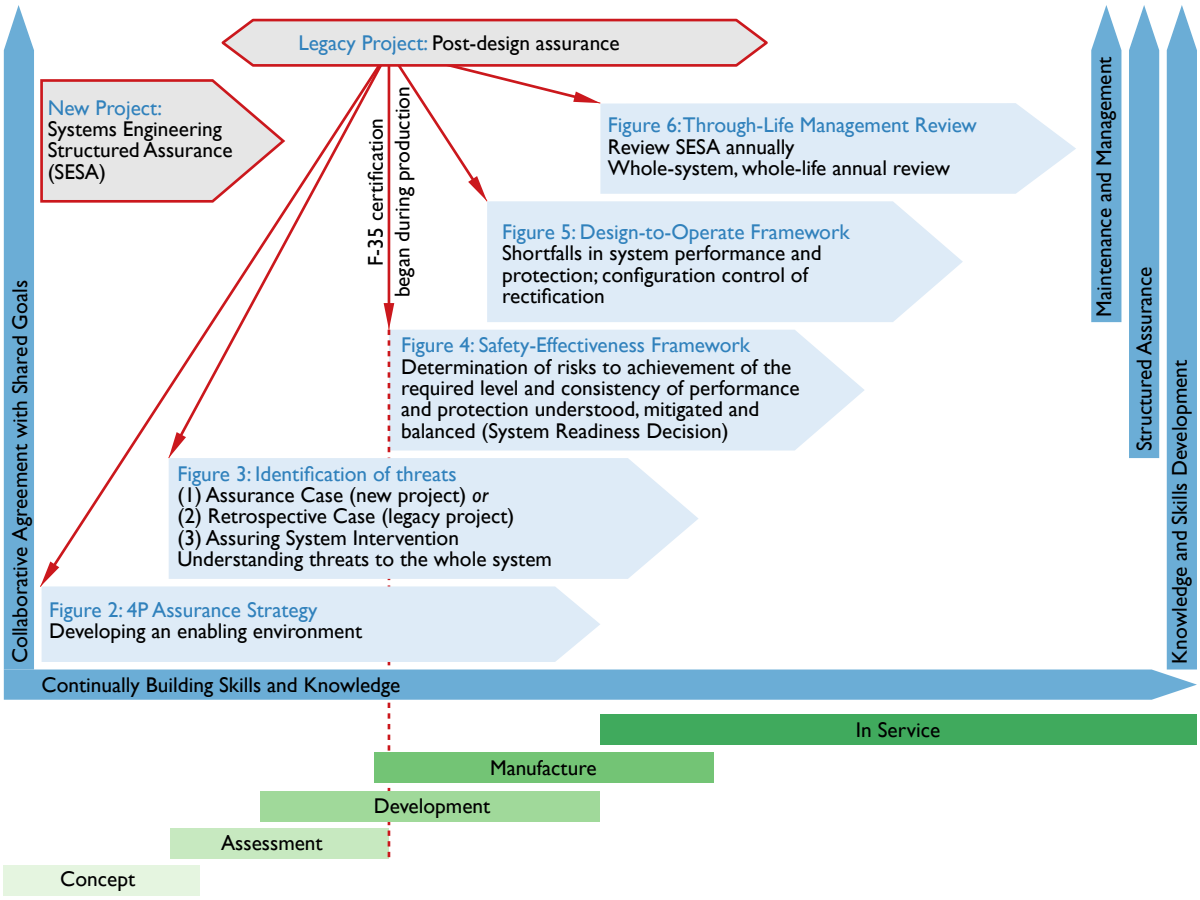

Figure 12: Systems engineering structured assurance (SESA) methodology maturity model

to identify and categorise the challenges, gathering all knowledge and expertise to develop system limitations in addition to identifying the best means of improving effectiveness across a variety of military tasks. The MFHBCF of the UK air systems faired well as the failure modes were encountered less often given the restrictions imposed. Hence, the MFHBCF of errant systems was effectively improved. Meaningful training was also developed to produce the required missions as part of the SESA programme.

The whole enterprise approach included the aircraft being fitted with comprehensive flight-test instrumentation thereby providing leading indicators to identify system degradation associated with potential safety-critical events as discussed above. These leading indicators could be adjusted as the system learning improved. Moreover, given that the aircraft was single seat, the SESA hazard analysis was employed to inform the scope of additional simulation training exercises utilising the SESA-determined hazards to be presented to the pilot [3]. Processes were aligned and tailored to both equipment and personnel maturity, where personnel included the pilot, maintenance operatives and engineering analysts.

The lessons identified during the application of the SESA methodology illustrate that it should be applied to the whole enterprise, embracing every aspect of equipment, people and process. We found that anything but an enterprise approach did not yield satisfactory results. The circumstances where SESA provides greatest benefit relate to the introduction of a complex system for first-time use, where the behaviour of the whole system may be unpredictable and uncertain. Experience with the F-35 illustrated that there was limited knowledge, experience and expertise, so developing these areas was critical for the UK.

The importance of maintaining the SESA discipline through life will represent a vital contribution in supporting the F-35's concurrent approach to development and the continued pressure of the production rate. Retrospective system fixes were, and will be, preserved for very special and rare cases; this required further investment and, as already explained, some system fixes may not be possible for many years. As a system fix is introduced careful monitoring is required for actual performance, unintended consequences, and emergent behaviour. Carefully tailoring the application of SESA as the enterprise matures, is vitally important for complex systems. Beginning the SESA effort at the design stage could introduce system performance and safety as bedfellows early enough to influence intended design.

The SESA framework provides a dedicated management arrangement that is persistent and easy to understand at every level. To deconstruct it requires a deliberate intervention including an explanation to the regulating authorities as to how decisions are taken to remove the controlling arrangements. The F-35 oxygen generation issue represents a single example of intervention across the whole range of safety-critical events. An operating restriction was identified as a mitigation alongside additional training measures, above and beyond that identified by the core F-35 programme. The SESA framework maintained the visibility of the threat whilst a long-term system intervention was planned; ideally, the introduction of an additional oxygen generation module to be installed within the system. This intervention would be delivered after a range of other system interventions to maximise effectiveness. The UK's involvement in the F-35 programme, exploiting SESA, has experienced little emergent behaviour; a controlled, constrained and understood operating environment, as a predictable and safe system of working, is the priority afforded through SESA.

5. Conclusions

In this article, we have described a methodology to achieve assurance in novel and complex systems. The literature review identified several issues that prevent the accurate prediction of system behaviour, with particular emphasis on poor system design and poor management (organisational drift), which is a key cause of accidents. Shortfalls in knowledge and practice also lead to inconsistencies in the approach adopted across the supply chain, where the supply chain is characterised as the designer, builder and user communities. The challenges associated with V-model testing and organisational drift were discussed, specifically their impact on the performance of complex systems and the associated time and cost implications. Rectifying the deficiencies introduced during V-model testing generates difficult business decisions, particularly where rectification is not cost effective and system shortfalls must be tolerated through-life. The literature also reported instances of practical drift in which systems were operated differently than intended, leading to potential system operating errors.

The acquisition of the F-35 stealth fighter by the UK authorities provided several challenges, given inconsistencies between US and UK socio-technical policies and in the arrangements necessary to address the variations in technical standards. These challenges, when combined with the need to access detailed system technical and test data retrospectively, alongside the numerous first-time technologies applied to the F-35 system, meant that significant effort was required to understand the system and thereby pave the way for certification. The F-35 case study achieved success through a collaborative approach, where the partners added value through the joint development of a common set of objectives representing system requirements and organisational ambitions. This approach creates an enabling environment, which is implemented most effectively through the 4P assurance strategy. Understanding the whole system and the threats associated with its introduction into the operational environment provides fundamental tenets for success. Following the adoption of appropriate enabling behaviours, the impact of the 4P framework should ideally be maintained and, if possible, improved. The 4P assurance strategy was applied to the

F-35 during the production phase of the life cycle (therefore representing a legacy application) that aimed to retrospectively address the design challenges. Clearly, it would have been more efficient and effective to introduce the 4P approach during the design phase, and early adoption of the SESA methodology is recommended.

Addressing system shortfalls and the associated recovery actions needed for performance and safety protection is best achieved by introducing an assurance methodology during the design and build phases and continuing its application through life. Moreover, the mapping of residual system problems and the parallel development of system fixes is achieved in an efficient manner using the safety and effectiveness framework. This offers the opportunity to plan, manage, and apply appropriate mitigation to the required implementation of system shortfalls. Thereafter, leadership frameworks can be used to identify, characterise and manage risk. This facilitates the development of control and coordination over the trade-space between performance and protection. The design-to-operate framework offers opportunities to both observe and coordinate the safety and effectiveness contributors at both the strategic and configuration control levels of application. The frameworks amalgamate into an overall 'systems engineering structured assurance' methodology providing a balanced approach for consistent applications to new systems or product designs. The development of the enabling environment and the identification of whole-system threats provide the first two steps towards balancing the requirements of the user community and the effort demanded of the design and build communities. The balancing of risk through the safety-effectiveness trade-off is a vital component that contributes towards a deep understanding of the system.

SESA methodology, when applied to a newly designed system, offers a means to optimise the relationship between stakeholders and improve the understanding of the characteristics and features of the systems engineering continuum. The US defence industries have cited shortfalls in system engineering expertise and a lack of whole-system technical leadership and knowledge. The SESA methodology discussed herein offers an opportunity to bridge the gaps in engineering expertise and transformational system knowledge by providing a systems leadership methodology.

References

References

[1] INCOSE. A world in motion: systems engineering vision 2025, 2014. URL `https://www.incose.org/docs/default-source/aboutse/se-vision-2025.pdf`. Accessed: 6 June 2018.

[2] K. J. Baldwin and D. S. Lucero. Defense system complexity: Engineering challenges and opportunities. *The ITEA Journal of Test and Evaluation*, 37:10–16, 2016.

[3] G. P. Farnell. *A new safety assurance methodology for complex systems*. PhD thesis, Cranfield University, Shrivenham, UK, 2017.

[4] A. Degani. *Taming HAL: Designing interfaces beyond 2001*. St. Martin's Press/Palgrave Macmillan, New York, 2004. ISBN: 031229574X.

[5] L. Steels. Towards a theory of emergent functionality. In J.-A. Meyer and S. W. Wilson, editors, *Proceedings of the first international conference on simulation of adaptive behavior on: From animals to animats*, pages 451–461. MIT Press, Cambridge, MA, 1990. ISBN: 0-262-63138-5.

[6] T. B. Sheridan and E. D. Nadler. A review of human-automation interaction failures and lessons learned. Technical Report DOT-VNTSC-NASA-06-01, US Department of Transportation Research and Innovative Technology Administration, Cambridge MA, October 2006. URL http://ntl.bts.gov/lib/34000/34700/34709/DOT-VNTSC-NASA-06-01.doc. Accessed: 4 October 2017.

[7] GAO. Joint Strike Fighter: Restructuring places program on firmer footing, but progress still lags. Technical Report GAO-11-325, United States Government Accountability Office, Washington DC, April 2011. URL http://www.gao.gov/assets/320/317548.pdf. Accessed: 4 October 2017.

[8] D. M. Zuckerman, P. Brown, and S. E. Nissen. Medical device recalls and the FDA approval process. *Arch. Intern. Med*, 171(11):1006–1011, June 2011. doi: 10.1001/archinternmed.2011.30.

[9] M. Crosse. Medical devices: FDA should enhance its oversight of recalls. Technical Report GAO-11-468, United States Government Accountability Office, Washington DC, June 2011. URL http://www.gao.gov/assets/320/319565.pdf. Accessed: 4 October 2017.

[10] ICAO. *Safety Management Manual (SMM)*. International Civil Aviation Organization, Montréal, Quebec, Canada, 3rd edition, 2013. ISBN 978-92-9249-214-4.

[11] CAIB. Report of Columbia Accident Investigation Board, Volume 1, 2003. URL https://www.nasa.gov/columbia/home/CAIB_Vol1.html. Accessed: 4 October 2017.

[12] S. W. A. Dekker. *Just Culture: Balancing Safety and Accountability*. Ashgate, Farnham, UK, 1st edition, 2007. ISBN 978-0754672661.

[13] K. Cowing. NASA reveals probable cause of Mars Polar Lander and Deep Space-2 mission failures, 2000. URL http://www.spaceref.com/news/viewnews.html?id=105. Accessed: 4 October 2017.

[14] D. Galin. *Software Quality Assurance: From Theory to Implementation*. Pearson/Addison Wesley, London, UK, 2004. ISBN 978-0201709452.

[15] B. Appleton. Technical debt – definition and resources, 2009. URL http://bradapp.blogspot.co.uk/2009/06/technical-debt-definition-and-resources.html. Accessed: 4 October 2017.

[16] P. A Judas and L. E. Prokop. A historical compilation of software metrics with applicability to NASA's Orion spacecraft flight software sizing. *Innovations Syst. Softw. Eng.*, 7(3):161–170, 2011. DOI 10.1007/s11334-011-0142-7.

[17] J. C. Mogul. Emergent (mis)behavior vs. complex software systems. Technical report, HP Laboratories, Palo Alto, CA, December 2005. URL http://www.hpl.hp.com/techreports/2006/HPL-2006-2.pdf. Accessed: 4 October 2017.

[18] N. G. Leveson. *Engineering a Safer World*. The MIT Press, Cambridge, MA, 2012. ISBN 978-0262016629.

[19] J. M. Gilmore. FY 2014 annual report. Technical report, The Office of the Director, Operational Test and Evaluation, Arlington, VA, 2014. URL http://www.dote.osd.mil/pub/reports/FY2014/pdf/other/2014DOTEAnnualReport.pdf.

[20] R. E. Ball. *The fundamentals of aircraft combat survivability: analysis and design*. AIAA, Reston, VA, 2nd edition, 2003. ISBN 978-1-56347-582-5.

[21] C. Perrow. *Normal accidents: living with high risk*. Princeton University Press, Princeton, NJ, 1999. ISBN 978-0691004129.

[22] E. Hollnagel, D. D. Woods, and N. G. Leveson, editors. *Resilience Engineering: Concepts and Precepts*. CRC Press, Boca Raton, FL, 2006. ISBN 978-0754649045.

[23] S. W. A. Dekker. *The Field Guide to Understanding Human Error*. Ashgate Publishing Ltd, Aldershot, UK, 2nd edition, 2006. ISBN 978-0754648253.

[24] P. Feiler, J. Goodenough, A. Gurfinkel, C. Weinstock, and L. Wrage. Four pillars for improving the quality of safety-critical software-reliant systems. Technical report, Software Engineering Institute, Carnegie-Mellon University, April 2013. URL `https://resources.sei.cmu.edu/asset_files/WhitePaper/2013_019_001_47803.pdf`. Accessed: 4 October 2017.

[25] W. Wagner and P. H. A. J. M. van Gelder. Applying RAMSSHEEP analysis for risk-driven maintenance. In R. D. J. M. Steenbergen, P. H. A. J. M. van Gelder, S. Miraglia, and A. C. W. M. Vrouwenvelder, editors, *Safety, Reliability and Risk Analaysis: Beyond the Horizon*. CRC Press, Boca Raton, FL, 2014. ISBN 978-1138001237.

[26] C. Haddon-Cave. The Nimrod Review: An independent review into the broader issues surrounding the loss of the RAF Nimrod MR2 aircraft XV230 in Afghanistan in 2006. Technical report, Her Majesty's Stationary Office, London, UK, 2009. ISBN 978-0102962659.

[27] HSE. *Reducing risks, protecting people: HSE's decision-making process*. Her Majesty's Stationery Office, 2001. URL `http://www.hse.gov.uk/risk/theory/r2p2.pdf`. ISBN 0717621510.

[28] N. G. Leveson. White paper on the use of safety cases in certification and regulation, 2011. URL `http://sunnyday.mit.edu/SafetyCases.pdf`. Accessed: 4 October 2017.

[29] R. Steinzor. Lessons from the North Sea: Should "safety cases" come to America? *B. C. Envtl. A . L. Rev.*, 38(2): 417–444, 2011. URL `http://lawdigitalcommons.bc.edu/ealr/vol38/iss2/10`.