Errors in Aviation Decision Making: A Factor in Accidents and Incidents

Judith Orasanu
Lynne Martin
NASA-Ames Research Center

This paper is concerned with errors in decision making, specifically in the aviation domain. Despite efforts to design systems and procedures to support 'correct' and safe operations, errors in human judgment still occur and contribute to accidents. In their analysis of 37 accidents where crew behavior was a causal factor, the National Transportation Safety Board (NTSB, 1994) concluded that tactical decision errors contributed to 25 of the accidents, or about two out of three cases.

Our discussion will address three issues: (1) What is the nature of decision errors in high-risk, engineered environments like aviation? (We consider aviation decision making to be a kind of "naturalistic decision making", Klein, Orasanu, Calderwood & Zsambok, 1993; Zsambok & Klein, 1997). (2) What factors contribute to decision errors in aviation? (3) What kinds of technologies might mitigate aviation decision errors?

I. Decision Errors in Naturalistic Contexts

Decision errors in naturalistic contexts like aviation typically are not slips[1] or lapses in carrying out an intention, but errors of intention itself (Norman, 1981). The decision maker acts according to his/her understanding of the situation, and the source of error is in the decision maker's knowledge base or in the process of reaching a decision. There are two major problems in identifying decision errors in naturalistic contexts. First, there often is no clear standard of "correctness", e.g. whether to land with a fault in a piece of non-critical equipment. The "best" decision may not be well defined, as it usually is in a highly structured laboratory task. Second, there is a loose coupling of event outcome and decision process, so that outcomes cannot be used as reliable indicators of the quality of the decision, e.g. in windshear conditions where one aircraft may land safely but the next be affected by windshear. Redundancies in the system can "save" a poor decision. Conversely, even the best decision may be overwhelmed by events over which the decision maker has no control, resulting in an undesirable outcome. These occasions may be labeled as pilot decision error, but as Woods et al (1994) point out, this is hindsight bias - the tendency to define errors by their consequences. So, how do you know when a "good" decision has been made or an error committed?

To address this question requires a brief discussion of the nature of "naturalistic decision making", and decision making in aviation in particular.

Broadly, naturalistic decision making (NDM) describes decision making by individuals with some level of domain expertise[2] in real world contexts (e.g., aviation, nuclear power, offshore oil process control, fire fighting, command and control). These contexts typically involve limited time, dynamically changing conditions, goal conflicts, and information sources of varying reliability. Decision makers may operate in team and organizational contexts, and have available tools or other information resources available to aid their decision making (Orasanu & Connolly, 1993). NDM typically involves recognizing that a problem exists and sizing up the situation to define the nature of the problem and relevant factors. A candidate solution is retrieved, evaluated and applied if it meets a criterion of adequacy (Klein, 1997). There is not usually an exhaustive evaluation of all options.

Satisficing decisions

In naturalistic situations, decision makers tend to "satisfice", or to choose an option that will meet their goal (Simon, 1957). In a skilled domain such as aviation, this is far from a "pot luck" exercise, as the decision maker is an expert in the domain and applies his/her knowledge to decision situations. A traditional definition of decision error shows naturalistic decisions to be non-optimal when compared with a normative mathematical model. Often naturalistic situations are not amenable to optimizing strategies. It may not be appropriate to apply a model such as Multiattribute Utility Theory (MAUT) because normative approaches require time and large computational capacity to thoroughly evaluate all options, neither of which may be available to the decision maker. Manipulating large amounts of data is not consistent with human information processing capability, especially under stress (Stokes et al, 1992). People work within a 'bounded rationality' (Simon, 1957); satisficing is a function of this.

From the perspective of bounded rationality, apparently irrational, biased or heuristic approaches to decision making are often an individual's response to the computational complexity of a decision task. In fact, this approach is not necessarily defective, given the constraints of the situation (Cohen, 1993; Payne, Bettman & Johnson, 1988). Many assumptions are built into normative theory, including that people will always seek to optimize outcomes, that their preferences are stable, and that information and options are known. These assumptions are not always valid in real world decision making (Brehmer, 1987) and typically ignore the contributions of the decision maker's expertise and flexibility.

Decision processes and errors

Considering the nature of decision making in naturalistic contexts, the question remains whether it is reasonable to say the decision maker made an error or simply did not select the best solution, e.g. choosing to attempt a landing in bad weather conditions rather than going around. Since the outcome of a decision cannot necessarily be used to identify error, it is more appropriate to look at the process by which decisions are made (Woods, et al., 1996).

Orasanu and Fischer (1996) described a decision process model for aviation that involves two components: situation assessment (SA) and choosing a course of action (CoA). We will use this rough process model as a frame for considering decision errors. Situation assessment involves defining the problem as well as assessing the levels of risk associated with it and the amount of time available for solving it. Once the problem is defined, a course of action must be chosen. The course of action is selected from the options available in the situation. Building on Rasmussen (1986), Orasanu and Fischer proposed three types of option structures: rule-based, choice and creative. In rule-based decisions, there is a single prescribed action to take in response to a particular condition. Once the problem situation is recognized, the solution should be evident to someone with domain expertise, e.g. icing conditions. The second type, choice decisions, involve multiple options with legitimate trade-offs between them, e.g. selecting alternative airports. Choices depend on prevailing goals and constraints. The third type, creative, pertains to situations in which no suitable options are readily available and the decision maker must invent at least one to meet the demands of the situation, perhaps using analogies to similar situations.

Thus, there are two major ways in which error may arise. People may (a) develop a wrong interpretation of the problem, which leads to a wrong decision because they are solving the wrong problem -- an SA error, or (b) establish an accurate picture of the situation, but choose the wrong course of action -- a CoA error.

Situation assessment errors can be of several types: situation cues may be misinterpreted, misdiagnosed, or ignored, resulting in a wrong picture; risk (threat or danger) levels may be misassessed (Orasanu, Dismukes & Fischer, 1993); or the amount of available time may be misjudged (Orasanu & Strauch, 1993).

Errors in choosing a course of action may also be of several types. In rule-based decisions, the appropriate response may not be retrieved from memory and applied, either because it was not known or because some contextual factor mitigated against it. In choice decisions, options also may not be retrieved from memory, or only one may be retrieved when in fact multiple options exist. Constraints or factors that determine the adequacy of various options may not be retrieved or used in evaluating the options. Finally, the consequences of various options may not be considered. The decision maker may fail to mentally simulate the possible outcomes of each considered option. Creative decisions may be the most difficult because they involve the least support from the environment. The absence of available options means candidate solutions must be invented to fit the goals and existing conditions.

Using this framework, we see that the confidence with which errors can be identified may depend on the type of decision structure and the component in which the error lies. Mistakes in situation assessment may be more clearly classified as errors than mistakes in choosing a course of action, since in principle it is possible to know the "correct" definition of a problem. Risk or threat levels in principle are knowable, as is the time available to make a decision. Confidence in identifying course of action errors varies depending on the determinacy of the options. In rule-based decisions, it may be possible to say that an error occurred because the correct response was not applied (assuming the situation was correctly assessed) e.g. being over take-off weight. However, in choice situations, defining the best option may be more problematic, e.g. one alternative airport may have the maintenance facilities you require whilst another would provide better connections for your passengers. These are cases in which satisficing may rule. Here, it may only be possible to say that the process was faulty if the decision maker did not consider readily available options, failed to take into account relevant constraints, or did not project consequences of the various courses of action. These would constitute process errors. In creative decisions, it is also difficult to say that an error occurred since no options exist. Any solution that met a goal and major constraints would have to be considered a success, even if other better solutions might be generated in hindsight. If no response at all were made, then we might call it an error.

Our analysis is supported by Klein's (1993) analysis of decision errors across a variety of naturalistic domains. He found three sources of error: lack of experience, lack of information, and inadequate simulation. Lack of experience may contribute to situation assessment errors if the decision maker does not have the right knowledge to construct an accurate representation. Experience can also influence availability of response options. Information errors are those in which the decision maker does not have enough information to make a decision--technically not cognitive errors. The third source of error is inadequate mental simulation which may lead to 'wrong choice' errors. Failure to simulate outcomes associated with various courses of action may lead to poor choices. This is a process error. Thus, Klein's content analysis supports our suggestion that there are two basic classes of error in naturalistic situations.

II. What Factors Contribute to Decision Errors?

To explore factors that contribute to decision errors, we examined cases in the NTSB's (1994) set of 37 "crew-caused" accidents that involved "tactical decision errors". A common pattern was the crew's decision to continue with their original plan when conditions suggested that other courses of action might be more prudent[3]. In other words, they decided to "go" in a "no go" situation, usually in the face of ambiguous or dynamically changing conditions, e.g. continuing with a landing when it might have been more appropriate to go-around. Four factors are hypothesized as possible contributors to these decision errors:

- The situations were not recognized as ones that should trigger a change of course of action, due to the ambiguity of the cues;
- Risk was underestimated;
- Goals conflicted (safety vs. productivity, mission completion or social factors);
- Consequences were not anticipated or evaluated.

A. Ambiguity: Cues that signal a problem are not always clear-cut. Conditions can deteriorate gradually, and the decision maker's situation assessment may not keep pace. If events occur infrequently, the decision maker may not have amassed the experience to recognize the signals associated with a different CoA. Flight crews are typically in a "go" frame of mind. A substantial weight of evidence is needed to change the plan being executed. Weather and certain system malfunctions can change dynamically and pose a challenge for situation assessment. For decisions that have consequences, such as rejecting a takeoff or diverting, the decision maker needs to justify a course of action that may entail a cost. If the situation is ambiguous, the decision is harder to justify than if the situation is clear-cut, which may work against a decision to change the CoA.

B. Underestimating risk: When faced with problems, crews typically assess the level of threat or risk associated with the situation, both immediately and down the line. If somewhat similar risky situations have been encountered in the past and the crew has successfully taken a particular course of action, they will expect also to succeed this time with the same CoA, e.g. landing at airports where conditions frequently are bad, for example in Alaska. Given the uncertainty of outcomes, in many cases they will be correct, but not always. Reason (1990) calls this "frequency gambling". Schuch (1992) investigated mid air collision accidents (MAC) which are interesting because they tend to involve experienced pilots. He concluded that because experienced pilots have made repeated flights without an incident they become desensitized and stop scanning the sky.

Hollenbeck et al. (1994) found that past success influences risk taking behavior. Baselines become misrepresented over time as a situation becomes familiar and the individual becomes more experienced (Schuch, 1992). Sitkin (1992) argued if you only have good experiences you have no baseline by which to determine when the situation is taking a turn for the worse.

A second factor that may increase apparent tolerance of risk is framing. Framing studies have found that people tend to be risk averse in gain situations, while risk seeking in loss situations (Kahneman & Tversky, 1984). This raises the question of whether deteriorating situations that imply a change of plan, for example, to divert or go around, are seen as loss situations and therefore promote risk seeking behavior.

C. Goal conflicts: Organizational factors emphasize productivity, e.g. on time arrivals and departures, or saving fuel, which may conflict with safety[4]. Pilots may be willing to take a risk with safety (a possible loss) to arrive on time (a sure benefit). Social factors are also influential: among pilots peer pressure may encourage risky behavior. Meeting organizational and social goals often appears to outweigh safety goals, especially in ambiguous conditions.

D. Consequences not anticipated: As situations degrade, risk and time pressure may rise. These conditions may limit the decision maker's ability to project the situation into the future and mentally simulate the consequences of a course of action. Stress can interfere with the retrieval of multiple hypotheses, and constrains working memory capacity, thus limiting evaluation of options (Stokes et al., 1992). Under stress, decision makers fall back on their most familiar responses, which may not be appropriate to the current situation. (Because abnormal events tend to be quite infrequent, the correct responses may not be familiar.) Stress may interfere with evaluation and recognition of the inappropriateness of wrong responses, e.g. shutting down the wrong engine.

Whilst these four factors highlight process errors, they were identified from NTSB accident reports, where the adverse outcome of the flight indicated the presence of an error. The question whether process errors can be identified in the absence of a negative outcome therefore remains unanswered. However, a similar analysis of ASRS incident reports may remedy this. The cues and processes identified from accident events could be traced and matched across other non-accident situations.

III. Support for Improved Decision Making

Having suggested some process errors inherent in reported aviation accidents, the issue becomes how to deal with these conditions and aid pilots to make better decisions. Is it possible to design systems that support decision makers in complex dynamic situations where we know humans are subject to decision "error"? What kind of support would be most helpful? Technologies to assist decision makers have not always had their focus on the decision maker and the areas in which they have difficulty. Decision aids may benefit from an interpretation in this light on areas of likely errors.

Wiener (1993) separates technologies for reducing human error into mechanistic and behaviorally based interventions. Mechanistic interventions include hardware interlocks, such as gust locks on aircraft controls, or software traps; behavioral ones include training, check-lists and procedures. While mechanistic approaches may be more reliable, behavioral ones are more adaptable. If we think in terms of the two major decision components, situation assessment and choosing a course of action, we can begin to suggest the kinds of aids that might be helpful under these two categories.

Assisting SA

Given that accident analysis suggests that a major impact could be made on decision making by improving situation assessment so pilots recognize situations that require revision of their CoA, what are the essential components? To improve pilots' understanding of their problem requires better diagnostic information and more accessible, comprehensible, integrated displays that show trends (see Flach et al. 1994, Norman, 1986). These improvements apply most clearly to system diagnostics (in which there has been great improvement already), providing more up to date information on unpredictable, dynamic attributes of the situation. It is more difficult to present weather information in the integrated format required to show trends, although this is not impossible as the new terminal area weather displays demonstrate (e.g. ITWIS, MIT/LL). However, weather information should highlight information critical to pilots depending on their phase of flight. Traffic displays are also being improved, especially in anticipation of "free flight" and will be integrated with weather information (e.g. NASA's AATT project).

A second avenue is to improve decision makers' experience by giving them better training: if they have a large number of exemplars to choose from, they will be able to select a model which more closely fits the problem. The importance of having a choice of exemplars in memory is illustrated by Klein (1997), amongst others, who notes a distinguishing feature of chess masters is the large number of board positions they have committed to memory.

SA will also be assisted by addressing two components of naturalistic situations: the risk and time available. It is difficult to embed information about risk and time in displays, since these often depend on the context. Aircraft system information, perhaps, could have risk estimates attached. Weather and traffic displays already have some risk information (color coding of weather severity; TCAS warnings for traffic). Aids providing judgments of time available for making a decision or taking action would require predictive models. They could indicate how long it will take for a condition to degrade to a critical state, e.g. fuel consumption or reserve battery life span, how soon weather will improve or a storm hit, or when traffic will dissipate in target region. Again, contextual factors are likely to make such predictions difficult.

Assisting the pilot's SA addresses two of the error sources identified by Klein (1993), those of lack of information and experience. Information can be provided at the point of decision in enhanced displays which are user-centered. Experience can be built through directed training which presents patterns of developing process errors to pilots, enabling them to recognize the trend of a sequence of events and hence act to prevent an incident occurring.

CoA Assistance

The second decision making component to aid is choosing a course of action, and with this the third source of error - inadequate simulation. To do this, aids may present options, constraints and the likelihood of outcomes. At the moment, flight manuals typically list options for dealing with system malfunctions. For example, when certain kinds of engine anomalies are encountered the conditions under which it would be most prudent to shut down the engine, reduce power to idle, or leave it running are described. Aiding may consist of prompting crews to consider options prior to jumping to action, to consider the disadvantages of the selected option, and the likelihoods of various outcomes. Klein (1997) included such recommendations in a training model for military decision making. Tools for doing "what if" reasoning and managing multiple hypotheses may also be helpful, encouraging forward thinking. Means et al (1993) stress the importance of making decision makers aware of the worst case scenario and training them to manage this.

It will be more difficult to present risk ratings for goal conflict situations. We know that people are notoriously poor at integrating numerical probabilities, faring better when information is presented as a graphical representation. Context-sensitive estimates of event likelihoods would be needed to enable trade-offs to be made.

Conybeare (1980), and others, have suggested that people recalculate the risks involved under a new system and compensate their behavior to bring the level of risk back to its previous level. This suggests decision aids may not have a lasting effect. However, McKenna (1985) refutes this homeostasis theory on a number of counts, among these are that people are unable to calculate risk probabilities in this way and that non-obvious safety measures could not be accounted for by the individual. Whilst agreeing with McKenna, there is anecdotal evidence that pilots (like motorists) will 'push the envelope' when new measures are introduced, altering their behavior to exploit the increased functionality of the systems, e.g. driving faster when safety belts became mandatory. An example of this is possibly the Airbus 320 crash as the Habsheim airshow. New systems may lead the pilot to change the way they evaluate options, or the weights they assign to attributes. This may not necessarily lead to more risky behavior but is likely to lead to a change in behavior and the process of pilot decision making.

Assisting pilots with decision making elements that are difficult for humans, may help them to avert potential incident. However, these measures do not guarantee that every incident situation will be caught. Changing the system in an effort to catch errors may also alter it in a way to change pilot behavior within it, the source of the error may therefore change. Thus, pilot decision error may be a moving target which will have to be periodically redefined as the field advances.

While better displays are already in evidence and automation already controls many aircraft functions, we still are left with human judgment in cases that cannot be automated. The challenge of the future is to provide support for the weak links in human decision processes, while exploiting the strengths and adaptability of the human agent through specifically designed tools and training.

References

Brehmer, B. (1987). The psychology of risk. In W.T. Singleton & J. Hovden (Eds.) *Risk and Decisions* (pp. 25-39). New York: John Wiley & Sons.

Cohen, M.S. (1993). The naturalistic basis of decision biases. In G. Klein, J. Orasanu, R. Calderwood & C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 51-99). Norwood, NJ: Ablex.

Conybeare, J. (1980). Evaluation of automobile safety regulation: the case of compulsory seat belt legislation in Australia. *Policy Sciences*, 12, 27-39.

Flach, J., Hancock, P., Caird, J. & Vicente, K. (Eds.) (1994). *Global Perspectives on the Ecology of Human-Machine Systems* (pp.1-13). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hollenbeck, J., Ilgen, D., Phillips, J. & Hedlund, J. (1994). Decision risk in dynamic two-stage contexts: Beyond the status-quo. *Journal of Applied Psychology*, 79(4), 592-598.

Jensen, R. (1998). Presentation at Training for safety in general aviation and air taxi operations: Research needs and opportunities. NASA-Ames Research Center, Moffett Field, CA. January 1998.

Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4) 341-350.

Klein, G. (1993). Sources of error in naturalistic decision making. In *Proceedings of the Human Factors and Ergonomics Society* 37th Annual Meeting, 1, 368-371. Santa Monica, CA: Human Factors and Ergonomics Society.

Klein, G. (1997). The current status of the naturalistic decision making framework. In R. Flin, E. Salas, M. Strub & L. Martin (Eds.) *Decision Making Under Stress: Emerging Themes and Applications*. Aldershot, UK: Ashgate.

Klein, G., Orasanu, J., Calderwood, R. & Zsambok, C. (Eds.). *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

McKenna, F. (1985). Do safety measures really work? An examination of risk homeostasis theory. *Ergonomics*, 28(2), 489-498.

Means, B., Salas, E., Crandall, B. & Jacobs, T.O. (1993). Training decision makers for the real world. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 51-99). Norwood, NJ: Ablex.

National Transportation Safety Board. (1994). A Review of Flightcrew-Involved, Major Accidents of U.S. Air Carriers, 1978 through 1990. (PB94-917001, NTSB/SS-94/01). Washington, DC: Author.

Norman, D.A. (1981). Categorization of action slips. *Psychological Review*, 88, 1-14.

Norman, D.A. (1986). Cognitive Engineering. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design* (pp. 87-124). Hillsdale, NJ: Lawrence Erlbaum Associates.

Orasanu, J. & Connolly, T. (1993). The reinvention of decision making. In G. Klein, J. Orasanu, R. Calderwood & C. Zsambok (Eds). *Decision making in action: Models and methods* (pp. 3-20). Norwood, NJ: Ablex.

Orasanu, J., Dismukes, K. & Fischer, U. (1993). Decision errors in the cockpit. In *Proceedings of the Human Factors and Ergonomics Society* 37th Annual Meeting, 1, 363-367. Santa Monica, CA: Human Factors and Ergonomics Society.

Orasanu, J. & Strauch, B. (1994). Temporal factors in aviation decision making. In *Proceedings of the Human Factors and Ergonomics Society* Meeting (pp. 935-939). Santa Monica, CA: Human Factors and Ergonomics Society.

Orasanu, J. & Fischer, U. (1997). Finding decisions in natural environments: The view from the cockpit. In C. Zsambok & G. Klein (Eds.) *Naturalistic Decision Making*. (pp. 343-357). Hillsdale, NJ: Lawrence Erlbaum Associates.

Payne, J.W., Bettman, J.R. & Johnson, E.J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 14 (3), 534-552.

Pidgeon, N. (1988). Risk assessment and accident analysis. In B. Rohrmann, L. Beach, C. Vlek & S. Watson (Eds.) *Advances in Decision Research* (pp. 355-368). New York: Elsevier Science Publishers.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on Systems, Man and Cybernetics*, 2, (SMC - 15) 234-243.

Reason, J. (1990). *Human Error*. Cambridge, UK: CUP.

Schuch, H.P. (1992). The influence of flight experience on midair collision risk perception. *Accident Analysis & Prevention*, 24(4), 655-660.

Simon, H. (1957). *Models of Man*. New York: Wiley.

Sitkin, S. (1992). Learning through failure: The strategy of small losses. *Research in Organizational Behavior*, 14, 231-266.

Stokes, A.F., Kemper, K.L. & Marsh, R. (1992). Time-stressed flight decision making: A study of expert and novice aviators (Tech. Rep. No. 4). Arlington, VA: Office of Naval Research.

Wiener, E.L. (1993). Intervention Strategies for the Management of Human Error (NASA Contractor Report No. 4547). Moffett Field, CA: NASA Ames Research Center.

Woods, D.D., Johannesen, L.J., Cook, R.I. & Sarter, N.B. (1994). Behind Human Error: Cognitive Systems, Computers, And Hindsight. Wright-Patterson Air Force Base, Ohio: Crew Systems Ergonomics Information Analysis Center.

Zsambok, C. & Klein, G. (1997). *Naturalistic Decision Making*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Notes

1 "slips ... are errors which result from some failure in the execution ... of an action sequence, regardless or whether or not the plan which guided them was adequate to achieve its objective." (Reason, 1990, p9)

2 Expertise can be viewed as a combination of skills, judgment and knowledge (Jensen, 1998).

3 Of course, we have no way of knowing how often pilots chose to continue in highly similar circumstances with no negative consequences.

4 Organizational decisions about levels of training, maintenance, fuel usage, keeping schedules, etc. may set latent pathogens that undermine safety, in the face of vocal support for a "safety culture." (Reason, 1990).

jo/lm, hew2, !/ /#