The Ohio State University
Eighth International Symposium on Aviation Psychology

What ASRS Incident Data Tell About Flight Crew Performance During Aircraft Malfunctions

Robert L. Sumwalt
Alan W. Watson

Battelle
Putting Technology To Work

Aviation Safety Reporting System
625 Ellis St. Suite 305 Mountain View California 94043

ASRS

NASA ASRS (Pub. 52)

"Oh no! We've Got a Problem"

WHAT ASRS INCIDENT DATA TELL ABOUT FLIGHT CREW PERFORMANCE DURING AIRCRAFT MALFUNCTIONS

Robert L. Sumwalt[1] and Alan W. Watson[2]
Battelle's Aviation Safety Reporting System Program Office
Mountain View, California

ABSTRACT

This research examined 230 reports in NASA's Aviation Safety Reporting System's (ASRS) database to develop a better understanding of factors that can affect flight crew performance when crew are faced with inflight aircraft malfunctions. Each report was placed into one of two categories, based on severity of the malfunction. Report analysis was then conducted to extract information regarding crew procedural issues, crew communications and situational awareness. A comparison of these crew factors across malfunction type was then performed. This comparison revealed a significant difference in ways that crews dealt with serious malfunctions compared to less serious malfunctions. The authors offer recommendations toward improving crew performance when faced with inflight aircraft malfunctions.

INTRODUCTION

Research from a major aircraft manufacturer states that a large number of aircraft accidents attributed to human error begin with an aircraft malfunction. (Wiegers and Rosman, 1986) Several of these accidents have been caused by the flight crew's fixation on the aircraft malfunction, which resulted in their overall loss of situational awareness. Examples include the December 1972 Eastern Air Lines L1011 crash in the Florida Everglades, and the December 1978 United Air Lines DC8 accident in Portland, Oregon. Both of these accidents are now well known cases, and are frequently cited in Crew Resource Management (CRM) classes worldwide.

The NASA Aviation Safety Reporting System (ASRS) database contains thousands of reports that cite aircraft malfunctions. This large number of related incident reports creates fertile ground for exploration of flight crew performance during aircraft malfunctions.

OBJECTIVE AND SCOPE

The objectives of this research study were twofold: to develop a better understanding of factors - both positive and negative - that can affect crew performance when faced with inflight aircraft malfunctions, and to offer recommendations designed to improve crew performance during these conditions.

In order for an ASRS report to be included in the study set, it must have involved a crew size of at least two pilots (including instructional flights) and involved the actual or perceived inflight malfunction of a major aircraft system or subsystem. Further it was stipulated that the mechanical malfunction must have created a relatively prolonged period of demand on aircrew communications, attention and procedures after the mechanical malfunction was discovered by the crew. This was to eliminate those situations that were immediately resolved by flight crew "reflex action" such as a runaway stabilizer malfunction, or an autopilot "hardover."[3]

APPROACH

Data

Our data set consisted of 230 ASRS reports that were submitted to ASRS between May 1986 and August 1994. The researchers were well aware that ASRS data, including those in this study, may reflect reporting biases. Chappell (1994) notes that reporters' incident descriptions are influenced by their individual motivations for reporting, and that reports often give only one perspective of the event which is not balanced by additional investigations or verification. Not withstanding these caveats, Chappell states, "If large numbers of reports on a

---

[1] Robert L. Sumwalt has served ASRS as a research consultant since 1992. Additionally, he is an active Captain for a major U.S. air carrier, where his duties have included airline check airman, instructor pilot and air safety investigator. He has published more than 40 articles and papers on aviation safety issues, including presentations at the Sixth and Seventh International Symposiums on Aviation Psychology.

[2] Prior to joining ASRS in 1994 as an analyst, Alan W. Watson served as a B767 international Captain for a major U.S. airline. He accumulated over 30,000 flight hours during his 29 years as an air carrier pilot and 6 years as a U.S. Air Force pilot. Prior to his retirement from airline operations he held positions of FAA Designated Flight Examiner and line check airman.

[3] Following our initial search of the ASRS database we found that a large number of retrieved reports did not meet the rigid scoping criteria of this study. To reduce this high number of "false positives," we conducted another ASRS database search and purposely excluded certain types of reports. Examples of excluded reports are those that mentioned rejected take-offs, loss of communications for reasons such as "stuck mics" and any auto-flight automation anomalies. While eliminating these types of reports, we simultaneously refined the search criteria to look specifically for certain malfunction key phrases, such as "engine fire," "landing gear malfunctions," and "flap/slat problems." Because of these search strategy manipulations, it cannot be assumed that the type or percentage of malfunctions cited in this study are a representative sampling of the total number of aircraft malfunctions in the ASRS database. Further, as with all ASRS data, no inference can be drawn as to how these numbers relate to the total population of all aircraft malfunctions.

topic are available, it is reasonable to assume that consistently reported aspects are likely to be true. It is doubtful that a large number of reporters would exaggerate or report erroneous data in the same way" (pp.154-155).

Method

Prior to initiating this research the investigators turned to several sources to determine the type of information that should be gathered to evaluate crew performance. One helpful source was FAA Advisory Circular AC 120-51A, *Crew Resource Management Training,* which notes that many successful CRM programs use three key cluster areas to evaluate flight crew performance: (1) Communications Processes and Decision Behavior, (2) Team Building and Maintenance, and (3) Workload Management and Situational Awareness.

The next step was to develop an extensive listing of potential aircraft malfunctions that researchers expected to find in their review of ASRS reports. Each of these potential aircraft malfunctions was then placed into one of two categories, depending on severity of the malfunction. This was done to allow statistical comparison of crew performance when dealing with serious problems versus less serious problems.

Type A malfunctions were those that we judged as being quite serious and posing the real or perceived threat of loss of life or equipment, (e.g., engine fire or failure, inability to extend landing gear and major flight control problems that grossly affect the ability to control the aircraft.). Type B malfunctions were those that were deemed to be less serious in nature. (e.g., flap problems, air-conditioning malfunctions and minor hydraulic system malfunctions.) We further distinguished Type A and Type B malfunctions by noting that malfunctions placed in the former category are resolved by many air carriers by use of "Emergency" Checklists, while those placed into the latter category are resolved by "Abnormal" Checklists.

By use of a six-page questionnaire, bits of relevant information were extracted from each ASRS report in the data set. The reports were analyzed as to which type of malfunction occurred and what crew factors were present. A comparison of crew factors across malfunction types was then performed.

FINDINGS AND DISCUSSION

Ninety-five percent of ASRS reports in this study involved air carrier operations; 92 percent involved passenger carrying operations. Two-thirds of these reports had a crew size of two pilots, while one-third involved three crewmembers.

Malfunction type

Of the 230 reports in the data set, 199 cited single malfunctions and 31 cited multiple malfunctions. Regarding multiple malfunctions, one report referenced five aircraft malfunctions, one cited four malfunctions, five reports referenced three malfunctions, and 24 reported dual malfunctions.

By design, we sought to evaluate approximately the same number of Type A and Type B malfunction reports. Type A malfunctions were noted in 105 of the 230 reports (46 percent) while Type B malfunctions were prevalent in 112 of the 230 reports (48 percent). A combination of Type A and Type B malfunctions were found in 13 of the 230 reports (6 percent). During analysis and throughout this paper we refer to these as "Type C" reports. Table 1 shows the four most frequent citations[4] of aircraft malfunctions for Type A and B categories.

TYPE A MALFUNCTIONS (105 of 230 Reports)	Citations
Engine fire and other serious engine problems requiring inflight shutdown	72
Landing gear - inability to extend (or verify down)	18
Smoke or fumes in cockpit or cabin	15
Rapid depressurization	4
TYPE B MALFUNCTIONS (112 of 230 Reports)	
Air-conditioning/pneumatic/pressurization system	33
Flap and slat	23
Hydraulic	17
Landing gear - non-major (anti-skid, brake pressure)	15

Table 1. Four most frequent aircraft malfunction citations for Type A and B Malfunctions

[4] A citation is where an ASRS report stated (or cited) a particular situation or occurrence. A single ASRS report may reference more than one situation or problem. Therefore, the total citations may exceed the total number of reports. For example, one ASRS report cited a hydraulic failure that resulted in the failure to properly extend the landing gear. In this example, one ASRS report yielded two "malfunction citations." Note that the data base search techniques did not necessarily produce a representative sample of aircraft malfunction types (refer to footnote 3 for additional information).

Crew procedural issues

We were interested to see if the crew followed prescribed procedures to deal with these malfunctions. 169 of the 230 reports in the data set provided this information. Table 2 compares the number of reports where crews followed prescribed procedures versus those reports where crews did not. Examples of improper procedural actions include failing to complete a checklist due to rushing, using the wrong checklist, and turning off the operative generator after a generator malfunction was discovered.

	A	B	C	TOTALS
Crew Followed Prescribed Procedures	90	54	5	149
Crew Did Not Follow Procedures	4	14	2	20
No Information Available	11	44	6	61
TOTALS	105	112	13	230

Table 2. Crew procedural usage according to malfunction type

Chi-square analysis revealed a significant difference between Type A and Type B malfunctions[5] regarding crews following (and not following) prescribed procedures. X^2 (df=2, N=230) =25.75, $p < .05$.

Eighty-eight of the 230 reports provided information concerning whether or not an emergency was declared following discovery of the mechanical malfunction. Of those 88 reports, 71 of the reporters specifically stated that they declared an emergency, while 17 wrote that they did not declare an emergency. Of those 71 reports where an emergency was declared, 40 indicated that an emergency was declared immediately or very soon after the problem was detected. Nine of these reports noted that an emergency was declared after a delay of some length. After discovering that their landing gear would not extend, one crew delayed declaring an emergency for 2 1/2 hours while they circled to burn excess fuel. In two cases the crew did not declare an emergency until on short final approach, and only then because ATC positioned another aircraft onto the runway just ahead of them. Some reporters wrote in retrospect: "Declaring an emergency may have allowed us priority handling, and hence, less traffic disturbance." (ASRS Record No. 211356) "It would have been much safer to inform ATC of our suspected problem early on." (ASRS Record No. 152994)

Communications processes and decision behavior

Positive Communications (75 of 230 reports)	
Captain's open solicitation of input	(38 of 75 reports)
Briefing concerning planned actions, solutions, or crew coordination issues	(26 of 75 reports)
Crewmembers providing input and/or voicing safety concerns	(20 of 75 reports)
Active participation encouraged in decision-making process	(9 of 75 reports)
Negative Communications (14 of 230 reports)	
Captain not receptive to crewmember input	(7 of 14 reports)
Strain or difficulty with crew communications	(5 of 14 reports)
Captain failed to keep others informed of intentions	(2 of 14 reports)

Table 3. Citations concerning crew communications (information supplied in 89 of 230 reports)

Previous NASA research has shown that the type and quality of crew communications are predictors of crew performance (Foushee & Manos, 1981; Foushee, Lauber, Baetge & Acomb, 1986). Because of this previous research, and because the importance of crew communications is widely emphasized in effective CRM programs, we were interested to see what crew communications information could be distilled from these ASRS reports. Only 89 of the study's 230 reports had information pertaining to crew communications. Reports citing instances of crews using "positive" communication techniques outnumbered the reports of "negative" communications by a factor of five to one. Table 3 describes some of the findings concerning crew communications.

Regarding the Captain's open solicitation of input (Positive Communications, Table 3), many reports indicated that solicitation of input was not just limited to cockpit crewmembers. Eight reports cited input from company maintenance facilities via radio, while 7 reported radio calls to the company dispatcher for input. Flight Attendant input was sought in 7 reports where information was needed about passenger status or problems visible in or from the passenger cabin.

---

[5] For each of the chi-square tests in this paper we computed these comparisons two separate ways. One comparison was made using Type A, B and C categories. Another comparison was made by first combining A and C categories, and then comparing that group against the B category. The comparison was significant in both cases. Our rational for combining A and C categories was that, by definition, Type C malfunctions were those that had a combination of Type A and B malfunctions. We hypothesized that a crew who experienced both serious and less-serious malfunctions would prioritize tasks and deal with the serious malfunction (Type A) first and most aggressively, then as resources became available, begin dealing with the less-serious malfunction (Type B). In some cases, dealing with more serious malfunction preempted resolution of the less-serious malfunction. Therefore, we felt that combining Types A and C for these comparisons was logical.

In addition to previously discussed unknown and known reporting biases, certain caveats are in order when making statistical inferences from these data. First, these reports represent only a portion of the total number of all aircraft malfunctions that have occurred. Therefore, conclusions drawn from this analysis are only valid for the reports in this study, and are not necessarily valid for the total population of all aircraft malfunctions. Also, care must be taken not to assume a cause and effect relationship among the elements compared (Chappell, 1994).

Reporters exemplified positive communications/decision behavior with statements such as "decision making in a collective environment, and coordination between us (and the cabin team) went extremely well." (ASRS Record No. 204057) An example of a "negative communications" citation came from a report where the First Officer informed the Captain that he was not comfortable with the situation, but the Captain continued the flight despite the input.

Workload management and situational awareness

Schwartz (1986) identified ten items that may serve as clues to mark loss of flight crew situational awareness. Schwartz referred to these as elements of an "error chain." By slightly modifying this list we used its components to seek evidence of crewmember loss of situational awareness for reports in this study. At least one "error chain" element was identified by this research team in 73 of the 230 reports. Figure 1 depicts the error chain clues, along with the number of their citations according to Type A, B, and C malfunction classifications.[6]

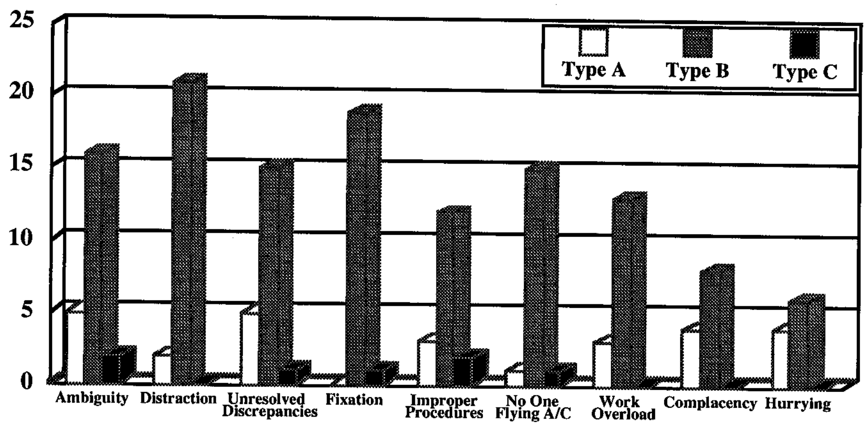

Figure 1. Citations of "Error Chain" clues (information supplied in 73 of 230 reports)

We theorized that having a number of simultaneous error chain clues could have a cumulative effect on decreasing crew performance during the resolution of malfunctions. Table 4 depicts the number of these simultaneous error chain clues, according to malfunction type. To determine if the number of simultaneous error chain clues was dependent of malfunction type, we performed a chi-square test for independence. This test showed a significant difference between Type A and B malfunctions. X^2 (df=4, N=230) =31.12, $p < .05$.

	A	B	C	TOTALS
One Clue	13	16	3	32
Two Clues	4	10	2	16
Three Clues	1	12	0	13
Four Clues	0	10	0	10
Five Clues	1	3	0	4
Six Clues	0	1	0	1
TOTALS	19	52	5	76

Table 4. Number of simultaneous error chain clues present according to malfunction type

Adverse safety problems

We were interested to see if the attention demands on the flight crew during resolution of the aircraft malfunction caused any adverse safety consequences. Of the 230 reports reviewed in this research, 192 (83 percent) provided no evidence of any further consequences or safety problems. The remaining 38 (17 percent) led to sundry problems. Table 5 shows the distribution of adverse safety consequence citations. We statistically compared two elements of this category - altitude deviations and course/track/heading deviations - to look for significant differences between Type A and B malfunction types. Chi-square tests show a significant difference between Type A and B malfunctions for both of these adverse safety consequence items. For altitude deviations, X^2 (df=1, N=230) =9.68, $p < .05$, and for course/track/heading deviations X^2 (df=1, N=230) =6.16, $p < .05$.

ADVERSE SAFETY PROBLEM	A	B	C	TOTALS
Altitude deviations	0	14	2	16
Non-adherence to ATC clearance	0	11	2	13
Course/track/heading deviations	0	9	2	11
Non-compliance with FARs/SOPs	0	7	1	8
Other[7]	3	5	0	8
TOTALS	3	46	7	56

Table 5. Number of adverse safety consequences according to malfunction type

[6] It should be noted that these error chain clues are not mutually exclusive, and some are closely related (for example, "distraction" and "no-one flying the aircraft"). During analysis we took measures to ensure that we had not simply "double coded" related error chain clues, thus artificially increasing their count.

[7] Included in the "other" category are reports of passenger or crew injury and aircraft damage.

CONCLUSIONS AND RECOMMENDATIONS

When viewed universally (looking at all reports in the dataset collectively), these data indicate that a large number of reports (83 percent) did not report any adverse safety consequences such as altitude deviations, aircraft damage, or passenger or crew injuries. Further, of those reports where information could be extracted concerning crew procedural issues, 88 percent revealed that crews followed prescribed procedures when faced with inflight aircraft malfunctions. However, unlike some businesses where above 80 percent may be considered "a passing score," aviation demands that safety margins be held to the highest values. Recently, the U.S. Secretary of Transportation held an industry-wide safety conference where he presented a challenge to industry to accept a goal of "zero accidents" (U.S. Department of Transportation, 1995). There is certainly room for improvement regarding crew performance during aircraft malfunctions.

Chi-square analysis revealed a highly significant difference (at the .05 level) between Type A and B malfunction categories in the areas of crew procedural issues, simultaneous error chain clues and adverse safety consequences (altitude and course/track/heading deviations.) This provides strong statistical indications that Type A and Type B malfunctions are different populations. The extremely small chi-square probabilities derived indicate that the type of malfunction experienced may be related to adherence to procedures, and other as yet unknown factors.

Additionally, a much less sophisticated look at raw numbers also points to some interesting observations. Merely totaling the number of error chain clue citations (Figure 1) shows that there were 25 citations for the 105 Type A malfunction reports, but 125 citations existed for the 112 Type B reports. Totaling the number of citations for adverse safety consequences also shows similar results (Table 5). For the 105 Type A reports there were 3 citations, while the 112 Type B reports had 46 adverse safety consequence citations.

The research team suggests that the widespread differences between these categories may be due to crew perception of the malfunction, as well as training. When faced with major mechanical malfunctions such as engine fires or complete loss of major aircraft systems, crews typically resort to highly practiced rules-based procedures, use of CRM principles, and some degree of heightened awareness. From analysis of this research, we theorize that the way a crew perceives a mechanical malfunction to some extent determines the way they will deal with the problem; i.e., serious problems demand a high degree of procedural usage and crew coordination, whereas less serious problems pose little threat so they can be handled less formally.

Rasmussen's skill-rule-knowledge (SRK) classification of human performance can be used to further explain these differences in crew performance (Rasmussen, 1993). Clearly, the majority of Type A malfunctions could be resolved by rules-based behavior, i.e., at the indication of an engine fire, crews should accomplish the following by immediate recall: thrust lever - closed, start lever - cutoff, engine fire handle - pull, engine fire bottle - discharge. Conversely, the nature of many of this study's Type B malfunctions had resolution procedures that were not as clear, and therefore may have required crews to revert to knowledge-based behavior. This level of behavior can require individuals to devote great amounts of time and effort to properly assess and resolve the situation. On occasion, this refocusing of tasks can result in reduced levels of procedural accomplishment, communications and situational awareness.

Apart from Line-Oriented Flight Training (LOFT) simulations, training and check flights almost invariably involve handling of major malfunctions, but have much less involvement with less serious malfunctions. We therefore suggest that emphasis be placed on enhancing the crew understanding that procedural issues and CRM principals need to be employed when dealing with less serious malfunctions, just as they need to be used when dealing with serious problems. We further recommend that developers of simulator training programs recognize the importance of simulating both serious and less serious malfunctions.

Error chain clues can denote reductions in, or loss of, situational awareness. We identified at least one error chain clue in 73 of the 230 reports. Fixation, distraction, no one flying the aircraft and work overload were found in a number of these reports, and are of particular concern because they have been identified in many fatal aircraft accidents. In reports in this study, we noted with consistency the tendency of crewmembers to become absorbed with resolving the malfunction, often at the expense of proper aircraft control. Stated one reporter, "No doubt flying the aircraft is the most important thing. We paid too much attention to a problem and forgot the most important thing - fly the airplane...One [person] should fly the airplane at all times, while the other crewmember solves the problem." (ASRS Record No. 124063) To minimize the possibility of such future occurrences, we recommend that crews practice controlling their FATE. We further recommend that flight crew training emphasize that an aircraft malfunction can serve as an immediate "red flag" to crewmembers, marking an occasion for possible loss of situational awareness.

FATE
Fly the aircraft
Assess the situation
Take appropriate action
Evaluate the results.

Many air carriers utilize LOFT scenarios to allow crews to practice and critique their CRM skills, often during simulations of aircraft malfunction resolution. Several of the ASRS reports reviewed in this study provided a wealth of information concerning problems encountered by crews dealing with malfunctions. Such reports are readily available to developers of LOFT scenarios to help incorporate *real* scenarios that have caused *real* problems for *real* crews.

One in ten of the 230 reports in this study provided evidence of crews using improper actions, such as not completing a checklist due to rushing, using the wrong checklist, and activating the wrong system control switches. Similar situations can be prevented by insisting that all crewmembers verify intended actions *before* initiation.

Although crew coordination and verification are topics usually stressed in training, furnishing crews with these findings may help provide insight that when faced with "the heat of the battle," crews may react in a manner contrary to training. If awareness is the first step toward behavioral change, then arming crews with this knowledge may better prepare them to avoid making these same mistakes.

Of the 88 reports that described whether or not crews declared an emergency, 9 indicted that the emergency was declared after a delay of some length. In two cases the reporters were forced to make this declaration at an inopportune time, because ATC did not fully appreciate the nature of their problem. It is commonly accepted that there exists a wide-spread reluctance within the pilot community to declare an emergency. Often cited reasons for failure to declare an emergency are "not wanting to fill out paperwork," and in general, not wanting to receive focus of attention from regulatory authorities or company management. It should be stressed with crews that the mere act of *declaring* an emergency does not, in itself, generate the automatic requirement to complete paperwork.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to the following persons for their assistance in making this project possible: Sherry L. Chappell, NASA/ASRS Research Director, for her technical expertise and input during the development of the coding instrument, and during the drafting of this paper; Loren J. Rosenthal, Battelle Program Manager for the ASRS, for supporting this project through funding and allocating necessary resources to assure its completion; and a very special thanks to Dr. F. Rowena Morrison, Battelle/ASRS Research Coordinator, for overseeing this project on a day-to-day basis, and who, along with Dr. Nancy Hassig of Battelle's Pacific Northwest Laboratory (PNL), assisted greatly with statistical analysis. The authors also wish to thank Gianna C. Siddens of Battelle/ASRS for her assistance with data analysis, and with the formatting of this paper.

REFERENCES

Chappell, S.L. (1994). "Using voluntary incident reports for human factors evaluation," in N. Johnson, N. McDonald and R. Fuller (eds), *Aviation Psychology in Practice,* Aldershot, England: Ashgate.

Foushee, H.C., Lauber, J.K., Baetge, M.M. and Acomb, D.B. (1986), *Crew factors in flight operations: III. The operational significance of exposure to short-haul air transport operations.* (NASA Technical Memorandum 88322), Moffett Field, CA: NASA.

Foushee, H.C. and Manos, K.L. (1981), "Information transfer within the cockpit: Problems in intracockpit communications," in C.E. Billings and E.S. Cheaney (eds), *Information Transfer Problems in the Aviation System,* (NASA Technical Paper 1875), Moffett Field, CA: NASA.

Schwartz, D.C., (1986), "CRM training for FAR parts 91 and 135 operators," in H.W. Orlady and H.C. Foushee (eds.), *Cockpit Resource Management Training. Proceedings of the NASA/MAC Workshop,* (NASA Conference Publication 2455), Moffett Field, CA: NASA.

Rasmussen, J., (1993), "Deciding and doing: Decision making in natural contexts," in G.A. Klein, J. Orasanu, R. Calderwood and C.E. Zsambok (eds), *Decision Making in Action: Models and Methods*, Norwood, NJ: Ablex.

U.S. Department of Transportation (1995), *Zero Accidents...A Shared Responsibility.* Washington, DC: DOT/FAA.

Wiegers, F.A. and Rosman, E.I., (1986), "Safety information system - A safety profile of wide-body commercial aircraft," (Douglas Paper 7780), *Proceedings of the International Air Safety Seminar,* Vancouver, BC: Flight Safety Foundation.