The art of measuring nothing: The paradox of measuring safety in a changing civil aviation industry using traditional safety metrics

Eric Arne Lofquist
BI Norwegian Business School

This is the author’s final and accepted version of a manuscript published in Safety Science, vol 48 (2010) 10:1520-1529
DOI: 10.1016/j.ssci.2010.05.006

Elsevier allows the author to “voluntary post a personal manuscript version of the accepted article on the author’s personal or institutional website including an institutional repository”. In this case BI Norwegian Business School’s institutional repository, BI Brage.
The art of measuring nothing: The paradox of measuring safety in a changing civil aviation industry using traditional safety metrics

ERIC ARNE LOFQUIST
Norwegian School of Management, BI Bergen, Norway

ABSTRACT Measuring safety as an outcome variable within the ultra-safe civil aviation industry during periods of deliberate organizational change is a difficult, and often fruitless, task. Anticipating eroding safety processes, based on measuring nothing happening over time, does not adequately capture the true state of an evolving safe system, and this is particularly relevant for leaders and managers in a civil aviation industry responsible for maintaining and improving ultra-safe performance while simultaneously managing demanding strategic business goals.

In this paper, I will look at the difficulties of measuring safety as an outcome measure in high reliability organizations (HROs) using the traditional measures of incident and accident reporting during periods of deliberate organizational change inspired by the results from a three-year longitudinal case study of the Norwegian Air Navigation Services provider – Avinor.

I will first review the current safety literature relating to Safety Management Systems (SMSs) used in the civil aviation industry. I will then propose a more holistic model that shifts the focus from the traditional safety monitoring mechanisms of risk analysis and trial and error learning, to the natural interactivity within socio-technical systems as found in High Reliability Organizations.

And finally, I will present a summary of the empirical results of an alternate methodology for measuring perceived changes in safety at the operational level as leading indicators of evolving safety at the organizational level.

KEY WORDS: Organizational change; Safety Management Systems; High Reliability Organizations; perceptions of safety

Background

This paper focuses primarily on the problem of measuring safety as an outcome variable in an ultra-safe civil aviation industry during periods of organizational change using traditional aviation safety metrics inspired by a three year longitudinal case study of the Norwegian Air Navigation Services provider – Avinor (Lofquist, 2008). The study focused on the effects of deliberate large-scale organizational change, also known as strategic change, on safety outcomes within one High Reliability Organization (HRO). The strategic change process observed was undertaken as part of a corporatization initiative of air navigation services in Norway by the Norwegian government, and was to be accomplished without causing any disruption in customer services, and while simultaneously maintaining or improving safety performance. However, early in the study it became increasingly clear that the traditional safety metrics of measuring incidents and accidents did not reflect the negative effects of change that were being observed within the organization at the individual and unit levels, or even after the deliberate change process had suddenly collapsed in December 2005 (Lofquist, 2008). And these observations were confirmed by the Norwegian Transportation Safety Board study (HSLB, 2005) where no changes in
traditional safety metrics were detected leading to the conclusion that safety was unaffected by the change processes, at least not in the short run.

Accordingly, the purpose of this paper is twofold. The first is to address the problem of measuring safety outcomes in ultra-safe, high-risk industries during deliberate organizational change processes using traditional safety metrics. I will argue that the problems encountered in measuring safety as an outcome measure using traditional metrics in the Avinor case can be generalized both to other organizations within the civil aviation industry, and to other high-risk industries (Lofquist, 2008). The second purpose of this paper is to propose an expanded integrated safety model based upon current Safety Management Systems (SMSs) already widely used within the civil aviation industry. This model depicts safety as an emergent property of a complex socio-technical system embedded within an organizational culture in three temporal phases that goes beyond traditional reactive measures, and gives a more robust picture of evolving organizational safety.

Introduction

In the spring of 2005, The Norwegian Ministry of Transport (NMT) received disturbing indications from several sources within the civil aviation industry that safety margins within the Norwegian air transport sector had been significantly reduced as a direct result of both strategic and incremental changes being implemented by many of the key actors within the civil aviation industry. In response to these warnings, the Norwegian Ministry of Transport instructed the Norwegian Accident Investigation Board (HSLB) to conduct a review of the civil aviation industry, and to report back on the state of safety within the evolving Norwegian Air Transport sector (NMT letter of 7 Oct 2004). In the final report, entitled “Safety in Norwegian Civil Aviation during Change Processes” (HSLB, 2005), the study concluded that safety levels within the air transport sector remained “high”, but the report also qualified this conclusion with the following disclaimer:

“The generally high safety level, and the correspondingly low number of accidents and serious incidents, makes it difficult, if not impossible on a national level, to utilize accident statistics to ‘measure’ or prove that flight safety has become better or worse due to the prior years reorganization/changes. Research and experience from other countries show that eventual negative effects of flight safety-related consequences seldom materialize in the form of accidents for several years after changes are implemented. It is, therefore, necessary to use as a basis, other types of indicators to be able to evaluate how flight safety is evaluated.” (HSLB, p. 6)

These findings were reassuring to many, and particularly to the politicians and the leaders within the aviation industry responsible for safe outcomes, but the report also left significant doubt as to the real state of safety in the aggregated civil aviation industry undergoing continuous change where signs of system stress had been receiving daily media focus. One area of particular concern revealed through the study was the perception by those most closely associated with safe operations, specifically the pilots and air traffic controllers, that air safety quality had been noticeably reduced over the previous 5-year period as a direct result of organizational changes within the industry (TØI Report 2005, p. 42). This observation, though not fully reflected in the
final findings of the HLSB report, raised questions as to the true state of safety in an aggregated civil aviation industry experiencing deliberate organizational change that had not yet demonstrated a measurable decrease in safety levels, at least not by using the traditional safety metrics of incident and accident reporting. This also leads us to ask the question, “is safety an outcome in itself, or is safety an emergent quality of a complex system producing desired outcomes that are safe?” If the latter is true, do traditional measures give us an accurate, complete and timely representation of safety as an emergent quality of a complex process during periods of strategic change?

Corporatization

Since the late 1960’s, there has been a gradual deliberate liberalization of the international civil aviation industry through an initiative known as corporatization. The International Civil Aviation Organization (ICAO) describes corporatization as “creating a legal entity outside the government to manage airports and/or air navigation services, either through a specific statute or under an existing general statute, such as company law. Normally, ownership of the corporation remains with the government. However, in a corporatized body, private sector participation is possible. It depends upon the provisions of the statute under which it is established and the policy of the government.” (ICAO, 2002: p. 3.). In short, corporatization is a systematic approach for the privatization or semi-privatization of national civil aviation activities within the international civil aviation industry. The specific goal of corporatization is to make national civil aviation activities more competitive and cost effective, while simultaneously maintaining and/or improving upon ultra-safe levels of flight operations. This has placed an increased burden on the organizational leaders and managers who find themselves responsible for managing a new strategic business model in a high-risk environment with potentially conflicting goals. These leaders must now balance highly visible and verifiable financial performance results with less visible, and even less quantifiable, safety outcomes in an ultra-safe civil aviation industry where incidents and accidents are rare by design. This often leads to conflicting organizational goals requiring leaders to make hard business choices that can directly impact safety outcomes, over time. An impact, that often does not necessarily manifest itself in the form of incidents and accidents until long after the changes have progressed to the state where safety margins can no longer ensure safe outcomes, and can lead to disaster. These disastrous consequences were alluded to in the HLSB report, but also observed in the Überlingen mid-air collision that took place on 1 July 2002 over the town of Überlingen, Germany (Johnson, 2004).

Literature review

Currently, much of the safety literature on high-risk environments is based on grounded theory from high-profile accident investigations (Perrow, 1984, Shrivastava, 1987; Weick, 1993a.;1993b.; Vaughan, 1996; Snook, 2002; Gehman, 2003; Johnson, 2004) based on dramatic system failures where much of the emphasis is devoted to showing post facto the structural and behavioral causes and precursors of operating failure (La Porte 1996: p. 60). Bourrier (1998) argues that “too often, organizational analyses are carried out only after a catastrophe has occurred. While very interesting, this perspective has serious limitations and it is always easier to explain and reconstruct events after they have taken place” (P. 133). And the literature also reflects that there is a general lack of proactive safety theory derived from longitudinal case studies of “safety management systems in companies being subject to reorganization in the wide-
open business environment”, and is described as “a black hole in research and literature” (Hale et al., 1998: p. 11). But this trend also seems to be changing as the concepts of High Reliability Organizations and Resilience Engineering begin to dig deeper into system approaches to safety management in high-risk industries where accidents are rare by design, but potentially disastrous (Rochlin et al., 1987; Roberts, 1990; Weick and Roberts, 1993; Weick and Sutcliffe, 2001:2006; Reason, 1997; Hollnagel, et al., 2006). These types of studies, in particular, are highly relevant for the ultra-safe civil aviation industry that continues to evolve in its attempt to corporatize, and provides a unique opportunity to observe, and to gain new knowledge from, change processes in high-risk environments that can have potentially devastating consequences, prior to an actual disaster.

In the past, most studies in safety management have been conducted primarily by scientists in the fields of sociology, psychology and engineering, all with their own particular scientific paradigms, and all with their own specific approaches for defining and measuring safety and/or safety culture (Perrow, 1984, Shrivastava, 1987; Weick, 1993a.:1993b.; Vaughn, 1996; Snook, 2002; Gehman, 2003; Johnson, 2004) but without a great deal of interactivity between academic groupings. These same academic disciplines have also been involved in studying safety system behavior based on multiple case studies that have introduced new safety concepts, such as: man-made disasters (Turner, 1978), high reliability organizations (Rochlin et al., 1987; Roberts, 1990; Weick and Roberts, 1993; Weick and Sutcliffe, 2001:2006), organizational accidents (Reason, 1997), and resilience engineering (Hollnagel, et al., 2006). And although all of these academic contributions have increased our understanding of the underlying organizational dynamics of how safe systems contribute to unacceptable outcomes, all fall short of defining a true systems perspective for measuring safety as a process within high-risk industries that captures all of the essential parts of a robust operative safety management system. And this is particularly relevant when defining safety within an expanded strategic business context where safety is one, but not necessarily the only, priority affecting business decisions. Rasmussen (1994) has contributed with the notion of organizational drift towards accidents under economic competitive pressures. This describes the threat to safety outcomes in business settings, but leaves business leaders and managers responsible for managing these complex systems with little prescriptive guidance on how to design and support proactive management structures that can help detect and identify potentially unsafe conditions in their developmental phases. And, more importantly, how and when managers should take appropriate proactive corrective actions based on leading safety indicators that fall well outside of the traditional historical metrics of incident and accident reporting. And this is also true during periods of demanding deliberate organizational change within a complex and changing environment where developing latent conditions (Reason, 1990) might be masked by competing priorities.

One problem for the civil aviation industry is the excessive focus upon incidents and accidents as extraordinary events. Perrow (1984), for example, studied the Three Mile Island accident and found that accidents in complex, socio-technical systems are in reality normal outcomes within design specifications that are often unanticipated due to system characteristics, such as, interactive complexity and tight coupling. This is considered an important point in that by being normal outcomes within a system’s design parameters, accidents are, in fact, just undesired outcomes from an otherwise properly functioning system. And though the consequences of such undesired and unexpected events are often dramatic, their causes are similar to non-dramatic, undesired events that could have developed into disaster but did not. This often leads to artificial
“fixes” based on detected errors in the form of rules and regulations that attempt to limit the danger, and potential adverse consequences of undesired system outcomes, without really fixing the process that allows the undesired outcome to occur in the first place.

This requires an expanded focus on what constitutes a safe system that goes well beyond the overly simplistic elimination of “breakdowns and errors” (Amalberti, 2001), even when conceptual designs generate systems with high theoretical performance (p. 110). Turner (1978) showed us how some accidents are caused by sloppy management, and can be dealt with through different types of control, but that others are caused by disaster preconditions created by normally functioning managerial and technical systems through the formation of “incubation periods” (p.215.). This is also reflected in Reason’s (1997) accumulation of latent failures. Still other literature focuses on man-made disasters such as: Bhopal (Shrivastava, 1987), Three Mile Island (Perrow, 1984), the Tenerife air disaster (Weick, 1993a), the Challenger (Vaughn, 1996) and Columbia (Gehman, 2003) space shuttle accidents, the “friendly” downing of two Blackhawk helicopters (Snook, 2002), and the Überlingen mid-air collision (Johnson, 2004), where organizations play a significant contributing role leading to undesired outcomes. The space shuttle accidents, in particular, demonstrate both the importance of organizational soft variables, such as organizational culture in safety outcomes, as well as, the difficulties in changing culture within socio-technical organizations over time (Gehman, 2003).

Recently, there have been “a few studies which have started to use organizational learning theory to look at the introduction, development and integration of safety management systems but even these have, up to now, been very limited and have only occasionally taken a longitudinal perspective to explain how the change works and what aids or hinders it” (Hale et al., 1998: p. 10). This is emphasised by Baram (1998) who states that “such deep organizational changes in a company which uses hazardous technology can reduce the effectiveness of its process safety management system unless the implications of change are consciously and carefully addressed” (p. 191).

These studies indicate the need for a systems approach to studying safety management systems in high-risk industries that goes beyond trial and error learning (Weick, 1987), and actively integrates reactive, interactive and proactive safety measures from both organizational and individual perspectives. And by an overall systems perspective in civil aviation, I am referring specifically to a socio-technical system that is becoming increasingly business-oriented, and where the balance between human interaction and technical complexity is high. Weick (1987) describes the air traffic control system as a system that “seems to keep the human more actively in the loop of technology than is true for other systems in which reliability is a bigger problem, and where the air traffic controller uses qualities such as discretion, latitude, looseness, enactment, slack, improvisation and faith - work through human beings to increase reliability” (p. 122). This places a greater reliance upon human operational strengths and weaknesses, and is, in fact, the human contribution to resilience in overall system performance.

So I argue that to achieve a genuine systems perspective on safety that looks at all of the respective parts of a safety management system, in context, all of the previously mentioned academic disciplines are essential in contributing to a holistic system safety understanding. But that is not enough, as we also need interaction and involvement from the system operators and business leaders and managers responsible for both system performance, and for safety outcomes.
prior to undesired events. This requires a robust safety management system that is integrated into the overall strategic business objectives of an organization within an expanded industrial business context that can anticipate changes in an operative environment while balancing safety with economic goals. And, though the consequences of change have often been identified as significant contributing factors in accident investigations (Vaughn, 1996; Snook, 2002, Gehman, 2003; Johnson, 2004), little of the safety literature focuses specifically upon how deliberate change, in particular, affects safety over time. And, more importantly, how organizational leaders responsible for making time-sensitive decisions based on conflicting organizational priorities should react to eroding safe systems prior to disastrous outcomes.

Civil aviation as a complex “social-technical” system

The safety literature for high-risk industries is dominated by studies involving large machine-bureaucratic organizations, such as the nuclear power and chemical industries, often with a very considerable investment in sophisticated defense-in-depth management systems in safety and environment (Hale et al., 1998). Civil aviation, on the other hand, has generally received less focus, yet few industries can rival the growth, introduction of new technologies, and increase in complexity that the civil aviation industry has experienced over the past years, and will continue to experience in the future (Shin, 2005). The civil aviation industry can be described as a complex system of overlapping socio-technical systems embedded within a highly competitive business environment, where safety is a primary, but not the only, goal. This is particularly true as international air navigation service providers corporatize, and become more economically competitive and, correspondingly, more vulnerable to conflicts of prioritization in a changing business environment. But unlike the tightly coupled machine-bureaucratic systems described by Perrow (1994) that lead to normal accidents, the civil aviation industry can be characterized as a system that is highly complex but loosely coupled, and relies more heavily on human interaction than most machine-bureaucratic organizations, and is highly influenced by human variation. In fact, Weick goes so far as to say that “one striking property of air traffic control is that controllers are the technology” (1987: p. 120) meaning that human variation is essentially an integral part of both the system design and system performance. But there are also benefits in that humans are also flexible and can accommodate subtle environmental changes real-time, and can also detect latent conditions or failures (Reason, 1990) that develop gradually over time. Loose coupling within a complex system, similar to feedback with delay in system dynamics terms (Forrester, 1961; Sterman, 2000), increases the complexity in understanding how causal relationships interact and develop over time, and these complexities are often missed in mishap investigations that often focus on finding the most observable root causes. This is also true during periods of strategic change in dynamic environments where system outcomes are based on systems that have evolved away from the original system design but, either due to operator accommodation, regulation or built-in resilience, has not demonstrated measurable undesired safety outcomes using traditional metrics.

And these causal relationships become even more complex and difficult to identify when examining overlapping safety management systems where external actors can obscure or mask system deficiencies (Lofquist, 2008). The civil aviation industry is such an industry consisting of a system of interacting safety management systems, often described as High Reliability Organizations (Roberts, 1990; Schulman, 1993; Weick, et al., 1999) that provide overlapping
safeguards that can both prevent, or contribute to, disaster. The so-called safety nets (Johnson 2004) provided by these overlapping high reliability organizations, often lead to the development of latent conditions (Reason, 1984) that hide potential disaster in unexpected ways (Turner, 1976). Many authors have recognized that small events, that are not necessarily complex in themselves, can link together to create disproportionate and disastrous effects (Weick, 1993a; Perrow, 1987; Vaughn, 1996; Reason, 1997).

Measuring safety in civil aviation

This leads us to the real problem. Providing a meaningful measure of safety based primarily on disastrous outcomes is a difficult task considering that the likelihood of a disastrous accident today is approaching an ultra-safe level where the risk of such an event is currently below one accident per million events (Amalberti, 2001, p. 111). However, focussing on disastrous events alone, places too much weight on the magnitude of the consequences, and precludes unintended outcomes that could have, but did not, produce such devastating consequences. This problem was probably best summed up by James Reason (1990) who stated that “safety is defined and measured more by its absence than its presence.” And this was also reflected in the findings of the Norwegian Accident Investigation Board (HSLB, 2005) in their study of the Norwegian civil aviation industry where growing concern over aviation safety were not reflected in measurable changes in traditional safety metrics during the timeframe studied. But measuring safety is elusive because it is a dynamic non-event where a stable outcome is produced more by constant change rather than continuous repetition (Weick and Sutcliffe, 2001). Nothing to measure, at least by current industry metrics, equates to no change in safety, which is counterintuitive in a strategic change environment where there are many examples of organizational change contributing to a sudden, and often unexpected, system failure (Weick, 1993a, Vaughn, 1996; Gehman, 2003; Johnson, 2004). When accidents do occur, we have a measurable indication that things are not safe, but when nothing happens, or there is nothing to pay attention to (Weick, 1987; Weick and Sutcliffe, 2001), we do not know if this is due to properly functioning safety processes, or due to good fortune. One problem is the focus upon what organizations label as errors, which are often associated with visible or measurable non-acceptable consequences instead of what psychologists define as erroneous acts, whatever the consequences, or level at which they are detected or recovered (Amalberti, 2001). This subtle difference often masks the true state of an eroding system until defined unacceptable events occur, often with disastrous consequences.

Safety in three phases

I have argued that the traditional safety metrics for measuring safety in the civil aviation industry of reporting incidents and accidents, though important inputs, do not fully capture the true safety state of an evolving organization, or even the industry as a whole, and are, at best, lagging indicators. Svedung and Rasmussen (1996) describe errors and accidents as not particular, separable phenomena, but must be studied as being the effect of normal, adaptive behavior drifting toward the boundaries of acceptable performance. Amalberti (2001) argues that larger incident and accident databases do not increase accident prediction capability, and also points out that these databases have drifted away from the original intent of safety monitoring toward focusing on literary or technical causes of accidents (p. 113). In addition, safety studies have
found that the quality of defining and reporting incidents and accidents vary both from organization-to-organization, but also between units within organizations based on organizational culture, so that these measures are often unreliable, or at least difficult to defend, statistically (Cabrera and Isla, 1998; Pidgeon, 1997).

Instead, I propose that a more balanced approach to measuring safety in a complex system is represented by a safety management system that monitors safety as a process in three temporal phases embedded in an organizational culture. These phases, depicted in Figure 1, include: proactive, interactive and reactive mechanisms and measures that are both separated in time, and should include both endogenous and exogenous environmental factors. Such a system will include both leading and lagging indicators that are both quantitative and qualitative in nature, and include operator perceptions of evolving safety systems during change. This will provide a more balanced view of a system’s true safety state at any given point in time and give organizational leaders leading indications from which to take proactive measures. This approach is also compatible with current safety management systems (SMS) already in place within most civil aviation-related organizations, but as evidenced in the Überlingen example, not necessarily functioning as an integrated system. Figure 1 below demonstrates these interrelated mechanisms:

Insert Figure 1 here. (Integrated Safety Management Model)

**Proactive phase**

In the proactive phase, I refer to the mechanisms and measures utilized in system design, and re-design, covering the inception, introduction and life of an operational system. Designed levels of performance are based upon a specific set of operational assumptions and limited by technological and economic realities that may or may not reflect actual system performance, overtime. But these discrepancies are often unknown prior to the system becoming operational, and it has been shown that system design can, in fact, contribute to accidents as either root causes, or as contributory factors (Kinnersley & Roelen, 2007). And, in any case, these operational assumptions will evolve over time requiring continual evaluation, modification and fine-tuning over the operative life of the system. And as pointed out by Hollnagel (2006) “in complex systems, performance is always a variable, both because of the variability of the environment and the variability of the constituent subsystems” (p. 12), and it is the interaction of these endogenous and exogenous variables that create a gap in expectation and actual performance. Yet, this performance variability is necessary if a joint cognitive system, meaning human-machine system or a socio-technical system, is to successfully cope with the complexity of the real world (Hollnagel & Woods, 2005). In this respect, a systems approach to organizational design, which includes technical interconnectivity and human variation, is critical in designing, evaluating, maintaining, measuring, and re-designing safety processes in high-risk environments involving socio-technical systems. However, understanding the gap between desired performance and actual performance in a high-risk system cannot be accomplished by engineers alone, and requires an active interface between system designers and the system operators/managers in identifying these gaps. And I will argue that it is this active interface that is lacking in most safety management systems today, and is the area of focus for this study.
To achieve this active interface, there needs to be a robust, active learning exchange between designers, operators and system managers which are embedded within the organizational culture. This will ensure that the operators and managers are both knowledgeable about system design and expected performance parameters under defined operational constraints, as well as, the proper procedures for operating the system in practice, and reporting system discrepancies. This will also enable operators to properly identify system performance deviations or latent conditions (Reason, 1990) at an early stage of system development, and provide meaningful feedback for system correction, both in the interactive and reactive phases to be discussed in the following sections. This bridge between design and operations is illustrated in Figure 1.

Some important proactive measures for system design and redesign already exist, but are applied in safety management systems with varying levels of success, and the value of these measures depends heavily upon organizational culture. Risk analysis is one area of significant value for evaluating safety potential during change processes, and is a cornerstone of civil aviation safety management systems. However, one weakness of risk analysis is that the process is sometimes haphazard and incomplete in its application, and administered inconsistently due to competitive organizational pressures. In addition, risk analysis methods rarely capture all potential undesired or secondary outcomes due to system complexity and are, therefore, incomplete. And, in any case, risk analysis is only valid under the assumptions that are used in the analysis and are vulnerable to system evolution.

**Interactive phase**

The interactive phase is in many respects the most critical phase for high-risk organizations and the area with the greatest potential contribution to improving safety outcomes, particularly during deliberate change processes. The interactive phase includes all of the mechanisms and measures required for the real-time operation, control and maintenance of the designed system, and where interactions are heavily affected by both the internal and external environments, and human variability. This is where designed system performance meets reality with potential devastating results (Weick, 1993a; Vaughn, 1996; Gehman, 2003; Johnson, 2004). In civil aviation, for example, this is an area of particular challenge due to variation in human performance and overlapping operative systems where external actors often determine, or influence, outcomes as demonstrated in the Überlingen mid-air collision. The interactive phase is also where the organizational leadership has the least direct control over real-time outcomes but where organizational structure (Schein, 1985: 1990); leadership commitment (Zohar, 1980), and organizational culture (Meyerson and Martin, 1987; Pidgeon, 1997) contribute to flexible decision making during unexpected outcomes. Accordingly, this is an area where leaders and managers can initiate proactive initiatives to minimize, or even remove, potential latent conditions before they develop into failures as illustrated in Figure 1 (solid arrow). The interactive phase is where deviations between expected system performance and real outcomes are directly observable in their formulation stages, and also where imperfections in system design are first detected. This is particularly important when systems evolve away from original system design, or where local initiatives potentially conflict with overall system performance. But it is also here where discrepancies and undesired outcomes must be dealt with on a real-time basis, based on education, training, understanding, and experience, and is largely affected by the organizational culture of the organization, and where operators and managers take local action to minimize any adverse affects to the system, and the organization (Rochlin et al., 1987).
The interactive phase depends upon a supportive organizational culture to achieve and sustain high levels of safety performance over time, and is where the High Reliability Organization school of thought makes a significant contribution. This is particularly true in promoting flexible line decision-making and effective organizational learning (Weick, 1987) based on historical factors in context (Rochlin et al., 1987). The interactive phase depends upon the level at which safety and performance-related issues are both noticed and acted upon prior to undesired events taking place as described in the concept of mindfulness (Weick and Sutcliffe, 2006). And this also includes deviations that fall short of incident and accident reporting. However, Pidgeon argues that safety culture, at least in a safety climate context (Zohar, 1980), can be “critiqued as a reduction to a combination of administrative procedures and individual attitudes ... which is critically missing the shared characteristic of social organization and culture” (Pidgeon, 1997: p. 6). It is exactly this concept of shared characteristics of social organization that are most dominant in a High Reliability Organization (Rochlin et al, 1987; Roberts, 1990; Weick and Roberts, 1993). And high-risk organizations that do not limit the effects of undesired events cannot afford to learn by trial and error, where the consequences can be disastrous (Roberts, 1990).

Safety process measurement techniques in the interactive phase of a safety management system are relatively new, with varying levels of agreement on methods and effectiveness. The use of safety climate surveys as part of safety audits are gaining ground and becoming an important qualitative tool for evaluating changes in safety processes (Zohar, 1980, Ciavarelli, 2003). However, despite the relative lack of agreement of the value of these tools, it is clear that this is an area with tremendous potential for improving safety outcomes by detecting potential failures in their formative phase, or when systems evolve away for original design parameters.

Reactive phase

The reactive phase of an integrated safety system involves the mechanisms and measures for enabling organizational learning after an undesired event has occurred. This is often referred to as trial and error learning (Weick, 1987), and depending upon the potential consequences involved, is not acceptable in high-risk organizations where one error can collapse the system. System corrections, based on either individual or an accumulation of events, can either take the form of restrictive measures to prevent future occurrences of such events based on the current system (dashed arrow), or redesign of the system (solid arrow) to remove the potential problem. Both are depicted in Figure 1. Introduction of restrictive measures that do not remove the potential problem do not provide real safety improvements, and may only provide local improvement while causing adverse affects in other parts of the system. Initiatives that use trial and error learning to initiate real system changes (solid arrow) remove system errors permanently through system redesign, and provide real safety improvements. As mentioned earlier, the reactive portion of the model is the area of system performance that has traditionally received the most focus for academic studies, and will not be addressed further here.

A summary of the theoretical contributions of High Reliability Organizations and Resilience Engineering to an integrated Safety Measurement Systems are reflected in Table 1 below:
Methodology

The Avinor study was a three year longitudinal case study following a strategic change initiative called Take-Off 05. In the study, I used a mixed-method approach to triangulate three separate quantitative and qualitative data sets administered over a two year period to take a closer look at the effects of change on perceptions of safety in the interactive phase of the Safety Management System Model described earlier. The data sets were taken at two distinct points in time (just before the Take-Off 05 change process was initiated and at the mid-point of the change implementation process), and was aggregated at two different levels (four embedded cases and the entire Air Traffic Controller/ATC Assistant population). The reasoning for choosing two levels of aggregation was to reduce potential bias effects in the final safety measurement model to be presented below.

The first quantitative data set used in the study was a leadership questionnaire administered internally by Avinor at each of the four Air Traffic Control Centers (ATCCs) in Norway in December, 2002. This questionnaire was administered three months before the announcement of the Avinor corporatization process known as the Take-Off 05 project. This questionnaire was then repeated at the mid-point of the Take-Off 05 implementation process in Nov/Dec 2004 after one of the ATCs had been closed down (Trondheim) and operations merged with a second ATCC in northern Norway (Bodø). In addition, during the timeframe that the second questionnaire was administered only three of the original four ATCCs remained, and a decision had been taken by the Avinor leadership to close a second ATCC (Oslo) to reduce the number of operative ATCCs in Norway from three to two. This meant that the three remaining ATCCs in 2005 were experiencing three different phases of a common change process: two ATCCs had already been merged into one (Bodø), one ATCC knew that they would be closed (Oslo), and one ATCC knew that they would survive but would have to absorb the second closed unit (Stavanger).

Participation in the leadership questionnaire at the respective units were:

2002
• Trondheim ATCC (n=30/48 for 63%)
• Bodø ATCC (n=33/68 for 49%)
• Stavanger ATCC (n=30/52 for 58%)
• Oslo ATCC (n=62/120 for 52%)

2004
• Bodø ATCC (n=46/85 for 54%)
• Stavanger ATCC (n=35/50 for 70%)
• Oslo ATCC (n=64/103 for 62%)

The leadership questionnaire results, combined with observations made during the longitudinal case study, were then used as a basis to design a semi-structured interview protocol to further investigate changes in individual attitudes and perceptions of four separate latent concepts considered important for understanding effects of organizational change on safety perceptions at the individual level over time. The interviews focused specifically on perceptions and attitudes.
related to: leadership commitment to safety, safety culture, attitudes towards organizational change, and individual perceptions of safety within Avinor’s three remaining ATCCs in 2005. These interviews were conducted just after the second leadership survey results were published and the HSLB questionnaire results were made public. These interviews (10 at each site) were then transcribed into a common Norwegian language, and the responses were then individually coded using the NVIVO 7 software. The three remaining ATCCs were chosen as units of analysis as they were similar in size and function, and directly responsible for the safe and efficient flight operations covering 100% of Norway’s enroute civil aviation structure.

The results from the second leadership questionnaire were then compared to coded data from semi-structured interviews conducted at each of the remaining ATCCs in 2005 (Bodo, Oslo and Stavanger ATCCs) but also included former members of the Trondheim ATCC. The triangulated leadership survey data sets were then used to verify changes in individual attitudes toward change and perceptions of the leadership’s commitment to safety using within-case and across-case analyses at two distinct points in time.

The second quantitative data set used in the study was provided by the Norwegian Transportation Safety Board (HSLB) from the “Safety in Norwegian Civil Aviation during Change Processes” study (HSLB, 2005). I was allowed to participate in the formulation of the study parameters, and was granted access to the results by both the HSLB and the Norwegian Ministry of Transport. This data set focused on how attitudes toward change, safety climate, and perceptions of safety had been affected by organizational changes that had taken place within the Norwegian civil aviation industry between 2000-2005, and consisted of over 4000 responses from eight distinct groupings within the civil aviation industry, including: pilots, air traffic controllers and ATC assistants, cabin crew, ground crew, maintenance personnel, engineers, the regulatory agency, and aviation leaders (HSLB, 2005). For this study, only responses from the air traffic controllers and Air Traffic Control assistants were used to conduct a confirmatory factor analysis on a conceptual structural equation model using Lisrel 8.7 software (Jöreskog et al., 1999). The objective was to test hypotheses within a conceptual model describing relationships between the latent constructs mentioned above.

Findings

The results of the leadership questionnaire data were analyzed using within-case and across-case comparisons and supported with quotations from the coded interview data. The findings showed that individual reactions to the change process, and correspondingly, the perceptions of the local leadership and local safety climate, varied predictably across units depending upon the respective phase of the change process each unit experienced. These changes are depicted in Tables 2 and 3 below.

Insert Table 2 here. (Leadership group statistics)

Insert Table 3 here. (Organizational climate group statistics)
However, the individual perceptions of the top leadership’s commitment were significantly similar across all cases and indicated a dramatic decrease across all units regardless of the phase of the change process experienced.

Insert Table 4 here. (Top leader statistics)

**Conceptual Model**

The conceptual structural equation model used in the study was constructed using four latent concepts taken from the literature related to perceptions and attitudes toward change and safety, and reflected the findings from the literature review, the leadership questionnaires, and the semi-structured interviews, and is depicted in Figure 2 below. In the model, I chose to use ordinal data instead of continuous data to accommodate for potential bias problems related to skewness and kurtosis. This approach was used on the advice of Jöreskog (2004) where the order of the responses was meaningful, and the distance between responses was not particularly interesting.

Insert Figure 2 here. (Conceptual Safety Measurement Model)

The individual latent constructs were then defined using individual items from the HLSB questionnaire based on the literature. The results are depicted in Figure 3 below:

Insert Figure 3 here. (Conceptual Safety Measurement Model Results)

The results show that there are strong positive direct correlations between the latent independent constructs (perception of leadership commitment, attitude toward change, and safety climate) and the main dependent latent construct (perception of safety). The results also show a moderate indirect influence of individual perceptions of leadership commitment on perceptions of safety through the intermediate latent construct attitude toward change. The total positive effect on perception of safety by individual perceptions of leadership commitment is 0.72. When this is combined with the results of the internal leadership questionnaires which show a significant decrease in individual perceptions of leadership commitment across all cases, it is clear that the potential effect on individual perceptions of safety based on the results of the structural equation model were significant, and in a negative direction.

**Discussion**

From the data collected during the Avinor study, it was clear that individual perceptions of safety were significantly reduced during the Take-off 05 process. This is also reflected in the three data sets used in the study. Yet, no changes in the classic safety measures of incident and accident reporting were recorded during the period as verified by the Norwegian Transportation Safety Board.

I have argued that the safety literature is of little help in providing alternate means of measuring safety outcomes in high-risk industries that are also ultra safe during periods of deliberate organizational change. This is because most of the literature focuses either upon accident
causation or human failure after a catastrophic event, and not on safe system operations per se. This is influenced by the primary focus on accidents or near-accidents as abnormal events, and not the occurrence of undesired events as otherwise normal but unforeseen outcomes in properly functioning systems. This difference, though seemingly minor, significantly changes the way we look at system operations and system outcomes, particularly during system changes that lead to latent failures (Reason, 1990) that may provide important leading signals of impending disaster. Paying more attention to leading indicators or even the development of latent conditions that do not have immediate negative consequences can produce significant benefits for both understanding safe system operations and preventing systems from becoming unstable. But due to the current focus in civil aviation on incidents and accidents as primary safety measures, leading indicators signalling potential system deficiencies often go unnoticed, either due to a weak supporting organizational culture, or due to gradual acceptance of eroding safety conditions, also described as deviance (Vaughn, 1996). And artificial fixes, in the form of new rules and regulations, often only mask the underlying problem and do not correct system deficiencies that allow the failure in the first place. And though several positive initiatives for measuring certain aspects of organizational safety performance outside of traditional measures are emerging, these have not, as of yet, been combined into an integrated safety evaluation methodology. And more importantly, safety has not been studied as one aggregate measure consisting of three temporal phases within a larger strategic business system where safety competes with other organizational outcomes.

Conclusions and implications

This paper has addressed the paradox of measuring nothing happening as an indication of system safety as an outcome measure during organizational change processes in the ultra safe civil aviation industry. I have shown that this is, in fact, a well-documented problem within the safety literature (Weick, 1993a; Johnson, 2004), and has been observed by numerous studies of disastrous events in other social-technical industries as well (Perrow, 1984, Shrivastava, 1987; Weick, 1993b.; Vaughn, 1996; Snook, 2002; Gehman, 2003), and is reflected in grounded theory from these events. But these studies, for the most part, focus on historical events, and it has been shown that there is a general lack of studies in vivo addressing the relationships between organizational change and safety outcomes due to decision-making processes, particularly in high-risk organizations that are also business-oriented.

The literature reflects several academic schools of thought that are contributing significant new understanding to expanded areas of safety management study at both the individual and the organizational levels. And although these different schools of thought on safety are considered significant contributions to safety management that fill a gap between the lagging indications of traditional trial and error learning processes, and a more proactive approach of using risk analysis and other methods to detect potential safety implications of change, neither provides a satisfactory, holistic view of safety measurement for strategic leaders to use in decision-making. However, I also believe that a robust integrated safety management system that includes proactive, interactive and reactive measures using a combination of approaches in an integrated model provides a much better evaluation by which organizational leaders can make better informed choices in business management in high risk industries that are also ultra safe.
I have argued that though proactive and reactive measures provide some level of system safety understanding, it is the interactive phase of system operations where system safety is affected most in social-technical systems, and it is also in this area where safety has the greatest potential for improvement. Bridging the gap between the interactive phase, or more accurately the real-time operations of a socio-technical system, to both the proactive and reactive phases is critical for improving system performance from a safety perspective, but also gives organizational leaders timely information and understanding of system performance during changing conditions that allows them to take more informed proactive decisions before undesired events take place.

Though the integrated concept is not unique, incorporating the many supportive concepts into a holistic, functioning model that provides early detection of potential system weaknesses during change processes has received little study. And although safety management systems are currently in place throughout the civil aviation industry, the literature shows that what exists on paper is often not supported by underlying mechanisms, particular those connected to organizational culture and, more specifically, safety culture and climate (Johnson, 2004). And even though many safety tools are already available off-the-shelf such as: safety audits and evaluation techniques, a combination of mandatory and voluntary reporting systems, enforced recommendations, protocols, and rules and significant changes in the governance of systems and in corporate safety cultures (Amalberti, 2006: p. 254.) little research reflects how these tools are contributing to improving safety during change. And Kirwan (1998) adds that “the problem with just focusing on the policy and procedures, i.e. the ‘paper system’, is that whilst this is necessary, it is not sufficient. This paper system may represent the ‘designed intent’ of the safety management system but does not necessarily represent its operational realization” (p. 68). And Kirwan also warns that “a good safety management system is necessary but not sufficient as there must also be commitment (from leadership) for implementation and application of sound safety management practice” (p. 74).

Follow-on studies need to focus on how organizations approach and implement organizational change in high-risk industries, as well as, how these changes affect safety outcomes short of disastrous events. And this is particularly true in ultra-safe environment were incidents and accidents are rare by design.
References


ICAO, 2002 International Civil Aviation Organization Meeting of Directors of Civil Aviation of Caribbean Region (CAR/DCA/1) (Grand Cayman, Cayman Islands, 8-11 October 2002)


Yerkes, R. M., and J. D. Dodson 1908 “The relation of strength of stimulus to rapidity of habit formation.” Journal of Comparative Neurological Psychology, 18:459-482

Table 1 - Theoretical contributions to Safety Management Systems

<table>
<thead>
<tr>
<th>Theoretical background</th>
<th>Specialization</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Reliability Organizations</strong></td>
<td>HRO structures</td>
<td>Rochlin et al., 1987</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Roberts, 1990</td>
</tr>
<tr>
<td><strong>Learning environments</strong></td>
<td></td>
<td>Weick and Roberts, 1993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Schulman, 1993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>La Porte, 1996</td>
</tr>
<tr>
<td><strong>Mindfulness</strong></td>
<td></td>
<td>Weick, 1995</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weick et al., 1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weick and Sutcliffe, 2001:2006</td>
</tr>
<tr>
<td><strong>Proactive Monitoring</strong></td>
<td></td>
<td>Ciavarelli, 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ciavarelli and Crowson, 2004</td>
</tr>
<tr>
<td><strong>Resilience Engineering</strong></td>
<td>Resilience</td>
<td>Hollnagel et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hale et al., 1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Amalberti, 2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dekker, 2006</td>
</tr>
</tbody>
</table>
Emergent systems
Woods, 2006
Dekker, 2006
Safety critical systems
Levesen et al., 2006
Managerial resilience
Flin, 2006
Sundström and Hollnagel, 2006

Table 2 – Leadership group statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Leader motivates</td>
<td>5.00</td>
<td>5.00</td>
<td>3.67</td>
<td>3.97</td>
<td>3.30</td>
<td>6.02</td>
<td>5.57</td>
</tr>
<tr>
<td>Leader discussion</td>
<td>5.38</td>
<td>4.81</td>
<td>3.87</td>
<td>2.97</td>
<td>3.62</td>
<td>6.38</td>
<td>5.54</td>
</tr>
<tr>
<td>Leader distributes</td>
<td>5.21</td>
<td>4.64</td>
<td>3.98</td>
<td>3.53</td>
<td>3.71</td>
<td>6.28</td>
<td>5.69</td>
</tr>
<tr>
<td>information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leader unity/commitment</td>
<td>5.03</td>
<td>4.34</td>
<td>3.56</td>
<td>3.37</td>
<td>3.20</td>
<td>6.37</td>
<td>6.06</td>
</tr>
<tr>
<td>Trust in leader</td>
<td>6.20</td>
<td>5.84</td>
<td>4.26</td>
<td>4.37</td>
<td>4.17</td>
<td>6.64</td>
<td>6.46</td>
</tr>
<tr>
<td>Leader performance</td>
<td>3.38</td>
<td>3.48</td>
<td>3.54</td>
<td>2.37</td>
<td>1.48</td>
<td>1.13</td>
<td>1.89</td>
</tr>
</tbody>
</table>

Table 3 – Organizational climate group statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal conflict</td>
<td>2.03</td>
<td>2.64</td>
<td>3.35</td>
<td>4.80</td>
<td>4.69</td>
<td>2.60</td>
<td>2.69</td>
</tr>
<tr>
<td>Work conditions</td>
<td>1.77</td>
<td>1.50</td>
<td>1.52</td>
<td>2.03</td>
<td>3.65</td>
<td>2.45</td>
<td>1.56</td>
</tr>
<tr>
<td>Cooperation problems</td>
<td>1.72</td>
<td>2.16</td>
<td>3.11</td>
<td>4.17</td>
<td>4.91</td>
<td>1.94</td>
<td>2.00</td>
</tr>
<tr>
<td>Employees harassed</td>
<td>1.45</td>
<td>1.94</td>
<td>2.81</td>
<td>4.63</td>
<td>3.48</td>
<td>1.41</td>
<td>2.74</td>
</tr>
<tr>
<td>Working environment</td>
<td>2.57</td>
<td>2.38</td>
<td>3.35</td>
<td>5.33</td>
<td>5.29</td>
<td>4.28</td>
<td>3.13</td>
</tr>
<tr>
<td>Power struggles</td>
<td>1.62</td>
<td>2.06</td>
<td>3.49</td>
<td>3.53</td>
<td>4.00</td>
<td>2.08</td>
<td>1.97</td>
</tr>
<tr>
<td>Poor motivation</td>
<td>2.87</td>
<td>2.30</td>
<td>3.00</td>
<td>3.33</td>
<td>4.26</td>
<td>5.16</td>
<td>2.46</td>
</tr>
<tr>
<td>Job burden</td>
<td>2.70</td>
<td>3.52</td>
<td>2.76</td>
<td>3.80</td>
<td>5.11</td>
<td>4.66</td>
<td>3.49</td>
</tr>
</tbody>
</table>
Private life

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low work morale</td>
<td>2.03</td>
<td>2.33</td>
<td>3.21</td>
<td>4.53</td>
<td>4.00</td>
<td>3.37</td>
<td>2.74</td>
</tr>
<tr>
<td>Different opinions</td>
<td>2.52</td>
<td>3.59</td>
<td>3.81</td>
<td>4.14</td>
<td>3.98</td>
<td>2.35</td>
<td>2.80</td>
</tr>
<tr>
<td>Pride in Avinor</td>
<td>4.67</td>
<td>5.38</td>
<td>4.46</td>
<td>4.00</td>
<td>2.57</td>
<td>1.56</td>
<td>2.94</td>
</tr>
</tbody>
</table>

Table 4 - Top Leader Statistics

Figure 2 – Conceptual measurement model

Figure 3 – Conceptual measurement model results
Leadership Commitment

Safety Climate

Attitudes Change

Perceptions Safety

RMSEA = 0.030