# ‘Systemic Failures’ and ‘Human Error’ in US NTSB and Canadian TSB Aviation Reports Between 1996 and 2003

C.W. Johnson; Dept of Computing Science, Univ. of Glasgow, Scotland, G12 9QQ.

C.M. Holloway NASA Langley Research Center, Hampton, VA 23681-2199, USA

Keywords: Causal Analysis, Accident Investigation, and Accident Analysis.

## Abstract

This paper describes the results of an independent analysis of the probable and contributory causes of aviation accidents in both the United States and Canada between 1996 and 2003. The purpose of the study was to assess the comparative frequency of a range of causal factors in the reporting of these adverse events. Our results suggest that the majority of these high consequence accidents were attributed to human error. A large number of reports also mentioned wider systemic issues, including the managerial and regulatory context of aviation operations. These issues are more likely to appear as contributory rather than primary causes in both sets of accident reports. The relative prominence of ‘systemic causes’ seems to have declined in recent years; this contradicts recent assertions about the importance of the latent rather than proximal causes of adverse events.

## Introduction

It has often been argued that investigators must look beyond the immediate causes of an accident to identify the underlying or ‘systemic’ factors that create the conditions in which a mishap is likely to occur. For example, Reason (Ref. 1) distinguishes the ‘person’ from the ‘system’ approach to accident analysis. Each of these perspectives implies a radically different view of causation. The ‘person approach focuses on the errors of individuals, blaming them for forgetfulness, inattention, or moral weakness’. In contrast, the system approach ‘concentrates on the conditions under which individuals work and tries to build defences to avert errors or mitigate their effects’. Similarly, Cook and Woods (Ref. 2) build on the earlier work of Mackie (Ref. 3) when they argue that accidents occur through the concatenation of multiple small failures. Each of these causes is necessary. However, they are each insufficient to cause the failure unless they occur in combination with other potential causes. Often these small failures have roots that extend well back from the moment when the accident is triggered. Cook and Woods are careful to distinguish between the operators who often trigger an incident ‘at the sharp end’ and the managers and regulators who often create the latent conditions for a failure ‘at the blunt end’. Managerial and regulatory problems make it possible for combinations of minor failures to build up over time and hence create the preconditions for failure.

It can be difficult to identify precisely which factors play a significant role in the latent causes of an accident or incident(Ref. 4). For example, the operational pressures of their everyday tasks may influence operator behaviour. The causes of these pressures can be traced back to particular management decisions distributed throughout the tiers of responsibility within a company. Often the systemic causes of adverse event will ultimately lead to the regulatory authorities and certification bodies that help to create the environment in which a management board operates. The proponents of the ‘systemic’ view can reasonably argue that regulators must ultimately bare responsibility for accidents in the industries that they regulate. However, this ignores the legislative and political constraints that limit the regulators’ scope for intervention. Similarly, it is important to question whether or not upper-levels of management can reasonably be expected to understand the detailed working practices that characterise the everyday operation of complex technology. For instance, previous studies of adverse events, including the Bristol Infirmary failures and the Challenger accident, have shown that middle and junior levels of management often find it difficult to pass bad news to their more senior colleagues (Ref. 4).

A number of senior executives have recently recognised the 'systemic' causes of adverse events. However, these insights have been used in ways that were not envisaged by the original proponents of this approach. Senior managers have argued that we must look beyond the immediate decisions of senior management to understand the latent factors that influenced their actions. Many managers have begun to use the language of 'systems failure' to defend their decisions. For instance, Joseph Berardino, the CEO at Arthur Andersen during the Enron collapse stated to the House Financial Services Committee that "the whole system needs to be looked at…it should be a felony to lie to or withhold facts from an accounting firm…I'm embarrassed by what happened at my firm" (Ref. 5). Similarly, a former NASA administrator summarised the causes of the Mars Surveyor mishaps by saying "in my effort to empower people, I pushed too hard... It wasn't intentional. It wasn't malicious. I believed in the vision... but it may have made failure inevitable…the team did not fail alone. As the head of NASA, I accept the responsibility. If anything, the system failed them." (Ref. 6). The administrator was prepared to accept person responsibility for the mission failures and yet at the same time he argued that the system failed them.

The regulations that govern the work of most accident investigation agencies do not explicitly distinguish between the 'person' or the 'system' view of failure. For example, the Code of Federal Regulations Title 49 Part 845.40 requires that the US National Transportation Safety Board (NTSB) "set forth the facts, conditions and circumstances relating to the accident and the probable cause thereof, along with any appropriate recommendations formulated on the basis of the investigation"**.** The NTSB's annual statistical reports provide further guidance on the nature of causes and contributory factors "the objective…is to discern the cause-and-effect relationships in the accident sequence. This could be described as *why* the accident happened. In determining probable cause of an accident, the Safety Board considers all facts, conditions, and circumstances. Within each accident occurrence, any information that contributes to the explanation of that event is identified as a finding and may be further designated as either a cause or factor. The term factor is used to describe situations or circumstances that contributed to the accident cause. The details of probable cause are coded as the combination of all causes, factors and findings associated with the accident. Just as accidents often include a series of events, the reason why those events led to an accident may be the combination of multiple causes and factors. For this reason, a single accident report can include multiple cause(s)" (Ref. 7). Under the Canadian Transportation Accident Investigation and Safety Board Act, 1989, c. 3, the Transport Safety Board (TSB) must identify "causes and contributing factors" to identify "safety deficiencies as evidenced by transportation occurrences". It is not the function of the Board "to assign fault or determine civil or criminal liability, but the Board shall not refrain from fully reporting on the causes and contributing factors merely because fault or liability might be inferred from the Board's findings.

Even though these guidelines fail to distinguish between Reason's 'person' and 'system' approaches, many investigators are aware of the importance of the systemic causes of adverse events (Ref. 4). However, a number of factors can prevent them from exploring the wide range of minor failures that together combine to create the preconditions for adverse events. In particular, resource constraints limit the scope of many investigations. Most investigation agencies operate with a relatively small core staff. They rely on external support to provide additional expertise. There are inevitable shortages of skilled personnel in several key areas, including software forensics. Further problems are created by the lack of recognised analytical techniques that might be used to guide and validate the 'systemic' analysis of adverse events. From this it follows that it can be difficult to determine whether or not investigators have considered an adequate range of causal factors during any particular investigation.

A number of leading accident investigators have written on the importance of 'systemic' factors in the causes of adverse events. For example, Strauch (Ref. 8) argues that the 'transformation of error perspective' from blaming the operator to identifying the contribution of system elements 'has, I believe, led to profound changes in the way we investigate, consider, and respond to accidents'. Similarly, Ayeko (Ref. 9) has argued that 'to learn a lesson from an accident one must understand not only the immediate cause but also contributing factors and underlying conditions of the accident'. He goes on to state 'it is my belief that, when we seek "cause" rather than "information about cause" in an investigation of an accident, the direction of the investigation often veers towards elements that are more likely to be linked to blame rather than the mitigation of risks'. The following pages determine whether these 'systemic' views of complex, technological failure have had a discernable impact on the work of accident investigation agencies.

## Method

It can be difficult to measure the impact that a particular view of accident causation has upon the working practices of an investigatory organisation. The views of leading investigators need not be shared by all of their colleagues. Similarly, it can be difficult to identify methods that might be used to support the application of a 'systems' or 'person' view to particular investigations. There are further methodological concerns. In particular, most investigatory organisations analysed a range of causal factors well before authors such as Perrow (Ref. 10) and Reason (Ref. 1) articulated the 'systems view'. It can also be argued that these authors provided their greatest service in developing a vocabulary to describe changes that were already slowly affecting investigatory practice. It is likely, therefore, that the impact of 'systemic' ideas can only be measured in terms of a relative change in the scope of any analysis rather than a dramatic or sudden change in investigatory practices. Finally, it is difficult to know what to *measure* in order to determine whether there has been any movement from the 'person' view to the 'systems' view of adverse events. Investigatory agencies such as the US NTSB and the Canadian TSB are responsible for several different modes of transportation ranging from pipelines through to aviation. The subsequent investigations produce many different forms of report that range from brief synopses of near-misses through to sustained documents on major accidents over hundreds of pages. It is unclear how one would identify the impact of a 'systems' view on such large and complex reports or across the mass of shorter documents on lower severity mishaps.

The method that was adopted in this study involved the two co-authors performing an independent analysis of all of the major aviation accident reports published between 1996 and 2003 by the US NTSB and the Canadian TSB. The investigators each had more than a decade's experience in the development of safety-critical systems. Each has been active in the analysis of system failures for more than five years. Aviation was chosen because this transportation mode is widely perceived to drive much of the innovation in accident investigation (Ref. 4). The start date was determined by pragmatism. It was felt that this provided a sufficiently large sample to support our analysis within the time available for our study. The sample yielded a total of 27 accident reports from the TSB and 26 from the NTSB. TSB documents relating to 2003 were unavailable at the time of the study. Hence our analysis of their investigations was based on documents from 1996 to 2002. The reports ranged from high profile, multiple fatality accidents such as the loss of ValuJet Flight 592 through to less severe loss-of-separation incidents. Our sample is relatively small compared to the 1,812 accidents and 1,374 incidents that were reported to the Canadian TSB in 2002. A considerable process of filtering was used by the investigatory agencies to select the most serious of these incidents for investigation. In consequence, our sample focuses on those higher risk mishaps, including near misses, which were deemed serious enough to warrant a subsequent investigation and report.

The analysis progressed by extracting the causal and contributory factors that were identified in the aftermath of each investigation. This preprocessing stage was necessary to insure that each of the analysts focused on the same source prose, given that many of the documents were hundreds of pages in length. The identification of all relevant sections in each report was performed as a collaborative activity between the two analysts. Before July 1996, NTSB reports grouped this information in a section labeled 'Probable Cause'. After that date, we used the probable and contributory causes that were explicitly listed in the abstract of each report. For example, the opening sections of the NTSB's 2000 report into the Guam crash contains the following analysis:

> "The National Transportation Safety Board determines that the probable cause of the Korean Air flight 801 accident was the captain's failure to adequately brief and execute the non-precision approach and the first officer's and flight engineer's failure to effectively monitor and cross-check the captain's execution of the approach. Contributing to these failures were the captain's fatigue and Korean Air's inadequate flight crew training. Contributing to the accident was the Federal Aviation Administration's (FAA) intentional inhibition of the minimum safe altitude warning system (MSAW) at Guam and the agency's failure to adequately manage the system". (NTSB AAR-00/01)

Canadian TSB reports also contain a section in their abstract that lists 'Findings As To Causes and Contributing Factors'. The analysis was less straightforward, however, because these documents did not distinguish between what was a cause and what was a contributing factor. For example, report A02F0069 lists the following observations in this section of the analysis:

1. The pilot not flying (PNF) inadvertently entered an erroneous $V_1$ speed into the MCDU. The error was not detected by either flight crew, despite numerous opportunities.

2. The PNF called "rotate" about 25 knots below the calculated and posted rotation speed.

3. The pilot flying (PF) initiated rotation 24 knots below the calculated and posted rotation speed and the tail of the aircraft struck the runway surface.

4. A glide path signal was most probably distorted by a taxiing aircraft and provided erroneous information to the autopilot, resulting in a pitch-up event. The pitch-up could have been minimized if the autopilot had been disconnected earlier by the PF. (Canadian TSB A02F0069)

Each analyst, therefore, had to distinguish probable causes from contributory factors in TSB reports so that some comparison could be made with the findings listed by the NTSB. The two investigators performed this analysis independently. All subsequent stages were also performed in isolation from each other until the results were available for comparison. We then went on to assign each of the probable causes and contributory factors to a number of common categories. We decided not to use any pre-defined taxonomy but to allow each of the investigators to independently assign their own terms to each of the 'causes'. The results of this process were then collated. There were some obvious differences in the terms used but there were also strong similarities. For instance, one analyst identified 'human error' while another distinguished between ‘aircrew error’, ‘ATM error’ and so on. Where such disagreements occurred we used a process of discussion to agree on a common term to support comparisons between the classifications. For example, we agreed to use the more general term ‘human error’. The term ‘ATM failure’ was used instead of ‘ATM error’ because it was often unclear whether a particular cause or contributory factor could be associated with the manager’s actions or with design problems in their information systems. Distinctions were preserved between different terms where no agreement could be reached between the two analysts.

## US NTSB Results

Tables 1 and 2 summarize the results of this classification process for both the probable causes and the contributory factors in the NTSB reports. The 26 incidents yielded a total of 40 probable causes for the first analyst. The mean number of probable causes was 1.5 with a standard deviation of 1. The second analyst also identified 40 probable causes with a mean of 1.5 and a standard deviation of 0.8. There were 52 contributory causes identified by the first analyst with a mean of 1.9 and a standard deviation of 1.5. The second analyst identified 55 contributory causes with a mean of 2 and a standard deviation of 1.5. The mode over all probable causes was 1 while the mode for all contributory causes was 2.

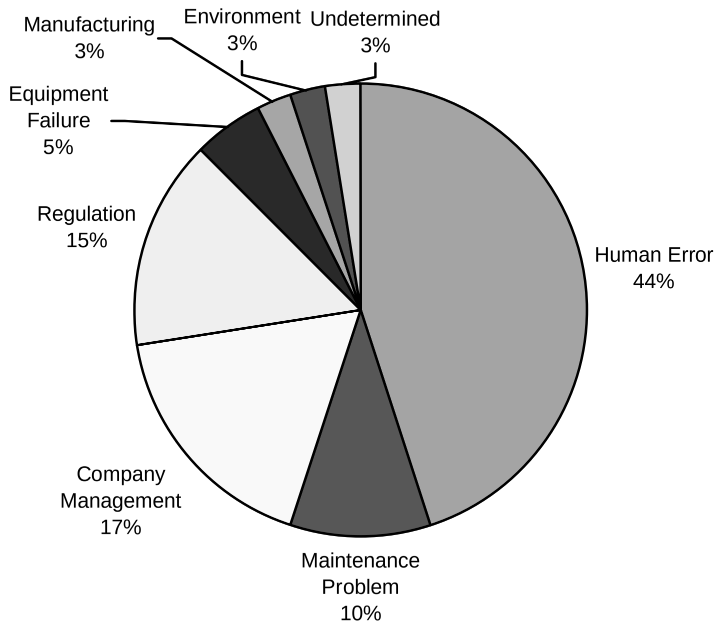
This slight disagreement over the total number of contributory causes between the investigators might appear to be confusing given that the NTSB explicitly labels probable and contributory causes. However, some probable causes described several different problems. For example, the Guam report (AAR-00/01) contains the following argument; “the National Transportation Safety Board determines that the probable cause of the Korean Air flight 801 accident was the captain’s failure to adequately brief and execute the non-precision approach and the first officer’s and flight engineer’s failure to effectively monitor and cross-check the captain’s execution of the approach”. Prior to our analysis, we agreed that compound statements would be interpreted to yield several individual causes or contributory factors. Two different forms of aircrew error can be extracted from the previous example. The first is ‘the captains failure to adequately brief and execute the non-precision approach’. The second cause is ‘the first officer’s and flight engineer’s failure to effectively monitor and cross-check the captain’s execution of the approach’. As can be seen, this form of analysis depends upon a degree of subjective interpretation within the statements of probable cause. Hence Tables 1 and 2 indicate a surprising level of agreement between the two analysts. This example illustrates that a single incident can yield multiple instances of the same cause within our taxonomies. It is for this reason that Tables 1 and 2 provide the total number of different incidents involving a particular cause in parentheses. Hence an entry in either table of the form 2(1) would indicate two instances of a particular cause from a single incident. While an entry of the form 5(2) would indicate that two incidents had yielded five different instances of the associated cause for that year.

| | 1996 | | 1997 | | 1998 | | 1999 | | 2000 | | 2001 | | 2002 | | 2003 | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M |
| Human Error | 3 (3) | 3(3) | 6 (5) | 7(5) | 0 | 0 | 0 | 0 | 5 (2) | 3(2) | 2 (1) | 3(1) | 0 | 0 | 2 (2) | 2(2) | 18 (13) | 18(13) |
| ATM Failure | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Maintenance Problem | 1 (1) | 1(1) | 0 | 0 | 1 (1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 (1) | 1(1) | 1 (1) | 1(1) | 4 (4) | 3(3) |
| Company Management | 2 (2) | 1(1) | 3 (2) | 3(2) | 2 (1) | 3(2) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 (5) | 7(5) |
| Regulation | 2 (1) | 0 | 1 (1) | 1(1) | 3 (1) | 3(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 (3) | 4(2) |
| Equipment Failure | 0 | 0 | 0 | 0 | 0 | 1(1) | 1 (1) | 1(1) | 1 (1) | 1(1) | 0 | 0 | 0 | 0 | 0 | 0 | 2 (2) | 3(3) |
| Aircraft Design | 0 | 1(1) | 0 | 0 | 0 | 0 | 0 | 1(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2(2) |
| Manufacturing | 0 | 0 | 0 | 0 | 1 (1) | 1(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 (1) | 1(1) |
| Environment | 0 | 1(1) | 1 (1) | 1(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 (1) | 2(2) |
| Undetermined | 0 | 0 | 0 | 0 | 1(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 (1) | 0 |
| Total | 8 (7) | 7(7) | 11(9) | 12(9) | 8(5) | 8(5) | 1 (1) | 2(2) | 6 (3) | 4(3) | 2 (1) | 3(1) | 1(1) | 1(1) | 3 (3) | 3(3) | 40 | 40 |

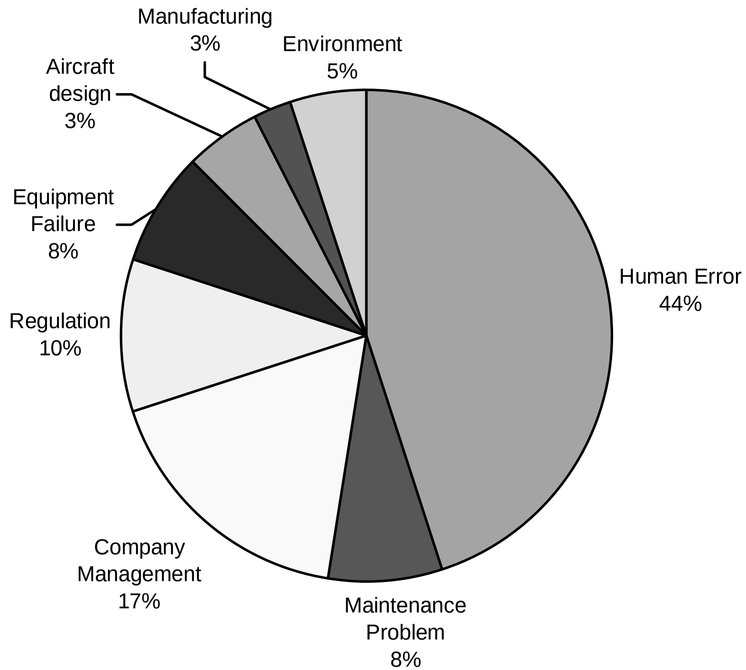
**Table 1:** Frequency of Probable Causes over Time (Analysts J and M, figures in parentheses represent number of different incidents)

| | 1996 | | 1997 | | 1998 | | 1999 | | 2000 | | 2001 | | 2002 | | 2003 | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M |
| Human Error | 3 (3) | 4(3) | 4 (3) | 5(4) | 1(1) | 1 | 0 | 0 | 2(2) | 2(2) | 3 (1) | 3(1) | 0 | 0 | 0 | 0 | 13 (10) | 15(11) |
| ATM Failure | 1 (1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 (1) | 0 |
| Maintenance Problem | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Company Management | 6 (4) | 8(5) | 3 (3) | 3(3) | 2(2) | 2(2) | 0 | 0 | 1(1) | 1(1) | 0 | 0 | 2(1) | 2(1) | 0 | 0 | 14 (11) | 16(10) |
| Regulation | 3 (2) | 6(3) | 4 (3) | 4(3) | 2 (1) | 2(1) | 0 | 0 | 4(2) | 4(2) | 0 | 0 | 2(1) | 2(1) | 0 | 0 | 15 (9) | 18(10) |
| Equipment Failure | 0 | 0 | 2 (1) | 2(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1(1) | 0 | 3 (2) | 3(2) |
| Aircraft Design | 0 | 0 | 2 (1) | 1(1) | 0 | 0 | 0 | 0 | 2(1) | 1(1) | 1(1) | 0 | 0 | 1(1) | 0 | 1(1) | 5 (3) | 3(3) |
| Manufacturing | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Environment | 1 (1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 (1) | 0 |
| Undetermined | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 14 (11) | 18(11) | 15 (11) | 15(12) | 5(4) | 5(4) | 0 | 0 | 9 (6) | 8(6) | 4 (2) | 3(1) | 4(2) | 5(3) | 1 (1) | 1(1) | 52 | 55 |

**Table 2:** Frequency of NTSB Contributory Causes over Time (Analysts J and M, figures in parentheses represent number of different incidents)

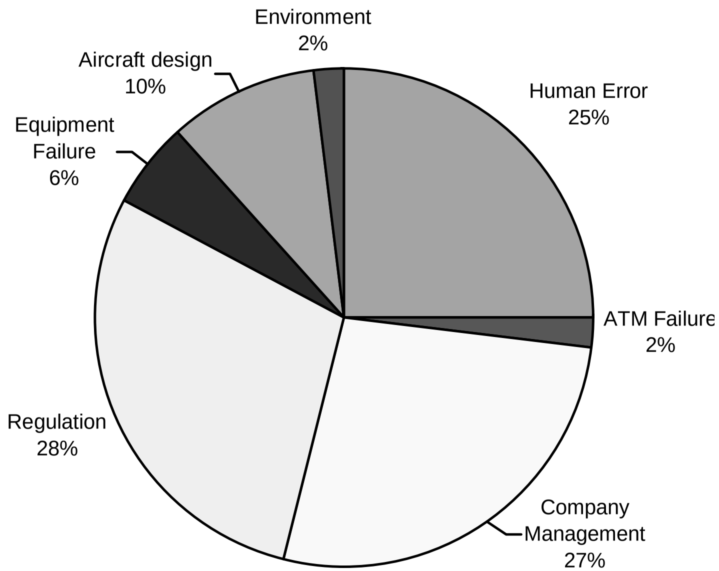



Analyst J

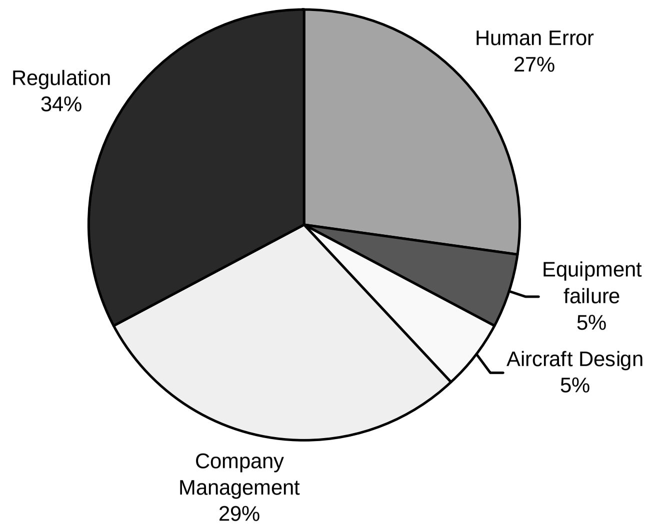



Analyst M

**Graph 1: Percentage of Probable Causes by Category (NTSB Data)**




Analyst J




Analyst M

**Graph 2: Percentage of Contributory Factors by Category (NTSB Data)**

As mentioned, the majority of the reports provided one probable cause. The mode for contributory causes was two. For instance, NTSB report AAR-96/05 contained the following summary:

> "The National Transportation Safety Board determines that the probable cause of this accident was the flight crew's failure to maintain the required minimum descent altitude until the required visual references identifiable with the runway were in sight. Contributing factors were the failure of the BDL approach controller to furnish the flight crew with a current altimeter setting, and the flight crew's failure to ask for a more current setting."

Both analysts identified the single probable cause as an instance of human error. They also agreed that there were two contributory causes. However, the first analyst identified one instance of an aircrew contributory failure and another associated with air traffic management. The second analyst categorised both under the heading of human error. In contrast to this incident, the standard deviations can be explained by exceptions to the mode that were characterised by a much higher number of probable or contributory causes. For instance, the NTSB report AAR 02-01 into Pacific Ocean Alaska Airlines Flight 261 provided the following summary of probable and contributory causes:

> "The National Transportation Safety Board determines that the probable cause of this accident was a loss of airplane pitch control resulting from the in-flight failure of the horizontal stabilizer trim system jackscrew assembly's acme nut threads. The thread failure was caused by excessive wear resulting from Alaska Airlines' insufficient lubrication of the jackscrew assembly. Contributing to the accident were Alaska Airlines' extended lubrication interval and the Federal Aviation Administration's (FAA) approval of that extension, which increased the likelihood that a missed or inadequate lubrication would result in excessive wear of the acme nut threads, and Alaska Airlines' extended end play check interval and the FAA's approval of that extension, which allowed the excessive wear of the acme nut threads to progress to failure without the opportunity for detection. Also contributing to the accident was the absence on the McDonnell Douglas MD-80 of a fail-safe mechanism to prevent the catastrophic effects of total acme nut thread loss."

Both analysts identified 1 primary cause and 5 contributory causes. In both cases the probable cause was listed as maintenance errors. They identified 2 regulatory failures, 2 management failures and a problem of aircraft design as contributory factors. This illustrates the considerable agreement over the probable and contributory causes for each incident. There were, however, minor differences in interpretation and classification. Given the limited sample size and the small number of analysts it is difficult to draw firm conclusions about the analysis of particular incidents.

Graphs 1 and 2 illustrate the overall percentage of probable and contributory causes identified by each analyst in each category. These diagrams provide arguably the most important insights from our work. There is considerable agreement over the proportion of causes that can be classified in each category. Both identified human error in 45% of all probable causes. There is no more than a 5% difference in any other category of probable cause. Analyst J identified regulatory issues in 15% of all probable causes compared to 10% for analyst M. For contributory factors there is a 2% difference in the proportion of causes associated with human error at 25% and 27%. The proportion associated with regulation also shows considerable agreement at 28% and 34%. Company management was identified in 27% of the contributory factors studied by analyst J while analyst M identified it in 29% of the factors. There are some small differences in the classifications used by each analyst. For example, Analyst J identified Air Traffic Management error in 2% of the contributory causes while analyst M assigned these causes within a wider definition of human error. Similarly, analyst J found environmental issues in 2% of contributory factors while several of these issues were classified as regulatory problems by analyst M. Overall, however, these differences seem remarkably small compared to the considerable agreement in the main categories of regulation, human error and company management.

There is no apparent bias towards 'blaming the operator' in earlier reports. The frequency of human error being identified by analyst J is: 3 probable and 3 contributory (1996), 6 probable and 4 contributory (1997), 0 probable and 1 contributory (1998), 0 probable and 0 contributory (1999), 5 probable and 2 contributory (2000), 2 probable and 3 contributory (2001), 0 probable and 0 contributory (2002), 2 probable and 0 contributory (2003). The frequency of distribution for analyst M is similar: 3 probable and 4 contributory (1996), 7 probable and 5 contributory (1997), 0 probable and 1 contributory (1998), 0 probable and 0 contributory (1999), 3 probable and 2 contributory (2000), 3 probable and 3 contributory (2001), 0 probable and 0 contributory (2002), 2 probable and 0 contributory (2003). The apparent peak in 1997 is not due to any single incident with a high number of human errors. Instead, it is explained by a series of reports that both analysts identified as being due to human error. These are AAR 97-05, 04, 03, 02 and 01. A similar rise in human error identified during 2000 makes it likely that these are semi-random artefacts stemming more from the types of incidents that occurred in those years. It seems unlikely that these peaks can be directly attributed to a subtler acceptance of a 'systemic view' over what Reason terms the 'person' approach to causal analysis. This view is further justified by the observation that the rise in human error is largely explained by two incidents in 2000, these are documented in AAR 00-01 and 00-02.

Our analysis of these diagrams and of the earlier tables also undermines some of the previous accusations made by safety researchers. Both analysts identified a large number of systemic causes and contributory factors throughout the sample of NTSB reports. Overall managerial failures were identified in approximately 17% of all probable causes and in 28% of all contributory factors. Similarly, regulatory issues account for 12% of all probable causes and 31% of all contributory causes. Managerial issues are the second most frequent classification for causal and contributory factors. Regulatory failures are the third most frequent probable cause and the *most frequent contributory factor* for both analysts.

It could be argued that fewer managerial and regulatory causes are identified after 1999. This contradicts the hypotheses that the proponents of systemic views have had an increasing impact on the work of investigatory organisations. As mentioned, the analysis of this data is non-trivial. For example, it could be argued that the residual 29% and 25% of contributory factors attributed to human error illustrate the dominance of a 'person' view of blame. This is, arguably, over simplistic and ignores the significance of managerial and regulatory factors throughout our sample reports.

| | 1996 | | 1997 | | 1998 | | 1999 | | 2000 | | 2001 | | 2002 | | 2003 | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M |
| Human Error | 3(3) | 3(3) | 2(2) | 4(3) | 6(4) | 6(6) | 9(3) | 10(4) | 4(2) | 4(2) | 1 | 1 | 6(3) | 5(3) | - | - | 31(18) | 33(22) |
| ATM Failure | 0 | 0 | 2(2) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 3(3) | 0 |
| Maintenance Problem | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | - | - | 1 | 1 |
| Loading Error | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 1 | 0 |
| Company Management | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 1 | 0 |
| Regulation | 0 | 0 | 0 | 0 | 2(2) | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 2(2) | 1 |
| Equipment Failure | 1 | 1 | 1 | 2(2) | 1 | 1 | 2(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 5(4) | 4(3) |
| Aircraft Design | 0 | 0 | 0 | 0 | 2(1) | 2(1) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 2(1) | 2(1) |
| Manufacturing | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2(1) | 0 | 0 | 0 | - | - | 3(2) | 0 |
| Environment | 0 | 0 | 1 | 0 | 2(2) | 2(2) | 0 | 0 | 0 | 0 | 0 | 0 | 2(2) | 1 | - | - | 5(5) | 3(3) |
| Total | 4(4) | 4(4) | 7(7) | 6(5) | 16(13) | 12(9) | 11(4) | 10(4) | 4(2) | 4(2) | 4(3) | 2(2) | 8(5) | 6(4) | - | - | 53 | 44 |

**Table 3:** Frequency of Probable Causes over Time, Analysts J & M. Parentheses represent number of different incidents, Canadian TSB.

| | 1996 | | 1997 | | 1998 | | 1999 | | 2000 | | 2001 | | 2002 | | 2003 | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M | J | M |
| Human Error | 1 | 2(2) | 3(2) | 5(2) | 1 | 6(4) | 0 | 0 | 1 | 3(1) | 1 | 1 | 1 | 2(1) | - | - | 8(7) | 19(11) |
| ATM Failure | 0 | 0 | 2(2) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 3(3) | 0 |
| Maintenance Problem | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 1(1) | 0 |
| Company Management | 2(2) | 2(2) | 5(1) | 12(3) | 3(3) | 5(4) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 10(6) | 19(9) |
| Regulation | 0 | 0 | 3(2) | 6(1) | 0 | 2(2) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 3(2) | 8(3)) |
| Equipment Failure | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 2(2) | 2(2) | - | - | 3(3) | 5(5) |
| Aircraft Loading | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 1(1) | 0 |
| Aircraft Design | 0 | 1 | 2(1) | 3(2) | 1 | 6(3) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 3(2) | 11(6) |
| Manufacturing | 0 | 2(1) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 0 | 3(2) |
| Environment | 0 | 0 | 1 | 4(2) | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | - | - | 3(3) | 6(4) |
| Total | 4(4) | 7(6) | 17(10) | 30(12) | 7(7) | 20(14) | 0 | 2(2) | 2(2) | 4(2) | 1(1) | 1(1) | 4(4) | 5(4) | - | - | 35 | 71 |

**Table 4:** Frequency of Contributory Causes over Time. Analysts J & M. Parentheses represent number of different incidents, Canadian TSB.

Canadian TSB Results

We were anxious to determine whether the US NTSB were atypical in the prominence of regulatory and managerial factors in their causal analysis of major accidents over the last decade. In consequence, we decided to replicate our study with a sample of incidents investigated by the Canadian TSB. Our work proceeded in a similar manner to that of the NTSB sample. The first stage was to make our initial selection of incidents from the many thousands of adverse events that are reported to the TSB each year. Again we focused our efforts on the higher profile occurrences that resulted in more sustained reports. These did, however, include near miss incidents as well as multiple fatalities. The heuristic that we adopted was to investigate every aviation incident report that was composed of distinct numbered sections between 1996 and 2003. The decision to focus on these reports was justified because they provide the closest analogy to the NTSB's major accident reports, as opposed to their accident briefs. However, unlike the NTSB the TSB had no published incidents for 2003 when we performed the analysis early in 2004. Similarly, there were no reports with distinct numbered sections for 2002. We, therefore, substituted a number of the less structured, reports for this year only. These were reports A02C0124, A02F0069, A02P0109, A02Q0130. This decision was justified by the need to avoid a gap in our sample for the last two years. We were also unable to determine whether this less structured format will provide a standard for future TSB reports. Our selection policy was further justified by the observation that this left a total sample of 27 TSB documents compared to 26 reports from the NTSB.

The next stage in our analysis was to extract those sections that dealt with probable causes and contributory factors. The TSB documents in our sample included separate sections on "Findings As to Causes and Contributing Factors", "Findings as to Risk", and "Other Findings". We focussed on the sections detailing causes and contributory factors, as we had done with the NTSB sample. Some reports only had two sections entitled "Causes" and "Findings"; again we focussed on the section describing the causes of the adverse event. We then independently used the same procedure to classify each causal and contributory factor. This proved to be more problematic than had been the case with the NTSB reports. Most of the difficulty stemmed from the way in which the TSB combined probable causes and contributory factors. For instance, the section of report Number A02F0069 entitled 'Findings as to Causes and Contributing Factors' contained the following list:

1. The pilot not flying (PNF) inadvertently entered an erroneous $V_1$ speed into the MCDU. The error was not detected by either flight crew, despite numerous opportunities.

2. The PNF called "rotate" about 25 knots below the calculated and posted rotation speed.

3. The pilot flying (PF) initiated rotation 24 knots below the calculated and posted rotation speed and the tail of the aircraft struck the runway surface.

4. A glide path signal was most probably distorted by a taxiing aircraft and provided erroneous information to the autopilot, resulting in a pitch-up event. The pitch-up could have been minimized if the autopilot had been disconnected earlier by the PF.

As can be seen, there is no indication as to which of these items is a cause and which is a contributory factor. Each analyst, therefore, had to use his own judgement. Both identified three causes relating to human error and one contributing factor relating to problems in equipment design. This reliance on individual judgement did, however, create disagreement over causes and contributory factors. The 27 incidents yielded a total of 53 probable causes for Analyst J. The mean number of probable causes was 1.9 with a standard deviation of 1.2. The second analyst identified 44 probable causes with a mean of 1.6 and a standard deviation of 1. There were 35 contributory causes identified by the first analyst with a mean of 1.3 and a standard deviation of 2.5. The second analyst identified 71 contributory causes with a mean of 2.5 and a standard deviation of 3.9. The mode over all probable causes was 1 while the mode for all contributory causes was 0.

The variance between the investigators could have been reduced if a more formal method for distinguishing causes from contributory factors had been used (Johnson, 2003). At the start of the study, we decided not to use this approach because the development of appropriate root cause analysis techniques remains an active area for research. We were also keen to employ the subjective criteria that might be employed by the readers of these documents. We are currently exploring the impact that more formal techniques might have on the results of our analysis. Having raised these caveats, it is possible to present the results from this part of the study.

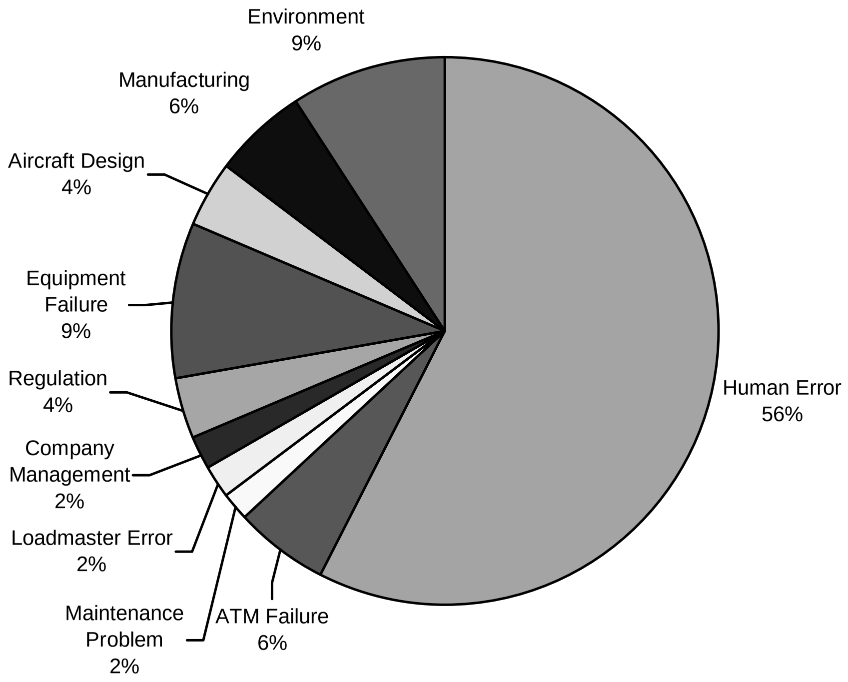
As before, there is a relatively high standard deviation associated with the mean results for both causes and contributory factors. Again, this can be explained in terms of a small number of reports, which were very different from the mode of one cause and zero contributory factors. In particular, analysts J identified two causes and thirteen contributory factors in report A97H0011. Analyst M identified the same number of causes but also identified twenty contributory factors. This incident involved a loss of control on go-around under adverse weather conditions. Analyst J identified human error and problems in air traffic management as the main causes. Analyst M identified two instances of the more general form of 'operator error'. The thirteen contributory factors identified by Analyst J included five instances of managerial failure, two human errors, two regulatory problems, two aircraft design issues, a maintenance failure and a problem relating to the operational environment in which the accident occurred. In contrast, analyst M identified three human errors, seven management problems, six regulatory failures, three environmental factors and one instance of equipment failure. A number of other instances also helped to pull the standard deviation away from the mode. For example, both analysts identified five instances of human error causing the incident described in TSB report A99Q0151:

> "The pilot flying did not establish a maximum performance climb profile, although required by the company's standard operating procedures (SOPs), when the ground proximity warning system (GPWS) "Terrain, Terrain" warning sounded during the descent, in cloud, to the non-directional beacon (NDB). The pilot flying did not fly a stabilized approach, although required by the company's SOPs. The crew did not carry out a go-around when it was clear that the approach was not stabilized. The crew descended the aircraft well below safe minimum altitude while in instrument meteorological conditions. Throughout the approach, even at 100 feet above ground level (agl), the captain asked the pilot flying to continue the descent without having established any visual contact with the runway environment. After the GPWS "Minimums, Minimums" voice activation at 100 feet agl, the aircraft's rate of descent continued at 850 feet per minute until impact. The crew planned and conducted, in cloud and low visibility, a user-defined global positioning system approach to Runway 31, contrary to regulations and safe practices."

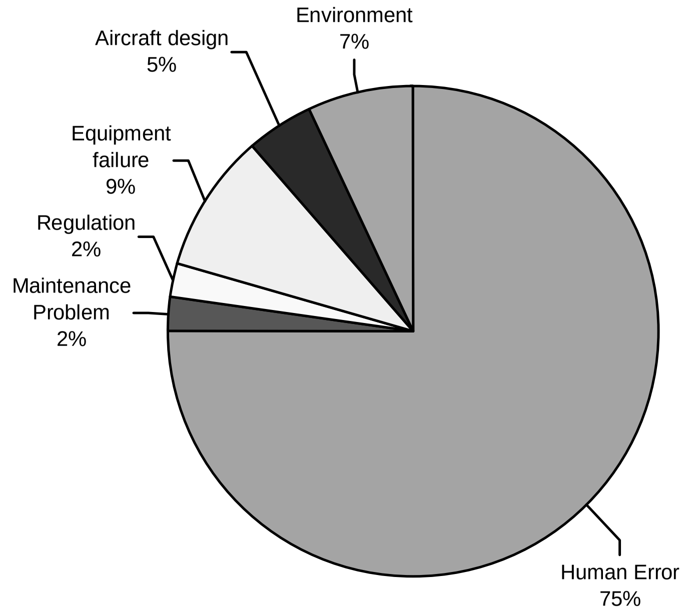
Graphs 3 and 4 provide an overview of the distribution of causes and contributory factors that were identified by our analysis. As with Graphs 1 and 2, they provide a number of important insights about the structure of our analysis and the causal findings of different investigation organisations. The most striking feature of the pie charts in Graph 3 is that they again show the relatively high level of agreement at this level of our analysis. Both investigators identified human error as the most common causal factor across the TSB sample at 56% and 75%. The relative difference between the proportions of incidents identified by the two analysts can be explained in terms of the broader range of categories that were considered by analyst J compared to analyst M. For example, analyst J also included 'loadmaster error' and 'ATM error' that were not included within the classification used by analyst M. With this caveat, the remaining results show considerable agreement; both analysts fall within one or two percent of their colleagues classification for environmental causes (9% and 7%), aircraft design (4% and 5%), equipment failure (9% and 9%), regulation (4% and 2%), maintenance (2% and 2%).

Graph 4 shows that human error plays a lesser role in the contributory factors that were identified by both analysts in the TSB sample (22% by analyst J and 28% by analyst M). Again, there is considerable agreement in terms of the overall percentages for each category of contributory factor. Analyst J identified 28% of all contributory factors as being related to company management. Analyst M found this in 27% of the factors in the accident reports. The agreement continued in regulation (9% and 11%), equipment failure (9% and 7%) and environmental factors (9% and 8%). There is a more noticeable disagreement over the role of aircraft design. Analyst J identified it in only 9% of these contributory factors. Analyst M identified design flaws in 15% of the factors in the TSB sample. This can be explained in terms of a cluster of incidents in 1998. Analyst M identified three aircraft design flaws in the contributory factors for A98H0003, two in A98H0002 and one in A98C0173. In contrast, Analyst J identified management failure as a contributory factor behind these design flaws.

In contrast to the NTSB reports, we identified a greater tendency to 'blame' the operator in the TSB sample. We did not, however, identify any trend away from this with the rise of 'systemic theories' of failure in recent years. The frequency of human error identified by analyst J is: 3 probable and 1 contributory (1996), 2 probable and 3 contributory (1997), 6 probable and 1 contributory (1998), 9 probable and 0 contributory (1999), 4 probable and 1 contributory (2000), 1 probable and 1 contributory (2001), 6 probable and 1 contributory (2002). The frequency of distribution for analyst M is: 3 probable and 2 contributory (1996), 4 probable and 5 contributory (1997), 6 probable and 6 contributory (1998), 10 probable and 0 contributory (1999), 4 probable and 3 contributory (2000), 1 probable and 1 contributory (2001), 5 probable and 2 contributory (2002). As mentioned, the peak in 1999 is due largely to A99Q0151. There are also relatively high levels of human error identified during 1998 and 2002. 2002 was similar to 1999, with a single incident documented as A02F0069 producing several different forms of human error. However, several different forms of human error distributed across a relatively high number of individual reports explain the rise in 1998: A98Q0192, A98P0303, A98H0011, A98H0003, A98H0002, A98C0173, A98A0191, A98A0067. As with the NTSB data, it is difficult to identify any trends that might characterise any change in the 'systemic view' over the 'person' approach to causal analysis, at least in terms of the distribution of human error between 1996 and 2002.

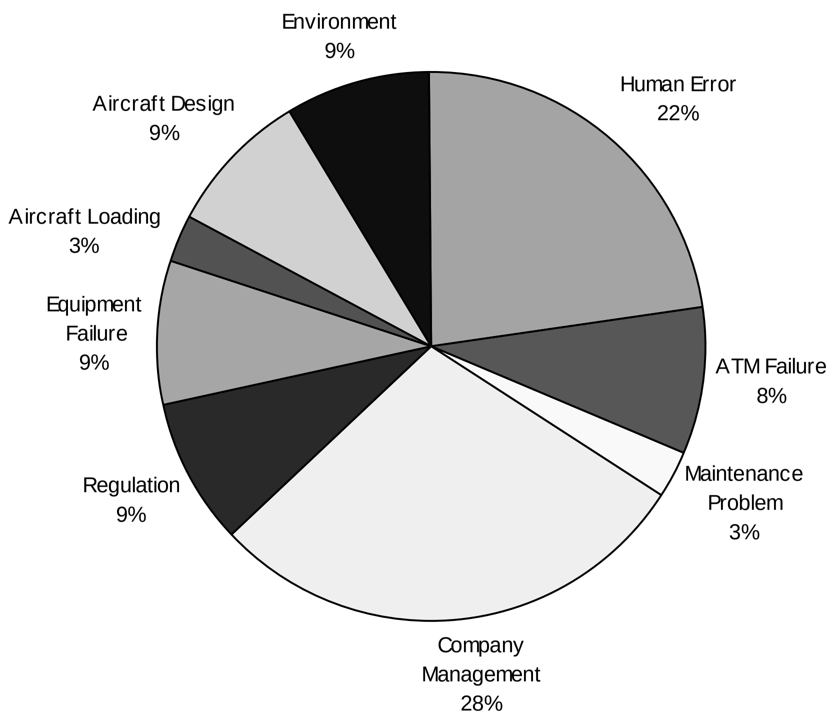



Analyst J

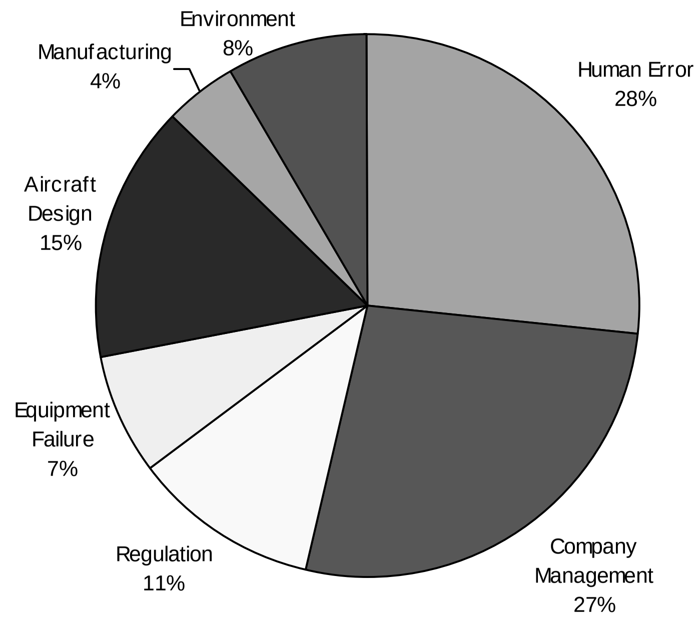



Analyst M

**Table 3: Percentage of Primary Causes by Category (Canadian TSB)**




Analyst J




Analyst M

**Table 4: Percentage of Contributory Causes by Category (Canadian TSB)**

Previous paragraphs have argued that our analysis of the TSB reports again confirms the relative prominence of systemic factors alongside human error. However, managerial and regulatory failures play a greater role in the contributory factors than they do in probable causes. Managerial issues account for around 1% of all causes and 27% of contributory factors across both analysts. Regulatory issues account for 3% of all causes and 10% of all contributory factors. This is a significant finding not so much for what it reveals about the practices and perspective of the TSB but for what it reveals about the opinions of the two analysts who were involved in this exercise. Recall that the TSB reports do not distinguish between contributory factors and causes. Hence we were making qualitative judgements about those failures that should be assigned to each general classification. This analysis suggests that we were predisposed to view human error as a more salient probable cause than either managerial factors or regulatory failure.

It is difficult to argue in favour of any rise in the 'systemic view' of failure. In contrast, there is a striking decline in the role of managerial and regulatory issues in the TSB reports. All of the causes and contributory factors that were classified according to these 'systemic' sources of failure come before 1999. In contrast, both analysts continued to identify 'person' causes between 1999 and 2002. Analyst J found 6 instances of aircrew 'error' in 2002 while their colleague found 5. The remaining causes and contributory factors were equipment failures, manufacturing flaws and environmental issues.

## Conclusions

When we began this analysis, we were keen to determine whether or not the 'systems' view of failure was having an impact on the reports issued by the Canadian TSB and the US NTSB. Prominent investigators in both of these organisations have argued that these factors must be considered when identifying the causes of adverse events (Ref. 8 and 9). Our results have shown that both the NSTB and TSB considered a wide range of causal and contributory factors in their reports. In particular, it seems clear that both organisations have a long tradition of considering the regulatory and managerial precursors to adverse events. However, both of these organisations focus on the role of human error as a potential cause in most of the adverse events that they investigate. It also seems that the prominence of managerial and regulatory issues has declined in recent years. This trend is more marked in the TSB data set than it is in that of the NTSB.

One explanation for the prominence of human error is that many investigatory agencies rely upon counterfactual definitions of root causes (Ref. 4). These are distinguished by arguments of the form 'if the probable cause X had not occurred then the mishap also would not have occurred'. Several existing investigation techniques, including Multilinear Event Sequence and Events and Causal Factors Charting, embody these arguments within their associated analytical techniques. Investigators look at the events immediately preceding the failure. They then ask whether the accident would have been avoided if that event had not taken place. If the accident is still likely to have occurred even if that event had been prevented then the investigation moves on to look at its precursors until a root cause is identified. One by-product of these techniques is that they are likely to identify direct operator errors as root causes because these are often identified as immediate precursors to an adverse event. The current NTSB accident coding manuals recommend a sequence of events modelling matrix that is very similar to both of the techniques cited above. However, further work is required to determine whether or not counterfactual explanations account for the results that are presented in this paper. Alternatively, the relative frequency and recency of human factors causes in other investigations may help to create a reinforcement cycle in which these factors are more likely to be cited in the future because they have often been cited in the past.

It is also possible to look beyond the trends that we have identified and propose possible explanations for the differences between the NTSB and the TSB. We have identified a greater proportion of managerial and regulatory causes and contributory factors in the NTSB reports. This might be explained by the relatively lower proportion of accidents involving major air carriers compared to small companies and private carriers in the Canadian data set. To summarise, we would argue that it is *inaccurate* to assert, as some have, that: (1) the operator is always blamed, (2) most investigations stop as soon as they find someone to blame, or (3) organizational causes are usually ignored.

This study has also taught us how difficult it can be to make inferences about the attitudes of complex, investigatory organizations from the reports that they produce. This should not be any surprise. It is, however, surprising that there have been so few previous attempts to conduct such a study. Without more detailed evidence it is difficult to sustain many of the criticisms that have been made about investigatory agencies and their attitudes towards the role of operator error. Particular problems stemmed from the difficulty of determining whether or not a particular cause was related to human error. Similarly, it can be hard to determine whether certain paragraphs described a single operator failure or a series of related errors. In practice, however, we were able to reach considerable agreement over these apparently intractable problems. The proportion of causes and contributory factors in each category over the eight years of this study were very similar, as revealed by graphs 1 to 4. Further problems stemmed from the difficulty of distinguishing causes from contributory factors in the TSB excerpts. These issues were grouped together in the same paragraph. In the process of identifying which issues were causes and which were contributory factors we also revealed an important analytical bias. We were more likely to identify human error as a cause and managerial or regulatory failures as contributory factors. This leads us to question whether we should attach such prominence to proximal cause rather than these more distal factors. It also reveals how deeply ingrained the 'person' approach may actually be within incident and accident investigators.

This paper describes the results of an independent analysis of the primary and contributory causes of aviation accidents in both the United States and Canada between 1996 and 2003. The purpose of the study was to assess the comparative frequency of a range of causal factors in the reporting of adverse events in aviation. Our results suggest that the majority of these high consequence accidents were attributed to human error. A large number of reports also mentioned wider systemic issues, including the managerial and regulatory context of aviation operations. These issues are more likely to appear as contributory rather than primary causes in both sets of accident reports. The relative prominence of 'systemic causes' seems to have declined in recent years; this contradicts recent assertions about the importance of the latent rather than proximal causes of adverse events.

## References

1. J. Reason, Human Error: Models and Management, British Medical Journal, 320:768-770, 18 March, 2000.

2. D. Woods and R.I. Cook, The New Look at Error, Safety, and Failure: A primer, Technical Report, US Veterans Association, National Centre for Patient Safety, 2004.

3. J.L. Mackie, The Cement of the Universe, Oxford University Press, Oxford, 1980.

4. C.W. Johnson, The Failure of Safety-Critical Systems: A Handbook of Accident and Incident Reporting, Glasgow University Press, Glasgow, 2003.

5. A. DeVogue, L.Douglass and A.B. Stoddard, Lawyer Says Ex-Enron Chief Will Appear; Lawmakers Rip Accountant, February 5th 2002. Available from: http://abcnews.go.com/sections/business/DailyNews/enron_lay_hearings020205.html

6. D. Goldin, "When The Best Must Do Even Better" Remarks by NASA Administrator Daniel S. Goldin At the Jet Propulsion Laboratory Pasadena, CA March 29, 2000, NASA Headquarters, Washington DC, USA, http://www.hq.nasa.gov/office/pao/ftp/Goldin/00text/jpl_remarks.txt, 2000.

7. NTSB, Annual Review of Aircraft Accident Data U.S. General Aviation, Calendar Year 1999, Washington, DC. Available from: http://www.ntsb.gov/publictn/2003/ARG0302.pdf

8. B. Strauch, Normal Accidents: Yesterday and Today. In C.W. Johnson (editor), Proceedings of the First Workshop on the Investigation and Reporting of Incidents and Accidents (IRIA 2002), Department of Computing Science, University of Glasgow, Scotland, 2002.

9. M. Ayeko, Integrated Safety Investigation Methodology (ISIM): Investigation for Risk Mitigation. In C.W. Johnson (editor), Proceedings of the First Workshop on the Investigation and Reporting of Incidents and Accidents (IRIA 2002), Department of Computing Science, University of Glasgow, Scotland, 2002.

10. C. Perrow, Normal Accidents: Living with High-Technology Accidents, Princeton University press, Princeton, 1999.