PILOT ERROR:
COGNITIVE FAILURE ANALYSIS

A thesis presented in partial fulfilment of the
requirements for the degree of Master in Aviation at
Massey University

Dmitri Victorovitch Zotov

1997

ABSTRACT

Rasmussen (1982) suggested that there was a need for a taxonomy of human errors based on the operator performing the task, rather than upon the task itself; the "internal human malfunction" (p. 323). This proposal was adopted by O'Hare, Wiggins, Batt & Morrison (1994) in a study of pilot errors derived from the New Zealand official Accident Reports.

O'Hare et al. (1994) found differences in the types of errors that led to major and minor accidents. These differences were at variance with the proposition by Billings and Reynard (1981) that the errors in accidents and incidents came from a common population, the outcome being due to chance. The results of O'Hare et al. (1994) cast some doubt on the validity of investigating incidents as a means of forestalling accidents.

Some of the accident reports used by O'Hare et al. (1994) had not been the result of independent investigation, but were self-reports by the pilots involved. The inclusion of these reports had the potential to produce the apparent dichotomy between the distributions of error types in major and minor accidents, found by O'Hare et al. (1994). It was therefore decided to revisit their work, using as a database the entire population of New Zealand official Accident Reports since 1965, which had been the subject of official investigation.

With the large database available, variability in the distribution of error types was also examined between different classes of aircraft, and between pilots of different levels of experience.

Some variability between major and minor accidents was found, but not enough to be of practical significance. No variability was found between pilots of different levels of experience. There was little difference between classes of aircraft, except in the case of fixed-wing agricultural aircraft. In the latter case, the difference in the distribution of error types from other classes of aircraft was marked, and further study to identify the reasons might assist in reducing the accident rate for agricultural aircraft.

ACKNOWLEDGEMENTS

The author's interest in the application of human factors to aircraft accident investigation was encouraged by Dr. David O'Hare of Otago University. Dr. O'Hare worked with the author on practical applications of psychology in particular accidents, notably the accident to ZK-SUN at Napier, and subsequently assisted him in the development of theory in this area.

The author gratefully acknowledges the assistance of the Flight Safety Trust, who have sponsored this and other studies in accident investigation.

TABLE OF CONTENTS

Appendices

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

There have been many studies which have examined the causal factors of aircraft accidents, and without exception the dominant factor has been found to be pilot error. Skjenna (1981) stated that in his experience, 80 to 90% of Canadian helicopter accidents arose from pilot error, a somewhat surprising figure when taken in conjunction with the known propensity of these devices to mechanical failure (Zotov, 1995a). Boeing (1985) surveyed commercial air carrier operations involving jet aircraft over a 24 year period (1959 - 1984). They found that crew (pilot and flight engineer) errors were implicated in 67% of all accidents, while Caesar (1987) reported that 76 % of all jet hull-loss accidents in 1986 were attributable to crew error.

Nagel (1988) has suggested that for general aviation in the USA, almost 90% of accidents were attributable to human causal factors, but O'Hare, Wiggins, Batt & Morrisson (1994) found, after a review of New Zealand accident reports, that 71% could be attributed to pilot errors. The difference may in part be due to the more unforgiving environment in New Zealand, as compared to the USA.[1] However, it may also reflect the difficulty of devising an error taxonomy that represents all forms of human error (Wiegmann & Shappell, 1996).

---

[1] The proportion of the New Zealand land mass suitable for forced landing is very much less than is the case in the USA.

Overall, about three-quarters of all aircraft accidents may be due to pilot errors. In order to make a significant reduction in the accident rate, it is necessary to find the "Causes of causes" (Gerbert & Kemmler, 1986, p. 1439), that is, we must be able to understand why pilots err, in order to attempt to devise methods of preventing accidents.

Rasmussen (1982) has suggested that to understand human error it is necessary to know the type, nature and origin of the error - what happened, how it came about, and why the error was made. More recently, the systemic causes behind human error accidents (the 'why') have been extensively studied, particularly by Helmreich (1990) and Reason (1990, 1991). However, despite this progress, there is little understanding of the failures of the crews who are often the final link in the accident chain.

Rasmussen (1982) categorized errors into skill, rule, and knowledge-based behaviours. Errors of skill are made at the manipulative level, for example, in applying the correct pressure to the control column for the correct time to produce a desired change of aircraft attitude. Such errors are likely to be commoner at the learning stage. Rule-based errors occur when the pilot applies an already-learnt rule to solve a problem, either applying the right rule incorrectly, or applying the wrong rule. Knowledge-based errors occur when no ready-made rule is available to solve a problem, and the pilot must apply his/her knowledge of how the system works to find a solution.

This categorization allows an insight into 'how' the failure may have occurred, but as Rasmussen (1982) explained, this is a different concept from 'what' the error was. To examine the types of errors, he proposed a seven component taxonomy derived from an information-processing model.[2] Rasmussen's proposed model was based not on the task, but on the task as performed by the human, and was a major simplification on previous taxonomies. Nagel (1988) subsequently suggested that even a three-part model, of information - decision - action, might suffice.

O'Hare et al. (1994) reviewed the New Zealand Accident Reports, over a nine-year period (1983-92), to establish a database. Their factorial analysis found that Nagel's (1988) three-element taxonomy did not adequately describe their data, and they subsequently derived a six-element taxonomy.

In the course of their investigation, O'Hare et al. (1994) found that the errors associated with fatal and serious injury accidents were generally different from those which had resulted in minor/ nil-injury accidents. They found that fatal/serious injury accidents were associated with incorrect selection of goals (for example, deciding to proceed to the planned destination in deteriorating weather), and minor/ nil-injury accidents with procedural errors such as performing a sequence of operations in the wrong order. The need for emphasis on decision-making (i.e. choice of goal and strategy) for the reduction of the number of severe accidents is generally accepted, and training in decision-making now forms part of the basis of pilot training (Hunt & MacFarlane, 1994). However, if there is a difference in the dominant error types between major and minor accidents, this

---

[2] There have been many attempts to produce error taxonomies (Wiegmann & Shappell, 1996). Early task-based taxonomies were cumbersome, having a great many categories (eg BASI (1984) - 146 categories).

could have implications for the optimum allocation of resources in accident investigation.

Investigative resources are generally limited, both by the number of investigators and by funding, and this has led to a concentration of resources on the investigation of major accidents. This is necessarily a reactive process, in that the lessons are learnt after a disaster has occurred. It would be better, if possible, to avert disasters by learning the lessons beforehand. Investigations of minor accidents, or incidents (where no accident occurred, but there was the potential for one) may provide such opportunity for learning. Billings and Reynard (1981) reviewed the records of anonymous self-reported incidents in the American National Aeronautics and Space Administration (NASA) database, known as the Aviation Safety Reporting System (ASRS). They found similarities between ASRS incidents and accident records, which suggested that errors reported as incidents are similar to actions that cause accidents. They commented that aircraft accidents involving operational and human factors are subsets of populations of incidents that contain the same elements.

Moves to anticipate accidents by investigating incidents have recently culminated in an International Civil Aviation Organization (ICAO) directive requiring States to investigate serious incidents on the same basis as accidents (ICAO, 1994). However, if actions which cause minor accidents differ in character from those which cause major accidents, it is open to question whether actions which lead only to incidents are likely to be the same as those resulting in disasters. If O'Hare et al. (1994) are correct in asserting that major and minor accidents are different in kind, there is some doubt as to the validity of the ICAO directive.

There may be some problems of validity with the database used by O'Hare et al. (1994). The minor accidents which were analyzed were almost all reported in an abbreviated ('Brief') format. Many of these 'Briefs' were not the result of formal investigation, but were simply transcripts of the accident notifications submitted by the pilots. As such, they were subject to the normal limitations of self-reporting (e.g. Dane, 1990). Some of the 'major accidents' also came from this self-reported population. A further limitation arose from the use of the ICAO term 'serious injury', which could include some quite minor injuries such as cuts and broken bones.

It is possible that deficiencies in the database may have produced the differences in distribution of error types that O'Hare et al. (1994) observed in accidents of different degrees of severity.

Instead of classifying accidents as 'major' and 'minor', they could be reclassified as 'fatal' and 'non-fatal'. This would have the effect of removing accidents that were really not very serious from the 'major accidents' category.

The present author conducted a pilot study based on the work of O'Hare et al. (1994), using a database limited to accidents known to have been investigated (Zotov, 1995b), and the 'fatal/ non-fatal' categorization. While the distributions of different classes of pilot error were different from those found by O'Hare et al, the results of the pilot study supported the view that there was a significant difference in the types of errors that led to different outcomes.

Accordingly, it was decided to conduct a full-scale replication of this aspect of the work of O'Hare et al. (1994).

Object of the Study

The objective was to study New Zealand official reports of accidents which had been professionally investigated, in order to determine whether there was a difference between the distribution of error types, between fatal and non-fatal accidents. The work of O'Hare et al. (1994) was extended by examining accidents to different classes of aircraft, to see whether error distributions were consistent between them. Also, errors made by pilots with different levels of experience were examined.

HYPOTHESES

The research hypotheses were:

1. There would be differences in the distributions of types of pilot errors, with various classes of aircraft:

   Fixed-wing, agricultural and non-agricultural
   Helicopters engaged in venison recovery, and other helicopters
   Gliders
   Light sporting aircraft - microlights, hang-gliders and autogyros.

2. There would be differences in the distributions of types of pilot errors, between fatal and non-fatal accidents.

3. There would be differences in the distributions of types of pilot errors, with various levels of pilot experience.

CHAPTER 2
REVIEW OF THE LITERATURE

The literature relating to pilot error taxonomy will be reviewed under the following headings:

Task-related taxonomies

Performance-related taxonomies

Effect of pilot experience on distribution of error types

The relationship of major accidents, minor accidents and incidents

Comparison of taxonomies

Task-Related Taxonomies

Wiegmann & Shappell (1996) have commented that there are as many taxonomies of pilot error as there are people interested in the topic. This is because many taxonomies have been designed for a particular study.

An early example of a taxonomy of pilot error, devised for a specific study, appears in the work of Fitts & Jones (1947). They examined the errors found when interpreting aircraft instruments and in operating flying controls. The errors in interpreting instruments fell into nine categories, such as misreading multipointer displays, and scale interpretation. Control operating errors were grouped into six categories, such as substituting one control for another, or reversing the direction of operation.

A more general classification of the same type was devised by Jensen and Benel (1977). They examined the United States National Transportation Safety Board (NTSB) records of General Aviation accidents for the period 1970 to 1974, and classified all aircrew errors into three major categories, based on behavioural activities:

(i). Procedural errors, for example the management of subsystems, or configuration of the aircraft (e.g. retraction of the flaps instead of the undercarriage).

(ii). Perceptual/motor errors, i.e. the manipulation of controls (e.g. overshooting the glidepath indication, when joining the Instrument Landing System approach path).

(iii). Decision tasks, such as flight planning, and in-flight hazard evaluation (e.g. failing to delegate tasks in an emergency, or continuing flight into bad weather).

The results from that study showed that 56% of non-fatal accidents involved perceptual-motor factors, whereas for fatal accidents the dominant factors were decisional (52%). This is in line with the subsequent findings of O'Hare et al. (1994), which suggest a significant difference in the causal factors of major and minor accidents, and tends to run counter to the traditional view exemplified by Billings & Reynard (1981), that accidents are a subset of incidents.

Nagel (1988) criticised task-based taxonomies on the ground that they were merely descriptive of behaviour. He suggested that in order to develop solutions to the problem of error, it was necessary to predict with some certainty which flight conditions were most likely to contribute to accidents. By 'predict' Nagel was suggesting that situations where certain types of error could be expected, recurred with significant regularity. In the 'action' domain, a well-known example is having the flap and undercarriage levers next to each other, and similar in shape and operation, leading to almost routine cases of pilots retracting the wheels instead of the flaps after the landing roll.

Task-based taxonomies are also both comprehensive, and complex. For example, the United States Navy (USN) aircraft accident database comprises 289 categories (Wiegmann & Shappell, 1996), and the Australian Bureau of Air Safety Investigation (BASI) database has 146 categories (BASI, 1984). They are so complex that coding errors are probable, so the accuracy of the database as a whole is dubious.

Performance-related Taxonomies

A new approach was introduced by Rasmussen (1980; 1982). He sought to devise a taxonomy based on a model of the person performing the task, rather than on a model of the task itself. He found little published work on generic psychological error mechanisms. In the earlier paper Rasmussen (1980) analyzed event reports from nuclear installations, and attempted to characterise human error in generic terms. In the later paper (Rasmussen, 1982), he outlined a multi-faceted classification system, dealing with three main variables:

(i). Why the human failed: the external work environment, and the task characteristics.

(ii). How the human failed: "the mechanism of human malfunction" (p. 325). This he defined in terms of errors at the skilled level of functioning, errors at the rule-based level, and errors of reasoning when there were no available rules (knowledge-based errors). These levels of behaviour are discussed below.

(iii). What failed: "the internal human malfunction" (p. 323).

The common analogy between machine error functions and human operators, where an input is perceived, mediated in some way, and then activates a motor function to produce an output, was held by Rasmussen (1982) to be unrealistic. For example, it does not show the selective filtering of errors due to the reversibility features of the task; that is, it does not take into account the ability of the human to detect errors and put them right. Also, it lacks the aspect of human intention and expectation.

In developing a model of the causes of human errors, Rasmussen (1982) distinguished between three levels of behaviour:

(i). Skill-based performance: that is, performance which is automated; more-or-less subconscious stored patterns of behaviour. Errors in this type of performance include variability of force, space or time coordination.

(ii). Rule-based performance: that is, performance in familiar situations, with stored rules for coordination of subroutines. Errors in this type of performance include the wrong classification or recognition of situations, or memory lapses in the recall of procedures.

(iii). Knowledge-based performance: that is, performance which is unique, and in unfamiliar situations, in which actions are planned from analysis and decision, based on knowledge of the system. Errors in this type of performance are defined in relation to the goal of the task, for example failures in functional reasoning, such as trying to determine how the machine will respond after a component has malfunctioned.

Rasmussen (1982) also sought a "general description of the internal mental function which was not performed as required by the external task" (p. 319). He stated that "in order to be able to identify the internal function which failed, on the basis of the external effects of the errors alone, this description must be independent of the level of human behaviour, and based alone on a rational breakdown of the decision sequence into the phases of detection, identification, decision etc". (p. 319). The decision sequence which he evolved is shown in Table 1.

Table I Decision Sequence for Identifying Human Errors (Rasmussen, 1982)

Action	**Type of Error**
Detection	Detection missing
Identification	Identification incorrect
Decision -	
Select Goal	Goal incorrect
Select Target State	Target state inappropriate
Select Task	Task inappropriate
Action -	
Procedure	Procedure incorrect
Execution	Execution erroneous

Analysing the sequence of events in this fashion may then show **what** failed, in terms of the internal human malfunction. (This analytical method was subsequently modified by Rouse and Rouse (1983) and by O'Hare et al. (1994), as will be discussed shortly).

The mechanism of human malfunction - **how** it failed - is a fundamentally different concept which should therefore be considered separately during event analysis.

Finally we need to know **why** the human failed, that is, the performance shaping factors such as company safety culture or inadequate training, and situation factors such as optical illusions. Rasmussen (1982) did not consider this field in detail, but it was addressed by Reason (1990, 1991) with his concept of latent failures. These are errors whose adverse consequences may not be felt for some time, in contrast to active failures, which take immediate effect.

Rasmussen's (1982) conceptual model formed the basis for practical application by Rouse and Rouse (1983). They sought insight for evaluating or modifying systems design and training programmes, by analysing errors in terms of causes as well as contributing factors and events. The algorithm which they developed (Fig 1) to represent a simplified view of the tasks of human operators in systems such as an aircraft, incorporated the assumptions that

(i). During normal operations, the human cycles through observing the system state, choosing procedures and executing procedures.

(ii) When an abnormality occurs (that is, variables are outside the normal range, or there is a warning) the operator resorts to problem solving, as in the algorithm.

This model, unlike Rasmussen's model (1982), does not presuppose any particular theory of human error: it is intended to be a practical tool relevant to the behaviour of human operators of complex man-machine systems.

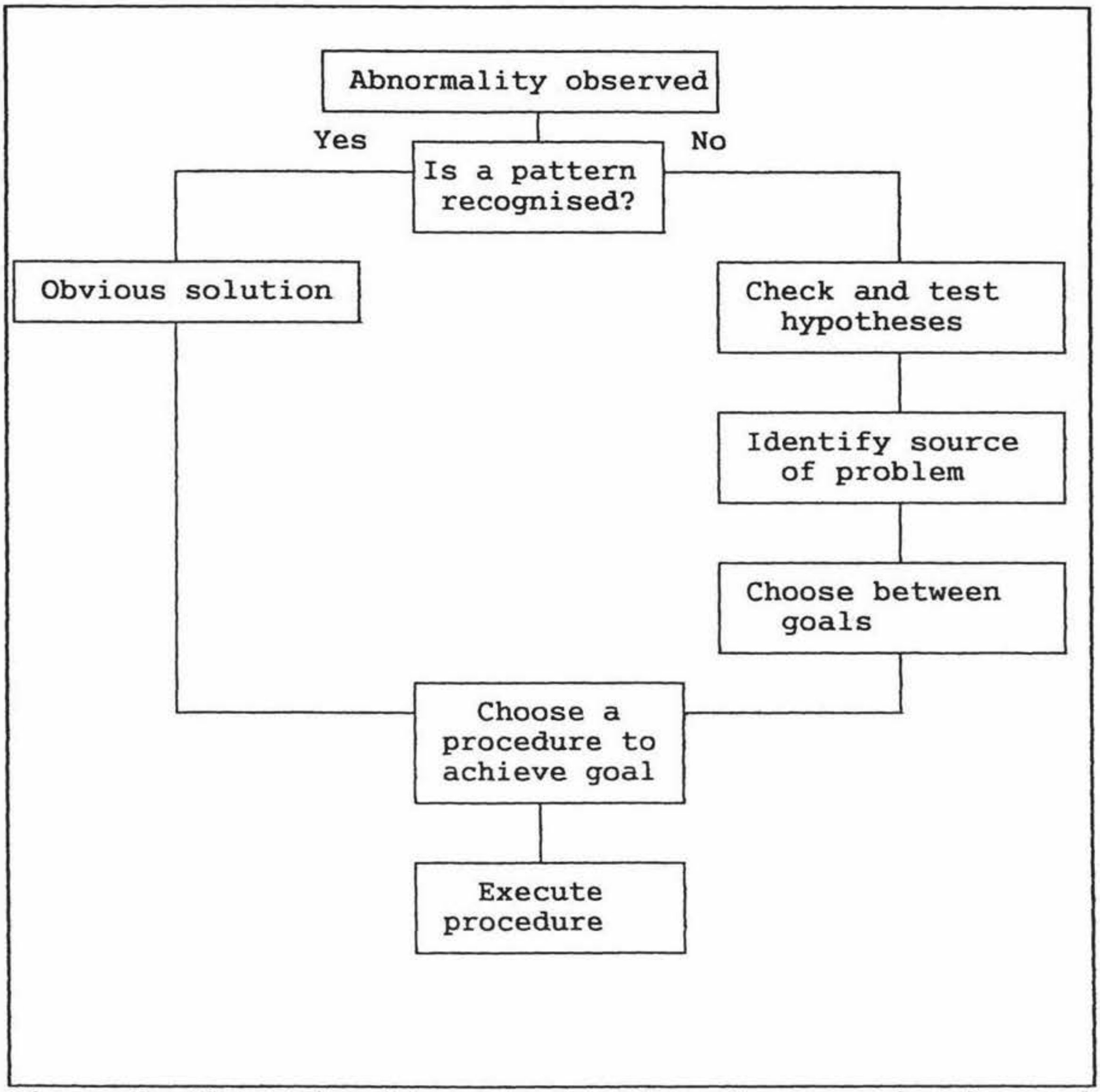

Figure 1 Algorithm for analysis of performance in abnormal conditions (after Rouse & Rouse, 1983)

Gerbert and Kemmler (1986) analyzed flight incident reports (on occurrences which the pilots considered to have had the potential to be accidents) submitted by 1448 German Air Force pilots. Their object was to discover when, and under what internal and external conditions, a restriction of capacity was experienced (for example, the inability to complete all tasks in the required time-frame), or an error was made; the hazard that resulted; and how it was corrected. They assumed a three-part model of performance, comprising attention/perception, cognition/decision, and implementation.

Gerbert and Kemmler (1986) assumed that there would be a factorial structure to errors, and examined sixty one error variables from the flight incident reports. The commonest twenty seven incidents were subjected to factorial analysis. From this, a readily-interpretable four-factor solution was found, namely:

(i). Vigilance errors (e.g. carelessness)

(ii). Information processing errors (e.g. wrong decisions, faulty action plan)

(iii). Perception errors (e.g. misjudgment of clearance from the ground or obstacles)

(iv). Sensorimotor/handling errors (e.g. non-application of procedures, poor coordination of controls).

These results applied to trained air force pilots, and so may not be directly comparable with similar results from general aviation pilots. The latter, generally, have a different level of training and flying experience compared with air force pilots. Also, because of the environments being dissimilar, the results may not be comparable with those which would be found in the case of airline pilots.

Another approach to understanding the function of the human in complex man-machine systems is the analogy between humans and computers, the 'information

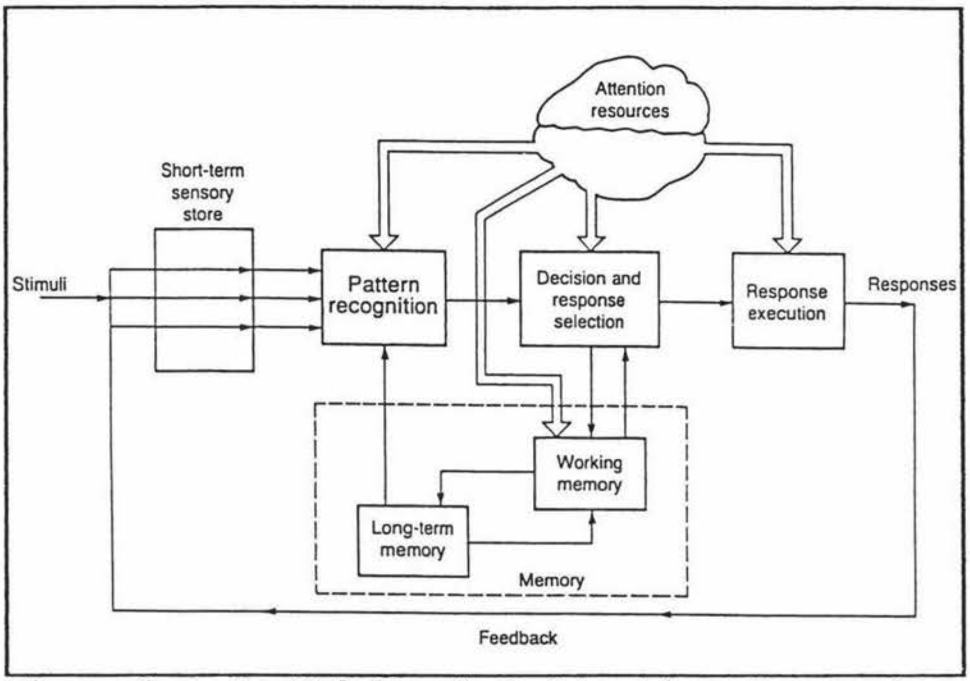

Figure 2 A Model of Information Processing. (Source: Wickens and Flach, 1988).

processing' approach. Such models assume that a series of stages occur between some stimulus such as sight or hearing, and the response to that stimulus; errors can occur at each stage. In the four-part model used by Wickens and Flach (1988) (Fig 2), the stimulus enters a short-term sensory store and then passes through pattern recognition where it is compared with patterns held in long term memory before being processed for decision-making and response selection using the limited working memory. Finally, a response execution signal is generated and sent to appropriate organs (voice, limbs etc).

Nagel (1988) proposed a very simple error model, also derived from information processing considerations. Actions in the cockpit proceed from gathering **Information**, to making **Decisions** based on the information, to **Actions** which implement those decisions; at each phase, errors can occur. The assumptions underlying this model are *rational intention* (not to commit errors) and *goodwill* (irresponsible behaviour is discounted); also that crews are *highly skilled and experienced,* so that high-level goals are appropriate.

For examining general aviation accidents, these assumptions may not be altogether valid. Irresponsible behaviour cannot be discounted: accidents arising from exhibitionism are not unknown. Also, unwise selection of goal (for example, proceeding to the planned destination in the face of deteriorating weather) is a common causal factor in general aviation accidents. Nagel's simplified model is thus restricted to air transport operations.

The taxonomies reviewed so far have centred on the actions of the pilots, but those actions are influenced by the systems within which the pilots operate, such as the company and regulatory environments.

The systemic causes of accidents (the 'why') were examined by Helmreich (1990). He reviewed the actions of the crew in a Fokker F28 accident at Dryden, Ontario, to assist the Commission of Inquiry held by Mr Justice Moshansky (Moshansky, 1992). The puzzling feature of the accident was that a very experienced crew had attempted to take off with thick snow on the wings.

In examining why the crew should make such a fundamental mistake, Helmreich (1990) looked at the various factors which could influence the crew or put them under pressure. He considered the various environments within which the crew had to work (which were envisaged as concentric spheres of influence): the crew itself, with interactions between the various members; the physical environment, e.g. the cockpit and weather; the company, which bought the aircraft, trained the crew, and had a responsibility to support their actions; and the regulatory authority, which should influence company and crew to operate in a safe manner (Fig 3).

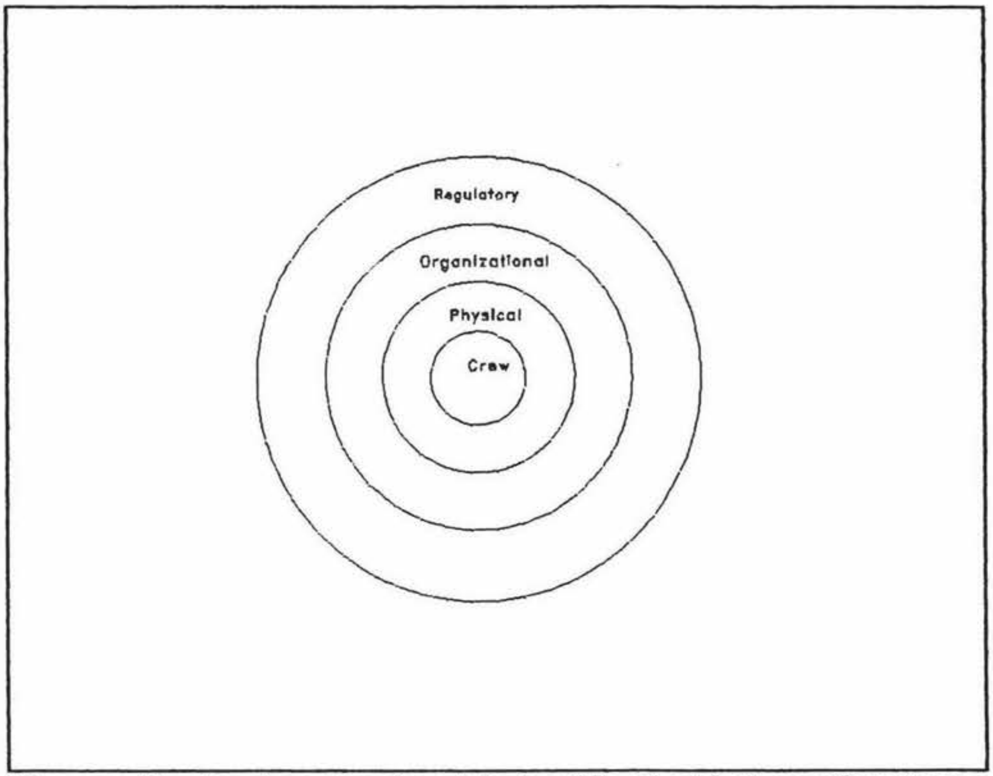

Figure 3 Environments which affect pilot performance (Source: Helmreich, 1990)

Helmreich (1990) found deficiencies in each of these environments. For example

(i). The crew were on their first trip together, a factor known to militate against crew effectiveness (crew environment)

(ii). The aircraft was not allowed to be de-iced with engines running (physical environment)

(iii). The company had despatched the aircraft to Dryden with unserviceabilities that would have precluded restarting the engines if both were stopped (organizational environment)

(iv). The regulatory authority had not produced guidelines which could be referred to by operators to determine how much snow on the wings was acceptable for take-off (Regulatory environment).

There was a long list of such items none of which, in itself, would have caused the accident. However, the combination of factors brought such pressure to bear on the crew that they made the wrong decisions. Helmreich (1990) concluded that such a series of factors, without any single proximal cause, merited the description of a *systems accident.*

An alternative approach to understanding systems accidents was developed by Reason (1990). He devised a model which incorporated the variable of unsafe acts, a different approach from the information processing models previously favoured. He considered that behaviour was divided into 'intentional' and 'unintentional' actions.[3] Unintentional acts result either from memory lapses, or attentional slips. Intentional acts that are erroneous may be either mistakes or violations: mistakes occur when previously learned rules or procedures are either misapplied (rule-based) or non-existent (knowledge based) (c.f. Rasmussen, 1982). Violations are intentional departures from rules or regulations.

Reason's conceptual framework, which he called the "generic error modelling system" (Reason, 1990, p. 53) was derived largely from Rasmussen's (1982) skill-

---

[3] Note that it is the act, not the underlying error, that is intended.

rule-knowledge performance classification. Reason (1990) sought to integrate two previously distinct areas of research - slips and lapses, in which actions deviated from intention due to execution or storage lapses; and mistakes, where the plan was inadequate to achieve the desired solution. He found some errors which did not conveniently fit either category, and described these as arising from the application of inappropriate diagnostic rules. He resolved the problem by differentiating between rule-based mistakes and knowledge-based mistakes. The main difference between Reason's (1990) generic error modelling system and Rasmussen's (1982) model lay in the attempt to present an integrated picture of error mechanisms operating at all three levels of performance.

Reason (1990) considered that errors at each of Rasmussen's three levels would be shaped not only by intrinsic effects such as attentional limitations, but also by extrinsic factors such as the structural characteristics of the task and context effects.

The concept of unsafe acts is only a part of what has become known as the Reason Model, however. Reason (1991) has argued that, in complex and tightly regulated systems such as airlines, the system is largely immune to a single failure by an individual, or a single mechanical failure. Rather, the aircraft is likely to fall prey to a systems failure. Unsafe acts by individuals are merely tokens of a generally unsafe condition. The unsafe acts trigger long-standing 'latent failures' (which are found generally within the province of management or organization) when the individuals are in the presence of unsafe environmental conditions (Fig 4).

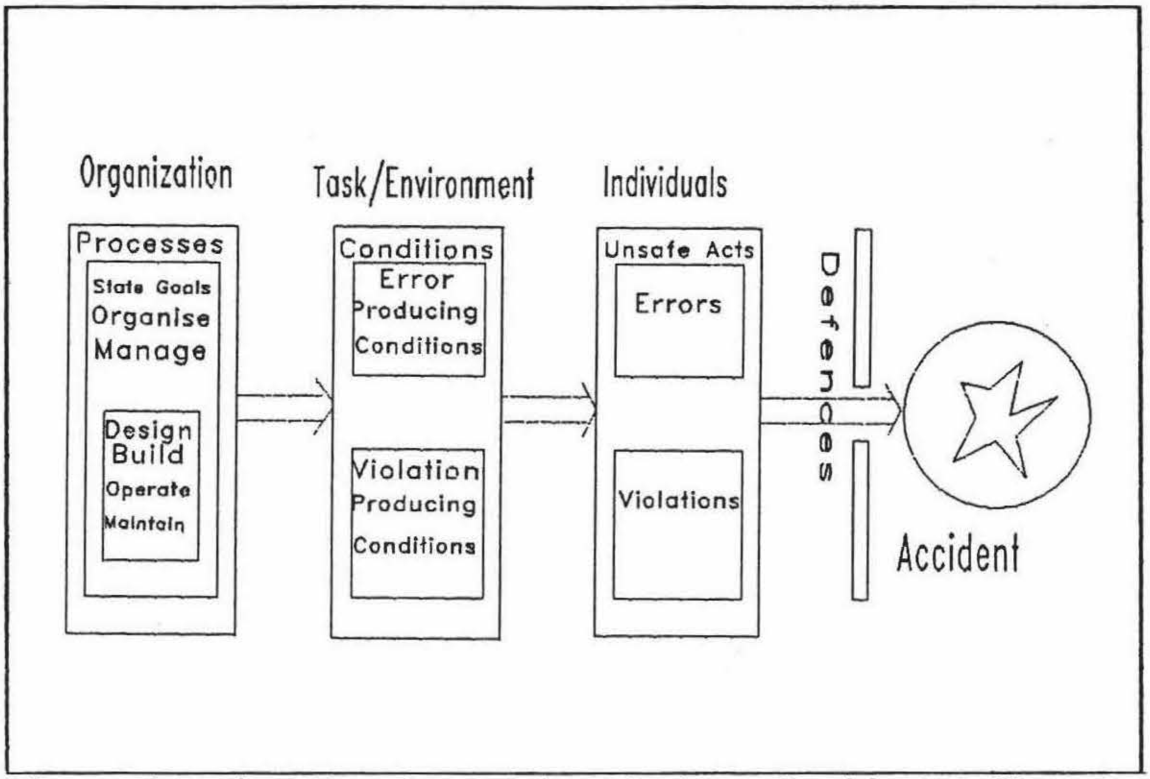

Figure 4 The Structure of a Systems Accident. (Source: Reason, 1991).

To illustrate these factors, the accident to the Air New Zealand DC10 in Antarctica (New Zealand Office of Air Accidents Investigation (OAAI), 1980) can be reviewed. The crew of the aircraft that flew into the side of Mount Erebus made errors in their methods of navigation (unsafe acts), in the presence of whiteout (an unsafe environment), but were the victims of deficient training and briefing (Vette, 1997) which were in turn the result of faulty communication within the organization (latent failures) (Mahon, 1981).

Helmreich's `crew-centred' approach to isolating the factors which led to the accident, has the potential to simplify the analysis advocated by Reason (1990; 1991) to identify the sources of latent failures (Zotov, 1996). While Reason (1991) considered that any technical organization is continuously involved in the related processes of designing, building, operating and maintaining, the aviation system is more complex than that. Interacting organizations such as engineering companies, manufacturing companies, air traffic control, and regulatory authorities may all contribute to an accident. Trying to track the sources of latent failures through such a network can be very complex. Preliminary examination of the accident using Helmreich's heuristic can indicate where the investigation should be focused.

The work of Reason (1990; 1991) and Helmreich (1990) has laid a foundation for the understanding of systemic factors in aircraft accidents, but in the view of O'Hare et al. (1994), the nature and causes of the 'active' failures (by pilots, air traffic controllers and engineers) remained relatively poorly understood. O'Hare et al. argued that there were three main sources of information on errors which led to aircraft accidents: anonymous reports such as the American Aviation Safety Reporting System (ASRS) (examined by Billings & Reynard, 1981); critical incident methodology such as the work of Fitts and Jones (1947) and Gerbert and Kemmler (1986); and the study of accident reports (O'Hare et al, 1994).

O'Hare et al. (1994) argued that both anonymous reports and critical incident reports suffered from the problems of self-reporting. Also, some types of errors may have been more liable to reporting than others, for example because they may have been more evident to the pilots, or because there may have been recent emphasis on particular errors in safety literature. The contribution of factors to accidents was therefore difficult to establish from these sources. Accident reports, which have been the subject of independent investigation, should not suffer these drawbacks.

Summaries of the salient features of groups of accidents, taken from official accident reports, had been the source of practical outputs such as design changes, and improvements in training, but these benefits were retrospective, not predictive. In order to forestall accidents, a theoretical basis for predicting the sorts of errors likely to occur was needed, and this is what O'Hare et al. (1994) sought to devise.

O'Hare et al. (1994) undertook two studies of human errors found in New Zealand aircraft accidents. In the first they sought to group the errors in accordance with the three step model advocated by Nagel (1988), and they also sought to replicate the findings of Gerbert and Kemmler (1986), of a four factor solution to the error variables. The second study used a more detailed analysis, based on the model outlined by both Rasmussen (1982) and Rouse and Rouse (1983) (see p. 14).

Two hundred and eighty four accidents to powered fixed wing aircraft (almost all of which were general aviation) were coded by O'Hare et al. in accordance with Nagel's (1988) classification of errors into information, decision and action categories. However, only 71% were able to be coded in this way, the remainder being chiefly technical failures with a few miscellaneous factors such as medical incapacitation. There was a clear dichotomy between those accidents where there was fatal or serious injury, and those where there was minor or no injury. Fatal/serious injury accidents were predominantly due to decision errors, that is, selection of a goal which was unwise in the circumstances, or an unsound strategy to achieve a selected goal; while minor accidents were largely due to poor choice or implementation of procedures necessary for the handling of the aircraft.

O'Hare et al. (1994) also coded the accidents for the 61 error variables obtained by Gerbert and Kemmler(1986). These were task-related variables, such as 'misjudgment of airspeed', 'misjudgment of clearance from obstacles' and 'failure to maintain airspeed'. Principal components analyses by O'Hare et al. extracted seven significant components, compared with the four found by Gerbert and Kemmler (1986), and O'Hare et al. considered that cockpit errors could be classified into at least five distinct categories, these being perceptual, decisional, procedural, monitoring, and handling.

The existence of five categories of cockpit error suggested that Nagel's (1988) three-element taxonomy of information, decision and action errors was an oversimplification. O'Hare et al.'s (1994) 'perceptual' and 'handling' categories corresponded to Nagel's 'information' and 'action' categories, but O'Hare et al. considered it necessary to expand Nagel's 'decisional' factor into 'decisional', 'procedural' and 'monitoring', because these factors played a key role in fatal accidents. O'Hare et al. (1994) therefore examined the applicability of the more detailed taxonomy, derived from the analysis of errors by nuclear power-plant operators by Rasmussen (1982), and Rouse & Rouse (1983).

In their second study, O'Hare et al. (1994) used the same database, expanded by the inclusion of later accident reports to a total of 323 accidents. The error categories they adopted were information, diagnosis, goal selection, strategy, procedure and action. (Rasmussen's (1982) 'target state' and 'task' were combined into 'strategy'). These categories were regarded as steps in an algorithm, each accident being examined for each of the error categories in turn. For example, if the pilot had detected information relevant to the aircraft state, the circumstances were then examined to see whether the aircraft state had been correctly diagnosed, and then whether the goal selected was appropriate to that state, and so on. The first and last stages in the algorithm adopted by O'Hare et al., 'information' and 'action', were the same as Nagel's (1988); the intervening four steps (diagnosis, goal selection, procedure and action) elaborated on the 'Decision' process.

O'Hare et al. (1994) coded the first failure to occur. In this second study, 70% of accidents were coded as cockpit errors, and the dichotomy between major accidents (33% goal errors) and minor accidents (28% procedural errors) was again found (Fig 5).

O'Hare et al. (1994) concluded that the six stage model derived from Rasmussen (1982) "provides a good account of human errors involved in a wide variety of aircraft accidents", and they advocated that investigators should target this six-step model "as key steps in their inquiry into the reasons behind the observed failures in the cockpit" (O'Hare et al., 1994, p. 1870).

The dichotomy found by O'Hare et al. (1994) in the error distributions between major and minor accidents might have been due to inherent characteristics of such accidents, but an alternative explanation could be that it was due to differences in the ways such accidents were investigated and reported. In their investigation, O'Hare et al. (1994)

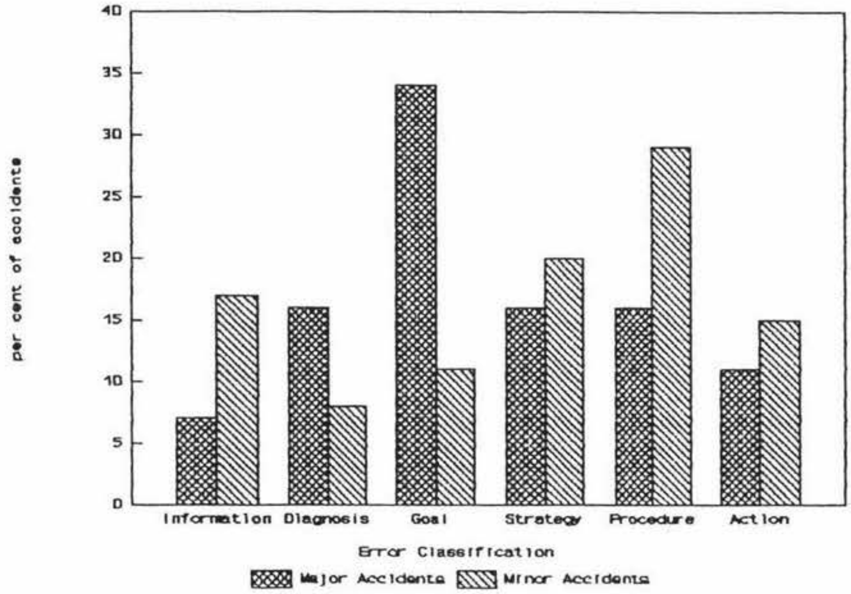

Figure 5 Comparison of errors - major and minor accidents (Source: O'Hare et al, 1994)

appear to have assumed that fatal accidents and non-fatal accidents were investigated in New Zealand with equal thoroughness. This was not so in the case of some of the reports they studied. Fatal accidents, reported in ICAO format, were all thoroughly investigated, but non-fatal accidents (which constituted the bulk of the database) were generally reported in 'Brief' format. Many of these non-fatal accidents were not subject to formal investigation, being just transcriptions of the pilots' accident notifications. New Zealand Office of Air Accidents Investigation Brief Reports serial numbers 84-010, 85-026, 85-030, 85-034, 86-075, and 96-013, are examples of such reports included in their database by O'Hare et al (1994), which examination of the official files shows to have been just transcriptions of the pilots' reports.

There are many more non-fatal accidents than fatal, so a substantial part of the study related to self-reports - the very thing which the researchers had hoped to avoid by studying accident reports.

One might expect pilots' reports to contain an element of self-presentation, that is, "concern for the impression one makes on others" (Dane, 1990, p. 11). Pilots may persuade themselves that events followed a sequence which shows them in a more favourable light than was really the case, or they may seek to misrepresent events.

In addition, pilots are subject to the usual limitations of witnesses generally. There is a desire to retain a cohesive and complete picture of events, so where there may be gaps in the witness's perception, there will be a tendency to create links in memory to fill those gaps. Also, where the event has been discussed with others, the witness can unwittingly absorb information from others and build it into 'recollections' of events (Rolfe & Bekerian, 1985).

Rolfe and Bekerian (1985) also took the view that pilot witnesses (whether involved in the accident or not) are likely to be unable to distinguish between what they saw, and their interpretation of those events. And when the pilot was involved in the accident, the effect of stress will have had the effect of concentrating the attention on the focus of the threat. Thus, if the aircraft was on fire, the pilot may be able to describe the flames, but not the flightpath the aircraft followed while the aircraft was on fire.

The evidence from pilots should therefore be treated with reserve, and investigators should seek corroboration from other witnesses or from physical evidence. Certainly, when some minor accidents have been investigated in detail the results have been at variance with the pilots' reports (OAAI, 1989a; 1989b; 1989c).

An additional source of uncertainty in the study by O'Hare et al. (1994) was the inclusion of accidents involving 'serious injury' with fatal accidents. 'Serious injury' is a standard International Civil Aviation Organization (ICAO) term; it includes quite minor injuries. For example, a person is said to be 'seriously injured' if suffering any fracture other than of fingers, toes or nose (ICAO, 1994). Since some of the 'serious injuries' are little different from 'minor injuries', it follows that the forces generated in the accidents that resulted in those categories of accidents may be not dissimilar, and in that case there is little reason to suppose that the events leading up to accidents in those categories are dissimilar. However, the inclusion of cases of 'serious injury' which were not in fact the result of major accidents would tend to blur any distinctions between 'serious' and 'minor' cases.

Also, some of the 'serious injury' accidents used in the study by O'Hare et al. came from self-reported Briefs, whose reliability is in question. It may be that a more useful cut-off would be the distinction between 'fatal' and 'non-fatal' accidents: these definitions would be unambiguous, and the fatal accidents were invariably the subject of formal investigation.

A possible source of bias was introduced into the study by O'Hare et al. (1994), by the inclusion of those groups of accidents where subsequent investigation has shown that causal factors were incorrectly determined. The earlier reports have seldom been amended retrospectively. Examples of such erroneous reports are the hang-glider 'downwind turn' accidents (OAAI, 1987), and the Pterodactyl microlight aircraft accidents (New Zealand Transport Accident Investigation Commission (TAIC), 1993). In both cases, the causes of the accidents were originally attributed to pilot factors, but were subsequently found to be aerodynamic, and beyond the control of the pilots.

A proportion of minor aircraft accidents are essentially trivial: they result from the normal process of training (learning by error - Reason (1990)), and apart from the expense they cause can hardly be considered accidents, being entirely predictable. There is little to be gained by investigating repeated errors by student pilots (Dr. R. B. Lee, personal communication). The inclusion of this pool of data would tend to skew the 'minor accidents' distribution in favour of skill errors, and might explain part of the dichotomy found by O'Hare et al.

The erroneous assumption, that all accident reports resulted from rigorous investigation, does not invalidate the error taxonomy derived by O'Hare et al. (1994). The assumption could only affect the proportions of the different types of error. However, the assertion that the errors in serious and minor accidents are different in character should be re-examined.

Effect of pilot experience on distribution of error types

Gerbert and Kemmler (1986) considered that one of the differences between skilled and trainee pilots might be that skilled pilots may have automated many processes. The process of receiving information, decision-making and implementation may be short-circuited at a lower cortical level, sensory perceptions being immediately transformed into appropriate actions. (Physiologically, the brain may be divided into *higher centres* which are the seat of consciousness, memory and will, and the *lower centres* which control many unconscious acts. Sensory inputs are received in the lower cortex, and movement is controlled from the motor cortex). Thus trainees would have more opportunity to make decisional errors, the decisions not yet having been automated. Also, they may experience higher workloads because all their actions must be fully processed, and usually several of these processes take place simultaneously.

Wickens and Flach (1988) also thought that decision quality could be expected to be affected by experience. They considered the most apparent difference between expert and novice lay in the long-term memory store of information; the expert should be able to interpret patterns of environmental cues with less effort so the information phase should take less time, and produce more accurate results. Also, the expert had a greater store of hypotheses and actions that could be generated in the search for a solution, so the goal and strategy phases should produce better solutions. However, Wickens and Flach (1988) found that experimental evidence did not strongly support the conclusion that expertise affected the overall quality of decision-making. One might perhaps look for similar error rates overall, as reflected in the accident rates, but the types of error committed at various levels of experience could vary.

Such a variation in the types of errors made by pilots with different levels of experience was found by O'Hare et al. (1994). Procedural errors were more common when pilots had less than 100 hours experience, action errors were more frequent than predicted when pilots had less than 500 hours, and goal errors were more frequent than predicted when they had between 100 and 1000 hours experience.

The relationship between major accidents, minor accidents and incidents

In considering human error, Rasmussen (1982) considered that humans learn, in part, by experimenting on a trial-and-error basis. Accidents might be considered to be unsuccessful experiments, which had undesirable consequences. Undesirable consequences could occur in an environment where the effects of inappropriate actions could not be seen and reversed before they led to an accident. Such an environment would exist where the effect of the experimental action was delayed in time, or dependent on further steps, or dependent on latent conditions.

A corollary to this view is that incidents are the results of errors where the effect is immediately observable and reversible by the operator, and therefore corrected without further notice. This provides some theoretical basis to the observations of Jensen and Benel (1977), O'Hare et al. (1994) and Wiegmann and Shappell (1996), that the errors involved in major and minor accidents are, at least in part, different in kind. Also, different sorts of errors may not be proportionately reported in anonymous reporting systems such as the American Aviation Safety Reporting System (ASRS), because some error types may be readily correctable by the operators, and so may not be perceived by them as being of any significance.

Nagel (1988) pointed out that a drawback to ASRS and similar systems is that reports are not made on a purely random basis. "Certain individuals (perhaps those more safety conscious) may report more often than others. Certain operational conditions (new regulations or safety assurance programs, for example) may induce people to report more frequently in some geographical areas, or during certain periods than in others ...We may learn a great deal about what errors are occurring, but not necessarily be able to determine much about how often errors occur" (Nagel, 1988, p. 270).

It is also possible that ASRS errors may over-represent those which can be detected and corrected by the pilot, and so do not represent a hazard.

Comparison of taxonomies

In a recent study, Wiegmann and Shappell (1996) examined the applicability of a variety of error taxonomies to the existing United States Navy (USN) aircraft accident database. In their study they quote

Senders and Moray (1991): "There are as many taxonomic schemes as there are people interested in the topic". This has happened because researchers in a particular area have devised schemes which suit their individual needs. The investigation of human factors in aircraft accidents is usually performed by people from outside the field of aviation psychology, or human factors psychology. The resulting databases are generally task-oriented, and are not conducive to traditional human factors analyses.

Wiegmann and Shappell (1996) took as an example the existing USN aircraft accident database, which is arranged in an arbitrary 'who', 'what' and 'why' format, with a list of 289 possible causal factors. This framework appears to answer many human factors questions, but it has no theoretical basis. The lack of a theoretical basis makes it difficult to infer specific causes of human error, a necessary step in developing interventions to reduce the occurrence or consequences of pilot errors.

Accordingly, Wiegmann and Shappell (1996) set out to see whether the existing USN aircraft accident database could be organized and analyzed by conceptual human factors frameworks. They coded the database in accordance with three models: that of Wickens and Flach (1988), a traditional four-part 'information processing' model; the taxonomic algorithm devised by O'Hare et al. (1994), which in turn was derived from Rasmussen's (1982) work and which seeks to diagnose the underlying cognitive failure responsible for an error; and Reason's (1990) model of unsafe acts.

The above models all examine direct causes of accidents, which may be broadly described as errors, in contrast with the environmental conditions (contextual), or such matters as supervisory deficiencies, known as latent failures by Reason (1990, 1991).

Each of the above models had its limitations when attempting to represent the information contained in the USN aircraft accident database. Wickens and Flach's (1988) information processing model could not encompass such matters as preflight planning, psycho-social variables of crew coordination, and physiological conditions like fatigue and disorientation. These factors accounted for about 20% of the database.

Rasmussen's (1982) model of the internal human malfunction, as modified by O'Hare et al. (1994), could not deal with psycho-social variables or physiological conditions, but did allow for goal or strategy errors. The factors unaccounted for were 12% of the database.

Reason's (1990) model of unsafe acts could not handle psycho-social factors or physiological conditions, nor sensory/information errors. However, this model does accommodate flight planning errors, and overall, the factors unaccounted for were about 16% of the database.

From these results, it is evident that the model of unsafe acts, and the model of internal human malfunction, account for a greater percentage of the factors recorded in the USN aircraft accident database than does the traditional information processing model.

Another interesting finding from this study was that "In general, it appears that major and minor accidents are due, at least in part, to qualitatively different problems. This finding tends to dispel the old adage that ... the difference is one of luck" (Wiegmann & Shappell, 1996, p. 14). This tends to support the findings of O'Hare et al. (1994), who found that there appeared to be different distributions of error types between major and minor accidents; rather than those of Billings and Reynard (1981) who considered that accidents and incidents came from the same population, the difference in outcomes being a matter of chance.

CHAPTER 3
METHODS

Sample

The accident database comprised the following official Reports of the New Zealand Office of Air Accidents Investigation (OAAI):

(i). All published reports of fatal accidents for the period 1965 - 1989 inclusive, 1989 being the last year for which consolidated reports were available.

(ii). All published Brief Reports for the period during which they were published, 1970 - 1990 inclusive, with the exception of those for 1975 and 1976 which were not available.

(iii). All incidents which had been examined, and on which reports were published. There were few of these, and they were included with the non-fatal accidents.

All files relating to non-fatal accidents (some 2500) were examined to determine which accidents had been subject to professional investigation; those which had not been were excluded from the sample.[4]

---

[4] While it would have been desirable to include only those non-fatal accidents which had been subject to a full investigation, there were insufficient of these. The criterion used was that there had been at least some field investigation, so that the Inspectors had the opportunity to form their own views about the facts of the accident.

Accidents known or suspected to have been incorrectly reported (as a result of subsequent investigation) were excluded from the sample.

The final sample comprised two hundred and seventy four reports of fatal accidents, and five hundred and forty nine reports of non-fatal accidents.

Procedure

Accident reports accepted for the sample were coded in accordance with the coding instructions developed by O'Hare et al. (1994) (Appendix A). The taxonomic algorithm is shown in Fig 6.

The algorithm progresses from the gathering of information, to making decisions based on the information, to implementing those decisions.

Starting with 'Information', this may be available, but the pilot may not seek or gather it. For example, in an accident to a B737-400 at Kegworth involving a high level of engine vibration, the unfamiliar 'glass cockpit' engine vibration display was correctly showing which engine was at fault, but the pilots did not refer to it (AAIB, 1990). Such a lapse would be coded as an information error.

Having gathered information, the pilot's picture of the state of the system should closely approximate the real world, but it does not always do so. A classical example was the B737 fire at Manchester, where the pilot's RT call - "It looks as though we have a fire in number one engine" - has to have been the understatement of the year. (AAIB, 1988). Misinterpretation of information in this way would be

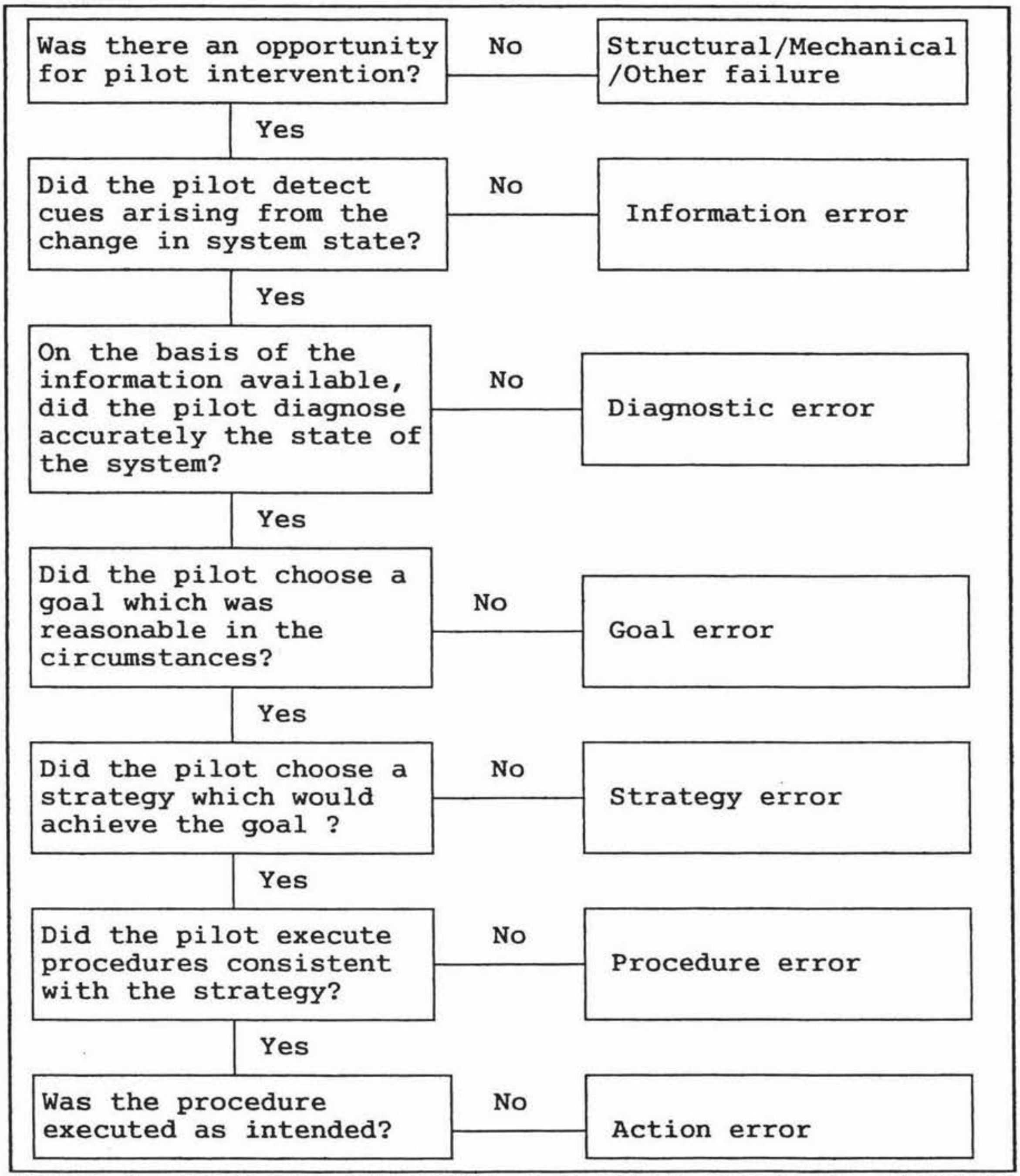

Figure 6 Taxonomic algorithm for pilot errors in aircraft accidents. (Source: O'Hare et al, 1994).

coded as a diagnosis error.

Next, the pilot may be faced with a choice between a number of possible goals. Goals are identified according to their information value. 'To arrive safely' would not be identified as a goal, since its informational value is low. However, 'to divert to an alternate' is informative by comparison with other goals which might be adopted such as 'to hold until the weather improves'. While the pilot will generally have some goal in mind, goal selection may not always be optimal; a sub-optimal goal selection would be coded as a goal error.

The means of achieving goals are referred to as strategies, that is, a general plan or approach. Thus, the goal of diverting could be achieved by climbing only to minimum en-route altitude (perhaps minimising traffic delays) or by climbing until the descent profile is intercepted (often giving the best fuel economy). The appropriateness of a strategy for achieving the pilot's goal can be judged; a sub-optimal strategy would be coded as a strategy error.

A procedure is a specification for conducting a set of components of a higher-level task. In the case of diversion from an instrument approach, the first procedure to be executed is a go-around, which includes the following:

Raise the nose
Apply go-around power
Raise the undercarriage
Accelerate to flap-raising speed
Raise the flaps

If the procedure performed is inappropriate to the chosen strategy, or an appropriate procedure is not performed at all, this would be coded as a procedure error.

Any action can, of course, be performed incorrectly, such as by raising the undercarriage instead of the flaps after landing. If the procedure is not executed as intended, then this would be coded as an action error.

Coded Event

The algorithm devised by O'Hare et al. (1994) deals with "the internal mental function which was not performed as required by the external task" (Rasmussen, 1982, p.319). Although accidents are generally due to a sequence of events, O'Hare et al. (1994) considered it necessary to restrict each analysis to one event, otherwise there would be problems in interpreting statistical data due to interdependencies. They argued that it was logical to use the earliest event identifiable, which set the causal chain in motion, i.e. the first identifiable point in the flight at which there was a significant departure from accepted and prudent[5] practice.

In the present study, O'Hare et al.'s (1994) method of selecting what might be termed the 'triggering' event was followed.

For example, in the accident to ZK-GSG (OAAI, 1986b), a motor glider suffered an engine failure after take-off, while at about 400 feet in the downwind leg of the circuit. The engine failure was caused by fuel exhaustion. The fuel gauge was faulty but the pilot was aware that the fuel state indicated

---

[5] Originally, the phrase 'accepted or normal practice' was used in the definition of the coding point, but this gave rise to considerable uncertainty. For example, it could be argued that in New Zealand it was 'normal' to fly single-engine aircraft over terrain on which a safe forced landing could not be made in the event of engine failure, though fatal accidents resulted from this practice. Likewise, it was almost 'normal' to ignore deficiencies in equipment required by Regulations. However, neither practice was prudent.

was not consistent with the likely consumption earlier in the day. Information on the fuel state was available to him (though not too easily) by dipping the tank, but he did not do so. The triggering error leading to the accident would therefore be coded as an information error; had the pilot not made this error, the engine would not have failed, so he would not have been in the position to make the subsequent errors which led to the accident.

However, prima facie an engine failure in a motor glider on the downwind leg of the circuit at an aerodrome ought not to lead to an accident. The factors which led to the subsequent stall on final approach were the decisions to leave the motor extended, so increasing the drag coefficient, and to fly at an excessively high airspeed, causing a severe degradation of the glide angle. These were procedural errors, and were the significant errors which led to this accident. Caution was therefore necessary for the coder to ensure that the earlier triggering error was coded.

Coding Reliability

Coding was done by the author of the present study. The author has a background of 25 years in military and civil aviation, including ten years in operational research. This was followed by seven years as an Inspector of Air Accidents with the New Zealand Office of Air Accidents Investigation, during which he investigated more than thirty fatal and numerous non-fatal accidents.

With any coding process it would be desirable to have multiple coders. The large size of the database made it impossible to find suitable coders who had sufficient time to take part. All samples were therefore coded by the author, and precautions were taken to ensure both consistency and close agreement with other studies.

During the pilot study (Zotov, 1995b), the first codings were done using only the coding algorithm developed by O'Hare et al. (1994). These codings produced distributions that bore little resemblance to those found by O'Hare et al. The detailed instructions at Appendix A were then provided by Dr. O'Hare, and after practice and critiquing by colleagues, consistent results similar in form to those obtained by O'Hare et al. were achieved.

The coding process is open to interpretation. It was therefore necessary to examine whether the coding instructions were robust, that is, whether different coders would achieve similar results. A non-probability convenience sample of the codings generated by O'Hare et al. (1994), was compared with the present author's coding for the same accidents. Reasonable agreement was found (kappa = 0.57)[6], similar to that reported by O'Hare et al. (1994). The differences were examined, and appeared to be caused by differences in perception, e.g. as to whether a goal was 'reasonable'. This may have been because O'Hare et al have a background in psychology with some flying experience, and may at times have taken a different view from the present author, who has extensive flying experience, but a more limited background in human factors.

---

[6] Fleiss (1981) described values of kappa over 0.75 as 'excellent' and values between 0.6 and 0.75 as 'good'.

After the entire database had been coded by the present author, the coding for the fatal accidents for one year was repeated (n = 22). Only one discrepancy was found, indicating that consistency in coding had been achieved.

The close resemblance between the distributions for major accidents to fixed-wing powered aircraft found by O'Hare et al. (1994) and that for fatal accidents found in the present study (see p. 59) gives confidence that differences between coders were not a dominant factor.

While there are still likely to be differences between individual coders, where the entire database is coded by one coder any bias is likely to affect all groupings similarly. Differences between groups are thus unlikely to arise because of individual bias.

Analytical Methods

The class of aircraft involved (powered fixed-wing, rotary wing or glider) may influence the outcome of particular types of errors in a number of ways. For example, helicopters are more unstable than fixed-wing aircraft (Saunders, 1975) so that more of a pilot's short-term memory capacity may be required for flying, leaving less available for decision-making.

Additionally, the environment in which a class of aircraft operates may exacerbate the effects of some types of errors. For example, agricultural fixed-wing aircraft operating at low level in mountainous terrain may be jeopardised by strategy or procedural errors which would be insignificant in other fixed-wing powered aircraft operations.

Because of the potential for variability between classes of aircraft and operations, classes were analyzed separately, as follows:

(i). Powered fixed-wing aircraft, other than agricultural or light sporting aircraft.

(ii). Fixed-wing agricultural aircraft.

(iii). Helicopters, not engaged in venison recovery.

(iv). Helicopters engaged in venison recovery.

(v). Gliders, other than hang-gliders.

(vi). Light sporting aircraft: microlights, autogyros and hang-gliders.

O'Hare et al. (1994) found that their samples showed variability with pilot experience. Experience groupings could be considered to be:

(i). 0-50 hours, approximating to a Student Pilot Licence (SPL)
(ii). 51-200 hours, approximating to a Private Pilot Licence (PPL)
(iii). 201-2000 hours, approximating to a Commercial Pilot Licence (CPL)
(iv). 2001+ hours, approximating to an Air Transport Pilot Licence (ATPL), or an experienced CPL.

These groupings were examined to see whether they showed different distributions of error types, in the aircraft categories which had sufficient samples - powered fixed-wing aircraft, agricultural and non-agricultural.

Statistical Analysis

The limited number of error classes precluded the use of multivariate analysis. Since the primary objective was to compare the distribution of error classes between fatal and non-fatal accidents, chi-squared tests were used to examine the effects from class of aircraft and level of pilot experience: where the effects were statistically non-significant, data could be pooled without biasing the results.

Comparison between aircraft types

A chi-squared test was used to establish whether samples came from populations with the same distribution. A 6 x 2 table was used to compare:

(i). The distributions of error types in fatal accidents, between various classes of aircraft, to see whether pooling was feasible

(ii). the distributions of error types in non-fatal accidents, between various classes of aircraft, to see whether pooling was feasible

Comparison of fatal and non-fatal accidents

A 6 x 2 table was used to compare the distributions of error types between fatal and non-fatal accidents, within various classes of aircraft, to see whether the differences found by O'Hare et al. (1994) between major and minor accidents were replicated when examining accidents that had been subject to investigation.

Effect of pilot experience

A 6 x 4 table was used to compare:

(i). The distributions of error types in fatal accidents, within various classes of aircraft, between various levels of pilot experience.

(ii). The distribution of error types in non-fatal accidents, within various classes of aircraft, between various levels of pilot experience.

Comparison of results from the present study, and from O'Hare et al. (1994)

A chi squared test for goodness of fit was used to establish whether the powered fixed-wing samples came from populations with the same distribution. A 6 x 2 table was used to compare:

(i). O'Hare et al.'s (1994) results for 'Fatal/ Serious Injury' with 'Fatal' results from the present study, to establish whether the removal of 'serious injury' accidents had any effect.

(ii). O'Hare et al.'s (1994) results for 'Minor/Nil Injury' accidents with 'Non-fatal' results from the present study, to see the effect of using only those reports where the accidents had been subject to investigation.[7]

---

[7] The non-fatal results were modified by exclusion of 'serious injury' cases, and agricultural aircraft results were included in both fatal and non-fatal categories, for compatibility with the results of O'Hare et al. (1994).

CHAPTER 4
RESULTS

After testing to ensure coder reliability as detailed at p. 43 (Kappa = 0.57), the accident database consisting of 814 cases was coded in accordance with the instructions used by O'Hare et al. (1994) (Appendix A). The distributions of error types were compared between aircraft classes, and the results were pooled where possible. The distributions, within aircraft types, were then compared between fatal and non-fatal accidents. Where there were sufficient samples within aircraft types (the fixed-wing powered aircraft groups) the distributions were examined for a pilot experience effect.

The results for fixed-wing powered aircraft were then compared with those found by O'Hare et al. (1994).

Accidents attributable to pilot error

The percentages of accidents attributable to pilot error, by aircraft class and accident severity, is shown in Table 2.

Table II Percentages of accidents arising from pilot error

Table 2

Aircraft Class	Severity			
	Fatal		Non-fatal	
	Total accidents	Pilot Error (%)	Total accidents	Pilot Error (%)
Fixed wing Power *	123	80	163	68
Fixed wing Agricultural	58	74	65	43
Helicopter **	29	65	138	32
Helicopter (Venison)	29	52	47	30
Glider	19	79	20	55
Light Sporting	16	50	9	78

NB: * = Excluding Agricultural Aircraft
** = Excluding Venison
Venison = engaged in deer recovery

Differences between aircraft classes

The distributions of error types for each class of aircraft were compared, both for fatal and for non-fatal accidents.

(i). The differences between agricultural fixed-wing powered aircraft and other fixed wing powered aircraft were significant. (Figs 7 and 8).

Fatal accidents: chi-squared = 12.38; df 5; $p < 0.03$.
Non-fatal accidents: chi-squared = 16.09; df 5; $p < 0.007$.

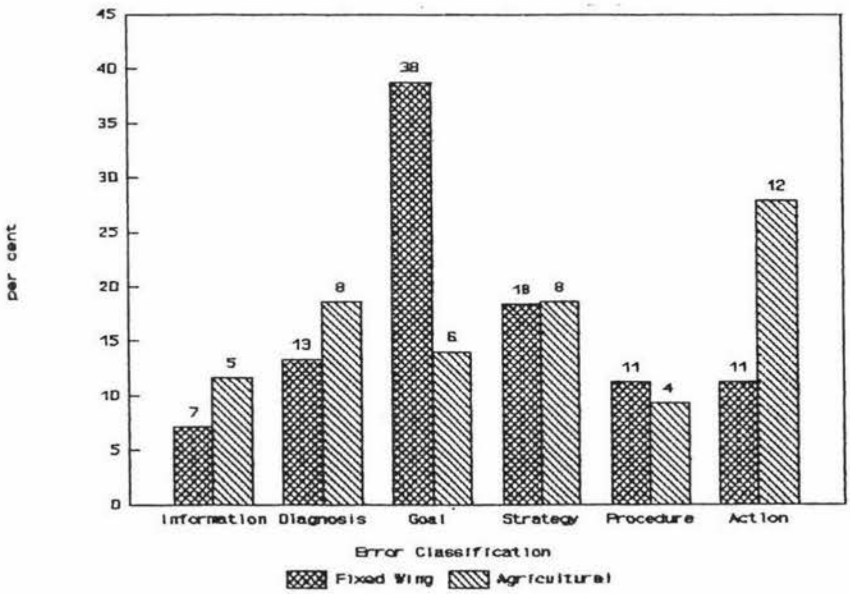

Figure 7 Fatal Accidents: Agricultural c.f. Other Fixed-wing Aircraft

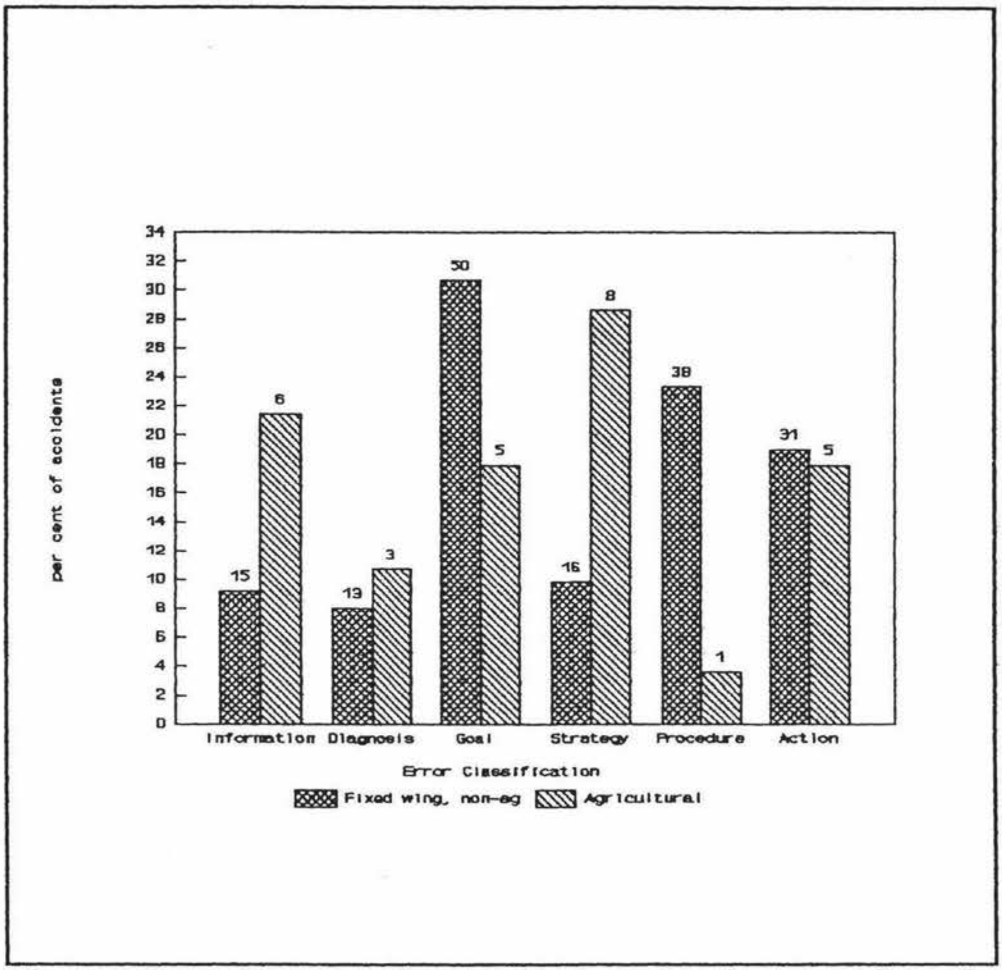

Figure 8 Non-Fatal Accidents: Agricultural c.f. Other Fixed-wing Aircraft

(ii). The differences between helicopters involved in venison recovery, and other helicopters, were not significant either for fatal or non-fatal accidents. Helicopter results were therefore pooled for subsequent analysis.

(iii). The pooled helicopter results were compared with those for non-agricultural fixed-wing aircraft, since the distributions were somewhat similar in form. For fatal accidents the differences were significant (chi-squared = 28.06; df 3; p < 0.005) (Fig 9). (The following error classes were pooled to give sufficient numbers in cells: 'information' with 'diagnosis' and 'procedure with 'action').

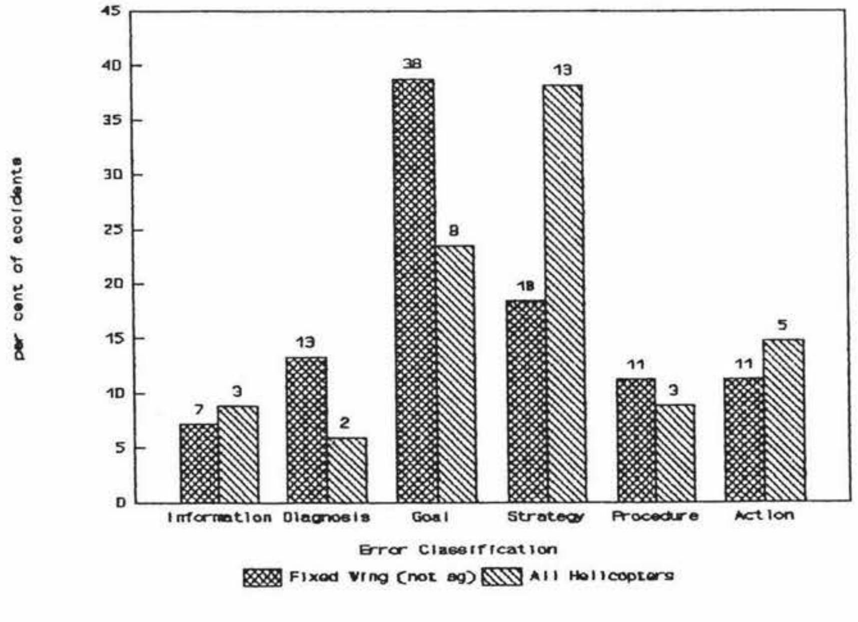

Figure 9 Fatal Accidents: Fixed Wing c.f. Helicopters

For non-fatal accidents the differences between helicopters and fixed-wing non-agricultural aircraft were approaching significance: chi-squared = 10.95; df 5; $p < 0.054$ (Fig 10).

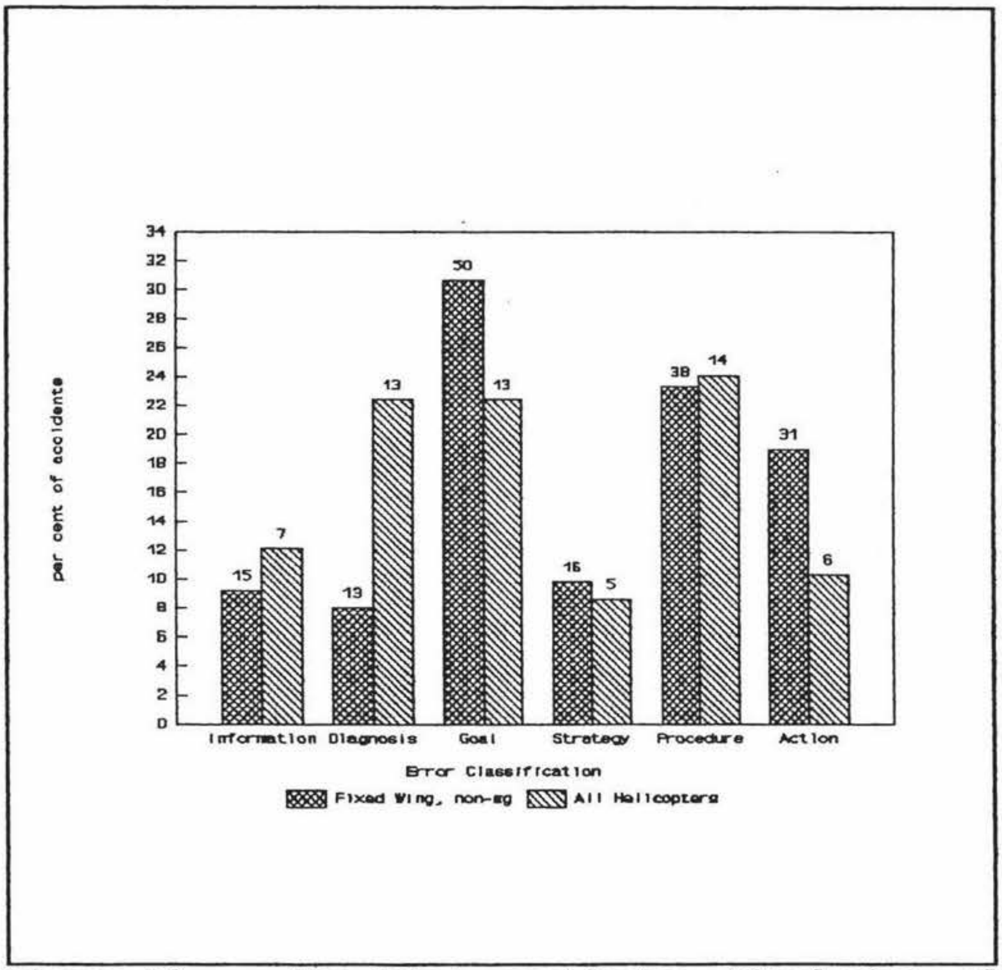

Figure 10 Non-fatal accidents: Fixed-wing cf Helicopters

There were insufficient results for gliders, and for light sporting aircraft, to make valid comparisons.

Differences between fatal and non-fatal accidents

The differences between fatal and non-fatal accidents within each aircraft class were examined.

(i). For non-agricultural fixed-wing powered aircraft, the differences were significant (chi-squared = 13.73; df 5; $p < 0.02$). (Fig 11).

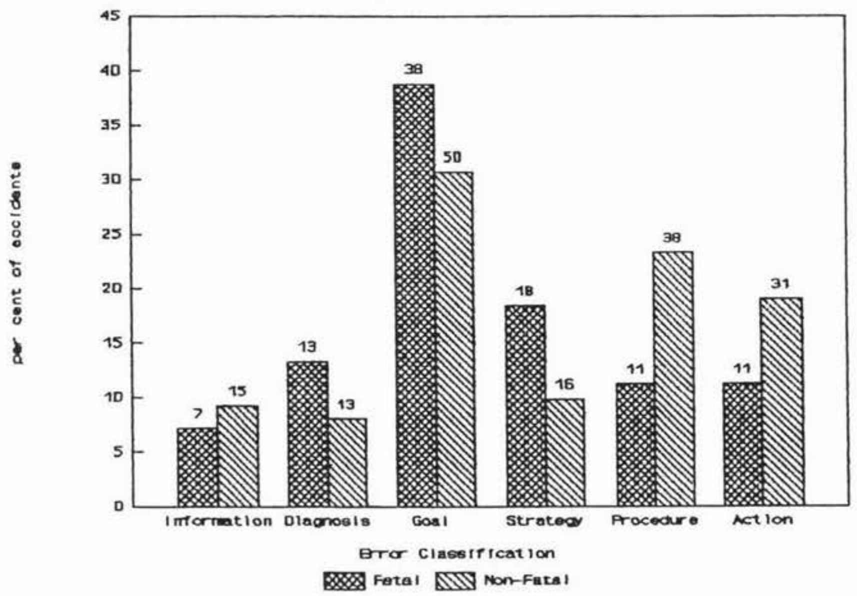

Figure 11 Fatal c.f. Non-Fatal Accidents: Fixed Wing Aircraft (not Agricultural)

(ii). For fixed-wing agricultural aircraft, the differences were not significant.[8] [However, the difference in the shape of the distributions for fixed-wing agricultural aircraft (Fig 12), compared with those for other fixed-wing aircraft and helicopters, may be important. This point is discussed later, pp. 63, 73.]

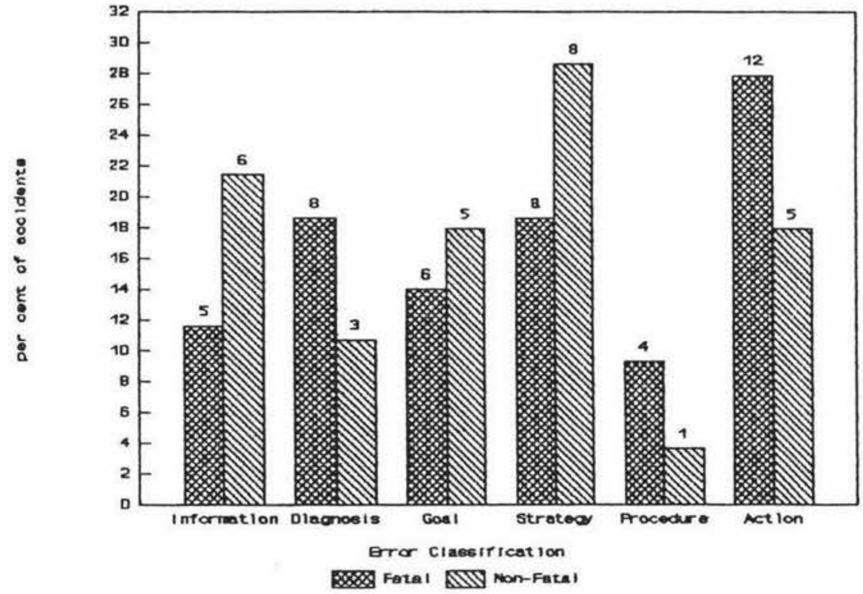

Figure 12 Fatal cf Non-Fatal Accidents: Fixed-wing Agricultural Aircraft

---

[8] Although from Fig 12 the distributions appear somewhat different, 5 of 12 cells had an expected frequency of less than 5.

(iii). For helicopters, the differences were significant (chi-squared = 16.48; df 5; p < 0.006) (Fig 13).

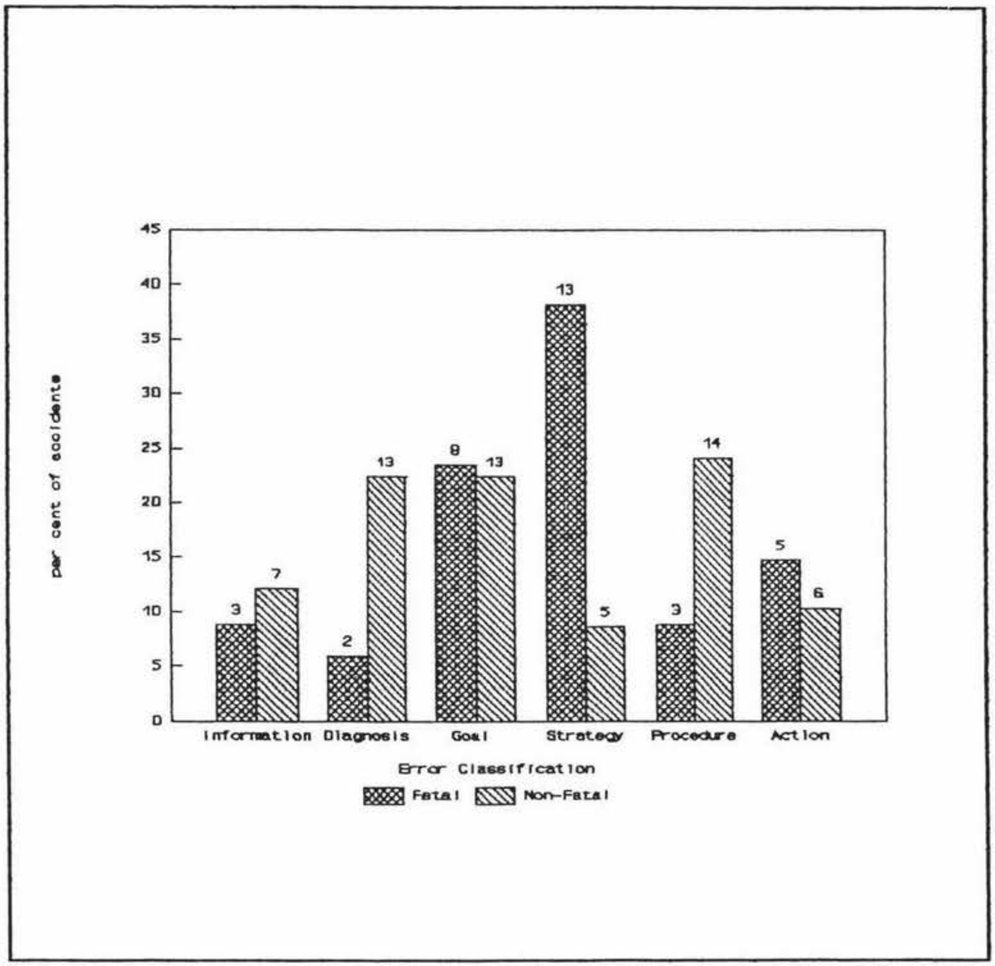

Figure 13 Fatal c.f. Non-Fatal Accidents: Helicopters

For gliders and for light sporting aircraft, the sample sizes were insufficient to draw valid conclusions.

The effect of pilot experience

The distributions of error types by pilot experience were examined within the fixed-wing powered aircraft classes. (There were insufficient samples to examine the effect of pilot experience within the other aircraft classes). In both agricultural and non-agricultural classes there were no significant results.

Comparison with the results obtained by O'Hare et al. (1994)

To produce comparable samples, some adjustments had to be made to the groupings from the present study. O'Hare et al. (1994) examined the distributions for all fixed-wing aircraft (that is, agricultural and non-agricultural combined), whereas in the present study these groups had been kept separate. Also, 'serious injury' accidents were included by O'Hare et al. within their 'major accidents' category, but in the present study these had been included in the 'non-fatal' category. The results from the present study were adjusted as follows:

(a). All fixed-wing powered aircraft results were pooled

(b). 'Serious injury' accidents were excluded from the 'non-fatal accidents' results.

(i). The 'Major accidents' distribution (O'Hare et al., 1994) and the 'Fatal accidents' distributions were closely similar in form, as will be seen from Fig 14, and the differences were not significant.

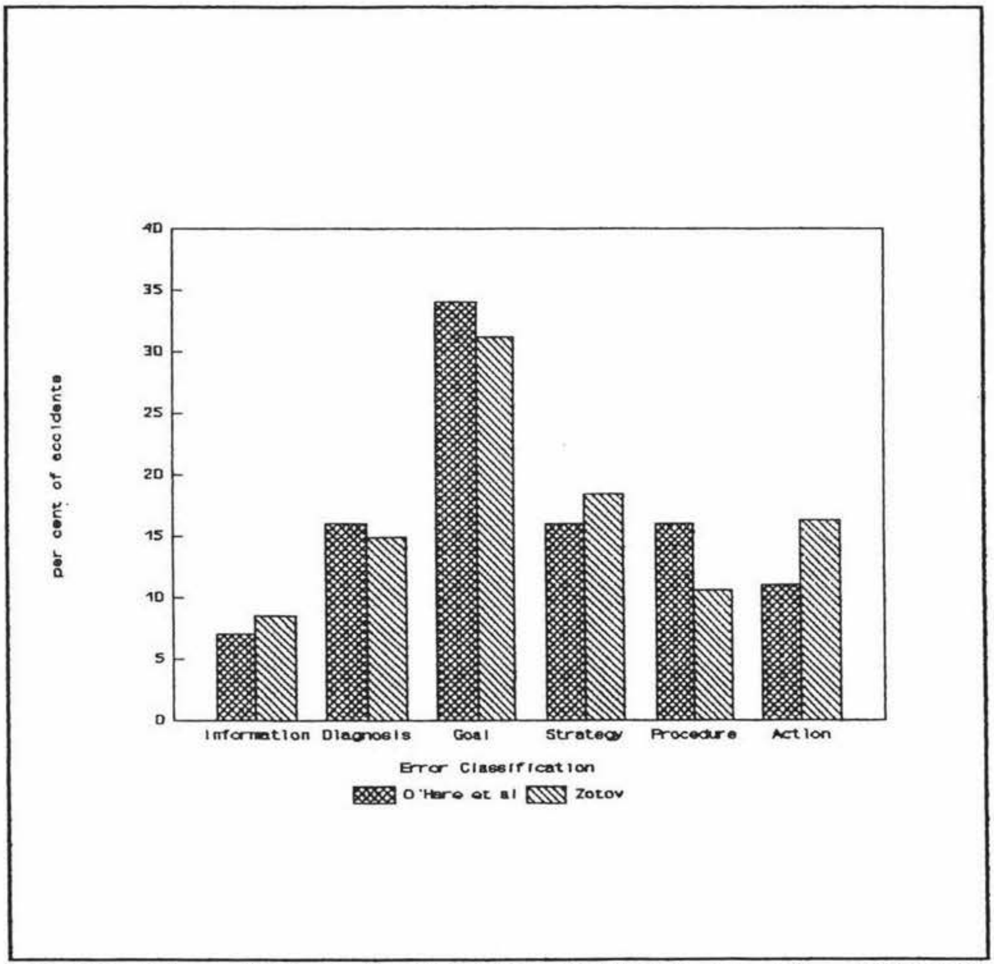

Figure 14 Major Accidents (O'Hare et al. (1994)) c.f. Fatal Accidents

(ii). The 'Minor accidents' distribution (O'Hare et al., 1994) and the 'Non-fatal accidents' distribution were different in form, and the difference was significant. (chi-squared = 71.14; df 5; $p < 0.005$) (Fig 15).

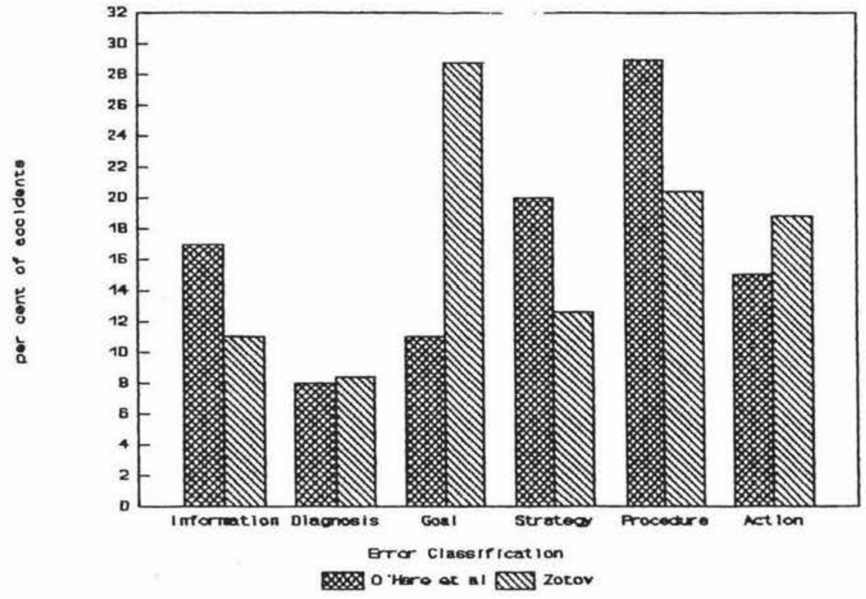

Figure 15 Minor Accidents (O'Hare et al. (1994)) c.f. Non-fatal Accidents

CONCLUSIONS

Accidents coded as pilot error

When considering fatal accidents, the general rule is that about three quarters of all accidents are attributable to pilot error. This rule is generally followed in this study, in the case of fixed-wing aircraft, powered or glider, and also for helicopters not engaged in venison recovery. There is, however, a clear trend towards an increasing proportion of accidents due to mechanical failure as the operating environment becomes more hazardous (Table 2).[10] In part this may be because in New Zealand, malfunctions which would result in a safe forced landing on open terrain, generally have catastrophic results if they occur over mountainous or bush-covered country. However, it also reflects the reliability of the aircraft which are engaged in these operations.

Agricultural aircraft, which have a somewhat higher proportion of fatal accidents due to mechanical failures than other fixed-wing aircraft, operate from rough airstrips, with a corrosive environment due to agricultural chemicals. Also, their engines repeatedly cycle between high power and idle, resulting in

---

[10] Although reliable figures are difficult to find (reporting of hours flown is not mandatory, in most cases) it is generally accepted that the accident rate for normal helicopter operations is significantly greater than that for fixed-wing aircraft. For helicopters engaged in venison recovery the rate is higher still. For 1982, when venison recovery operations were in full swing, the accident rates per 10 000 hours were:

Aeroclub (fixed-wing) 1.29; other fixed-wing (general aviation) 1.25; agricultural fixed-wing 1.96.
Agricultural helicopters 6.32; other helicopters including venison recovery 18.45. (OAAI, 1982).

thermal shock and increased wear. Helicopters, which suffer many mechanical failures, have to be lightly constructed because the engine must overcome the weight of the aircraft, rather than just the drag as in fixed-wing aircraft. For the same reason their engines operate at higher power settings than the engines of fixed-wing aircraft, and so are more prone to mechanical failure. Also, helicopters are more prone to vibration, and thus fatigue, than fixed-wing aircraft because of the large rotating mass of the rotor and the fluctuating forces thereon. There is thus a propensity for agricultural aircraft and helicopters to suffer mechanical failures, and when such failures occur, there may be little opportunity for the pilot to make a safe landing.

The apparent high level of mechanical failures in the 'non-fatal' accident category does not imply that the pilot error rates were lower in non-fatal than in fatal accidents. Mechanical failure was one of the criteria in deciding to allocate investigation resources to non-fatal accidents, since this was perceived to be an area where safety could be progressively enhanced by investigation. Overall, the pilot error rates for non-fatal and fatal accidents appeared to be similar, although no detailed count was made.

The proportion of fixed-wing general aviation accidents attributable to pilot error, found in the present study, is similar to that found by O'Hare et al. (1994). For fixed-wing non-agricultural aircraft, that proportion is not too different from that suggested by Nagel (1988) for accidents in the United States; the higher level of accidents due to mechanical failure in New Zealand being probably due to the less forgiving New Zealand environment, as already discussed (p. 1).

However, it is difficult to find environmental differences between New Zealand and Canada which could account for the difference between the helicopter pilot error accident rates found in the present study and those suggested by Skjenna (1981). The factors leading to the high incidence of mechanical failure in New Zealand would be likely to apply universally. It may be that Skjenna's figure for the percentage of accidents attributable to pilot error, which was only an estimate, is a little high.

Distributions of types of error, with aircraft class

The marked differences between the distributions of error types for agricultural aircraft, and other fixed-wing powered aircraft, means that the results for these two classes of aircraft must be kept separate. While the general forms of the distributions for non-agricultural aircraft resemble those found by O'Hare et al. (1994), those for agricultural aircraft are different, showing a greater proportion of action errors in fatal accidents, and a low proportion of goal errors in all accidents. In part, the markedly different operating environment of the two classes of aircraft could influence the types of errors which lead to accidents. The hazardous environment of agricultural operations is likely to penalise action errors harshly. Much agricultural flying is over hilly or mountainous terrain, there is little physical room for error, and because the operations are at low level there is little time to recover from the effects of an error.

Following the same line of reasoning as when comparing agricultural aircraft and other fixed-wing aircraft, one might expect to find significant differences between the types of errors causing accidents to helicopters engaged in venison recovery, and to other helicopters. A possible explanation for the lack of significant differences between the helicopter classes may be the small sample sizes: although there were many helicopter accidents, the proportion attributed to mechanical failure meant that the pilot error samples were small.

When comparing fatal helicopter accidents with fatal accidents to fixed- wing powered aircraft (non-agricultural), the fixed-wing accidents had a predominance of goal errors, while the helicopter accidents had a predominance of strategy errors. No explanation for this difference has been found. The general similarity in the distributions of error types for these two classes of aircraft is somewhat surprising, given the differences in their operating characteristics and operating environments.

In general, the first research hypothesis is supported by the present study. There were different distributions of types of pilot error, with various classes of aircraft. Helicopters showed different patterns from fixed-wing aircraft, and agricultural aircraft had markedly different distributions from other fixed-wing aircraft. However, no difference was found between distributions for helicopter venison recovery operations, and other helicopter flying. There were too few samples to identify differences in the case of gliders, and light sporting aircraft.

Error distributions for fatal and non-fatal accidents

For fixed-wing powered aircraft other than agricultural aircraft, the difference in the distributions of error types between fatal and non-fatal accidents is statistically significant, meaning that it is unlikely that a difference of that magnitude could have arisen by chance. However, there is a broad overlap, in agreement with the proposition of Billings and Reynard (1981), that the errors found in accidents of different degrees of severity are broadly similar. The marked shift found by O'Hare et al. (1994) from a predominance of goal errors in major accidents to procedural errors in minor accidents, was not found in the present study.

Helicopter accidents showed something of the dichotomy found by O'Hare et al. (1994) for fixed-wing aircraft, but the peaks occur at strategy selection (fatal accidents) and procedural errors (non-fatal accidents).

In the case of agricultural aircraft, the absence of a significant difference in the distribution of error types between fatal and non-fatal accidents may be due to the large number of cells with expected frequencies less than five (p. 56). A larger sample might have shown a significant difference.

The second research hypothesis is also generally supported by the present study: there are differences in the distributions of error types, for fatal and non-fatal accidents. For fixed-wing aircraft not engaged in agricultural operations, the difference was statistically significant, but of limited practical importance as the distributions were broadly similar. Helicopters showed a more noticeable difference. However, agricultural aircraft showed no significant difference. There were again too few samples to draw conclusions about gliders and light sporting aircraft.

Pilot experience

Gerbert and Kemmler (1986) and Wickens and Flach (1988) advanced theoretical grounds for suggesting that pilot experience might affect the distributions of types of pilot error. The absence of any detectable effect in the present study is, however, in accordance with Wickens and Flach's comment that there is little experimental support for the theoretical proposition.

O'Hare et al. (1994) found such an effect in their study of powered fixed-wing aircraft. Possible sources of this discrepancy could have been their inclusion of repetitive accidents by low-time pilots, and the inclusion of errors by agricultural pilots with errors by pilots of other powered fixed-wing aircraft. It is possible that repetitive errors by low-time pilots could have resulted from procedural errors often enough to produce one of the effects seen by O'Hare et al (1994) (more procedural errors than expected by pilots with less than 100 hours). However, since agricultural pilots are more experienced than most general aviation pilots, the effect of combining agricultural pilot errors with those of pilots of other powered fixed-wing errors would have been to produce a higher proportion of action or strategy errors by experienced pilots. This was not what O'Hare et al. found (more goal errors than expected, by pilots with 100 - 1000 hours). Alternatively, the effects found by O'Hare et al. may have been brought about by the inclusion of self-reported accidents in the 'minor accidents' sample, as is discussed further in the next section.

The third research hypothesis, that pilot experience would affect the distribution of types of pilot error is not supported by the present study.

Comparison with the results of O'Hare et al. (1994)

The close similarity of the distributions of error types for 'Major accidents' (O'Hare et al. (1994) and for 'Fatal accidents' from the present study argues against the possibility that it is differences in coding that brought about the differences in the 'Minor accidents' and 'Non-fatal accidents' distributions. Any bias in the 'Major accidents' results from the inclusion of 'serious injury' appears to have been unimportant, perhaps because there were not very many such accidents, although this aspect was not investigated.

The present study postulated limitations in the database used by O'Hare et al. (1994). It has already been shown (at p. 62) that agricultural aircraft accidents (fatal or non-fatal) need to be separated from accidents to other fixed-wing aircraft. However, a more serious limitation was thought to be the effect of self-reporting (described on p. 28) on the results for the 'minor accidents' category.

The discrepancy between the distributions for 'Minor accidents' (O'Hare et al. (1994)) and 'Non-fatal accidents' from the present study is that the marked peak at procedural errors found by O'Hare et al. is not found in the present study. Indeed, in the present study, both 'Fatal' and 'Non-fatal' distributions have the highest number of errors at goal selection.

The 'Non-fatal' samples in the present study were adjusted to have identical criteria to O'Hare et al.'s (1994) 'Minor accident' sample, before making the comparison. Remaining factors that could have caused the discrepancy are the inclusion by O'Hare et al. of trivial training accidents (which were not subject to investigation and so were excluded from the present study), the criteria used to determine whether a non-fatal accident should be investigated, and self-reporting in the pilots' reports of minor accidents that were not subject to official investigation.

If the training accidents were a major influence, one would expect proportionally more action errors in the distribution found by O'Hare et al. (1994), but this was not the case. (It may be that the overall number of minor training accidents was low, but this facet was not examined in the present study). We may conclude that training accidents did not influence the discrepancy between the results of the present study and those of O'Hare et al. (1994).

The selection by the investigating authority of non-fatal accidents and incidents for detailed professional investigation was not done on a systematic basis. Rather, the decision to investigate was made by considering a number of criteria. Firstly, Air Transport operations were likely to be investigated. Secondly, severe injury accidents were likely to be investigated (so that, should the pilot die, the mandatory fatal accident investigation would not be hampered by lack of on-site investigation). Thirdly, apparently severe impacts which the pilot survived might be investigated, to learn about the factors alleviating injuries. Fourthly, an accident where the pilot survived, similar in nature to other accidents which resulted in fatalities, might be investigated to gain insight into the causes of fatal accidents. Fifthly, accidents apparently resulting from a technical deficiency would usually be investigated, though they might turn out to be due to Human Factors. And lastly, accidents "of mysterious origin" might be investigated (e.g. OAAI, 1986a).

Of these criteria, the only one likely to cause the 'non-fatal' error distribution to differ from the 'minor accident' distribution found by O'Hare et al. (1994) was the inclusion of the 'severe injury' category in the 'non-fatal' grouping, whereas these accidents would have been included by O'Hare et al. in their 'Major accident' category. However, these 'severe injury' accidents were excluded from the sample of 'Non-fatal accidents' before the comparison was made, so eliminating any effect when comparing 'non-fatal' with 'minor accidents'. Thus, it seems unlikely that the criteria by which accidents were selected for official investigation would have caused the difference between the error distribution for non-fatal accidents in the present study, and the 'minor accidents' distribution found by O'Hare et al. (1994).

The remaining factor that could account for the discrepancy is self-reporting in pilots' reports of minor accidents that were not subject to professional investigation. It is concluded that self-reporting may have affected the results for 'Minor accidents' used by O'Hare et al. (1994).

The alternative view is that the differences between the 'minor accidents' and the 'non-fatal' accidents distributions of types of error could be due to the coding in the present study having been done by a single coder. The close similarity between the 'major accidents' and the 'fatal' accidents distributions argues that this was not the case. (This point is discussed in 'Limitations', p. 75).

While the results from the present study support the proposition by O'Hare et al. (1994) that there is a difference in distributions of types of errors, between major and minor accidents, the marked dichotomy which they found is attributed to the effects of self-reporting in many of the accident reports they used. In the present study a broad overlap between the 'Fatal accident' and 'Non-fatal accident' distributions was found, which lends credence to the views of Billings and Reynard (1981) that the errors which lead to accidents come from a general population, the outcome being due to chance.

CHAPTER 5
DISCUSSION

The policy of investigating incidents with a view to forestalling accidents depends for its validity on the proposition expressed by Billings and Reynard (1981), that accidents are a subset of incidents and the eventual outcome is due to chance. On the other hand, Rasmussen (1982) thought that incidents might differ from accidents by the elements of feedback or reversibility (that is, the ability of the pilot to correct an error), which could allow the pilot to intervene to prevent an accident from developing. Billings and Reynard (1981) found support for their view in the broad similarity in distributions of error types in accidents, and in incidents recorded in the ASRS database. It would be difficult to test this proposition by comparing the errors in accidents with those found in incidents which have been professionally investigated. Throughout the world, not very many incidents have been thoroughly investigated, because resources have generally been allocated to the investigation of accidents, but the distributions of error types in accidents of different degrees of severity could allow the proposition to be tested.

Jensen and Benel's (1977) finding, that different factors were associated with fatal and non-fatal accidents, was supported by the clear dichotomy found by O'Hare et al. (1994) in the distribution of error types between major and minor accidents. O'Hare et al. found that the preponderance of triggering factors in major accidents was faulty goal selection, whereas in minor accidents, incorrect performance of a procedure was the dominant factor.

However, while the present study supports the view of O'Hare et al. (1994), that there are differences in the distributions of errors between major and minor accidents, the clear dichotomy which they found in the case of powered fixed-wing aircraft was not found in the present study. Rather, there was a broad overlap with some change of emphasis. Although O'Hare et al. intended to eliminate self-reporting, in fact a large part of their database comprised self-reported accidents, and it appears that this may be the reason for the apparent dichotomy they found.

The finding of the present study, that the types of errors that lead to fatal and non-fatal accidents are broadly similar, does nothing to detract from the view of Billings and Reynard (1981). It is possible that while the differences in the factors associated with major and minor accidents, reported by Jensen and Benel (1977), and Wiegmann and Shappell (1996) were statistically significant, they may not have been of great practical importance. Whereas a marked dichotomy would suggest that lessons learned from investigating minor accidents or incidents would be of limited value in eliminating the causes of major accidents, broadly similar distributions of error types would indicate that the lessons were of general applicability.

In reviewing the non-fatal accidents for the present study, it was very clear that self-reporting had its limitations. In some cases the pilots may have been genuine in their belief of what happened, but (whether from incomplete perception or lack of knowledge) that belief was at variance with the facts found by the official investigators. In other cases the investigators openly expressed their disbelief in the pilots' veracity. The difference between O'Hare et al.'s (1994) results for 'Minor accidents' and the 'Non-fatal accidents' distributions in the present study, may be largely due to this effect.

This casts some doubt on the value of self-reported incident reports such as ASRS, in addition to the doubts raised by Nagel (1988). It would seem desirable to make some check on the extent to which such databases are affected.

The distributions of error types in accidents to fixed-wing agricultural aircraft were quite different from those involving other aircraft classes. One potential explanation is that pilots attracted to this facet of aviation, with its evident high risks, are likely to be risk-takers by nature. The alternative proposition that the difference in error patterns may be due to the more hazardous environment in which such flying is done, is not supported by the lack of such patterns in the even more hazardous field of helicopter operations. (However, the small numbers of helicopter pilot error accidents may have masked any difference). There does appear to be something which distinguishes fixed-wing agricultural flying from other types of aviation. Given the high accident rates in agricultural flying, finding the cause of this difference could be a fruitful line of investigation, since it seems that remedies particular to agricultural flying may be needed. The author is aware of only one study of the causes of accidents to agricultural aircraft in New Zealand. This study was performed by the New Zealand Office of Air Accidents Investigation in the 1960s, but the present author has been unable to locate a copy of it, or any file material.

In a human factors investigation, frequently the first stumbling block is the question, "what was the pilot trying to do?" It is unhelpful to substitute what the investigator thinks the pilot ought to have been trying to do: the investigator knows the outcome, and is subject to 'hindsight bias' (Reason, 1990, p. 215). The taxonomic algorithm designed by O'Hare et al. (1994) can be of assistance in solving such problems, and in the present author's opinion it will prove to be a valuable practical investigation tool.

The present study supports the view of Wiegmann and Shappell (1996) that the six-part taxonomy which O'Hare et al. (1994) derived from the work of Rasmussen (1982) and Rouse and Rouse (1983) is able to describe the great majority of pilot errors. Indeed, the author found no pilot error accidents in the New Zealand database which could not be coded with it.

However, the present study would also suggest that there is some difficulty in applying the algorithm. Even when used with the detailed instructions (Appendix A), repeatability was only achieved after considerable practice. In other words, like any other tool, training in its use will produce better results.

Dr. D. O'Hare has suggested (personal communication) that the instructions need further refinement, and instanced the example of coding the accident to ZK-GSG (p. 41), where a different coding event would be chosen, depending on the view of where the 'flight' began. ICAO defines the flight as commencing when persons first board the aircraft for the purpose of getting airborne (ICAO, 1994). If this definition was used, the accident to ZK-GSG would be coded as a procedural error, because pre-flight actions by the pilot would be excluded. This definition would exclude errors in flight planning, assessment of meteorological conditions and pre-flight preparation, all regular contributors to accidents, and in the present study the author considered that the flight began when the pilot began to prepare for it. Precise definitions would help to ensure that different coders would come to the same conclusions.

Limitations

Given the somewhat subjective nature of the coding procedure, it would have been desirable to have had more than one coder. The time required to code the large number of accidents made this impracticable. However, codings proved repeatable and were in reasonable agreement with those of O'Hare et al. (1994). In the case of the major accidents to fixed wing aircraft, the overall agreement was very close. While one coder might perceive factors differently from others, this will not affect the differences between sets coded by one person, provided that consistency is achieved.

Even though the database comprised virtually the entire population of reports of investigated accidents in New Zealand, there were insufficient samples in some categories for valid analyses. To overcome this limitation, it would be necessary to refer to larger databases such as that of the United States National Transportation Safety Board.

Further Study

Incident Reporting Systems

The present study has not found differences of practical importance between the distributions of errors in major and minor accidents. The concern that such differences, found by O'Hare et al (1994) might imply differences between incidents and accidents (so that the study of incidents might not be a valid means of reducing the number of accidents) is thus allayed. However, the present study provides practical support for the theoretical views expressed by Nagel (1988) that the value of confidential incident reports might be reduced because of the limitations of self-reporting by the pilots involved. It would seem desirable that this concern be investigated. Two possible approaches would be either an investigation on similar lines to the present study, but including incidents from a large database such as ASRS; or professional investigation of incidents which are reported to an organization whose incident reporting system could enable such a follow-up.

The Cognitive Failure Algorithm

The algorithm devised by O'Hare et al. (1994) has the potential to be a practical tool for use in the course of investigations. It should enable investigators to understand 'what happened' in terms of cognitive failure by the pilot, in addition to the present ability to analyze 'how it happened' and 'why it happened', so giving a comprehensive picture of the factors behind the pilot errors which are at the root of so many accidents. However, before it could be used in this way, it would be necessary to improve the instructions so that consistent and reliable coding could be achieved. It would also be desirable to devise a presentation to show practising investigators why, when and how to use the algorithm.

Agricultural Aircraft Accidents

The fixed-wing agricultural aircraft fatal accident rate in New Zealand is about 1 per 5000 hours. (See, for example, OAAI, 1982). The present study has shown that the pilot errors which led to these accidents are different in kind from the errors in other types of general aviation. A study aimed at finding the reasons for this difference might well lead to measures with the potential to reduce substantially the accident rate for agricultural aircraft.

REFERENCES

Air Accident Investigation Branch. (1988). Aircraft Accident Report 8/88. *Report on the accident to Boeing 737-236 G-BGJL at Manchester Airport on 22 August 1985.* London: HMSO.

Air Accident Investigation Branch. (1990). Aircraft Accident Report 4/90. *Report on the accident to Boeing 737-400 G-OBME at Kegworth on 8 January 1989.* London: HMSO.

Billings C. E. & Reynard, W. D. (1981). Dimensions of the information transfer problem. In C. E. Billings and E. S. Cheaney (Eds.) *Information Transfer Problems in the Aviation System*. Moffet Field, CA: NASA Ames. Research Centre, pp. 9 - 14, NASA-TP-1875. Cited in Rouse and Rouse (1983), and in Nagel (1988).

Boeing Commercial Airplane Company. (1985). *Statistical summary of commercial jet aircraft accidents, worldwide operations, 1959 - 1984.* Seattle: Author.

Bureau of Air Safety Investigation. (1984). *Aircraft accidents and air safety incidents: data recording system.* (3rd ed.). Canberra: Author.

Caesar, H. (1987). Safety statistics and their operational consequences. *Proceedings of the 40th International Air Safety Seminar.* (pp. 13-20). Arlington, Virginia, USA. Flight Safety Foundation.

Dane, F. C. (1990). Research methods. Pacific Grove, California: Brooks/Cole.

Fitts, P. M. & Jones, R. E. (1947). Psychological aspects of instrument display: analysis of 270 "pilot error" experiences in reading and interpreting aircraft instruments. In H. W. Sinnaiko (Ed.). (1961) Selected papers on human factors in the design and use of control systems (pp. 359-396). Cited in Wickens and Flach (1988), and O'Hare et al. (1994).

Fleiss, J. L. (1981). Statistical methods for rates and proportions. New York: Wiley. Cited in O'Hare et al. (1994).

Gerbert, K., & Kemmler, R. (1986). The causes of causes: determinants and background variables of human factor incidents and accidents. Ergonomics. 29. 1439-1453.

Helmreich, R. L. (1990). Human factors aspects of the Air Ontario crash at Dryden, Ontario. In Moshansky, V. P., Final Report of the Commission of Inquiry into the Air Ontario Crash at Dryden, Ontario. (Technical appendices, pp. 319-348). Ottawa: Minister of Supply and Services.

Hunt, G. J. F., & Macfarlane, R. (1993). Aviation human factors: applying human factors to the flight deck. Palmerston North, NZ: Massey University.

International Civil Aviation Organization. (1994). International standards and recommended practices: aircraft accident and incident investigation. (8th ed.). Ontario: Author.

Jensen, R. S. & Benel, R. A. (1977). Judgement evaluation and instruction in civil pilot training. DOT/FAA Report RD-78-24. Springfield, Va: NTIS. Cited in O'Hare et al. (1994).

Mahon, P. T. (1981). Report of the Royal Commission to inquire into the crash on Mount Erebus, Antarctica, of a DC 10 aircraft operated by Air New Zealand Limited. Wellington: Government Printer.

Moshansky, V. P. (1992). Final report of the commission of inquiry into the Air Ontario crash at Dryden, Ontario. Ottawa: Minister of Supply and Services.

Nagel, D. C. (1988). Human error in aviation operations. In E. L. Wiener and D. C. Nagel (Eds.) Human Factors in Aviation. (pp. 263-303). New York: Academic Press.

Office of Air Accidents Investigation. (1980). McDonnell-Douglas DC10-30 ZK-NZP. (Report No. 79-139). Wellington: Government Printer.

Office of Air Accidents Investigation. (1982). New Zealand civil aircraft accidents, 1982. Wellington: Government Printer.

Office of Air Accidents Investigation. (1986a). Bell 206b Jetranger ZK-HXT. (Report No. 86-081). Wellington: Government Printer.

Office of Air Accidents Investigation. (1986b). Siren Pik 30 motor glider ZK-GSG. (Report No. 86-096). Wellington: Government Printer.

Office of Air Accidents Investigation. (1987). Lancer 4L hang glider. (Report No. 87-078). Wellington: Government Printer.

Office of Air Accidents Investigation. (1989a). Cessna C188b ZK-CSC. (Report No. 89-088). Wellington: Government Printer.

Office of Air Accidents Investigation. (1989b). Cessna C188b ZK-DMA. (Report No. 89-045). Wellington: Government Printer.

Office of Air Accidents Investigation. (1989c). Fokker F27-100 ZK-BXG. (Report No. 88-0-438). Wellington: Government Printer.

O'Hare, D., Wiggins, M., Batt, R., & Morrison, D. (1994). Cognitive failure analysis for aircraft accident investigation. Ergonomics. 37 (11). 1855-1870.

Rasmussen, J. (1980). What can be learned from human error reports? In: K. D. Duncan, M. Gruneberg & D. Wallis (Eds.). Changes in working life. New York: Wiley.

Rasmussen, J. (1982). Human errors: a taxonomy for describing human malfunction in industrial installations. Journal of Occupational Accidents. 4. 311-335.

Reason, J. (1990). Human error. Cambridge: Cambridge University Press.

Reason, J. (1991). Identifying the latent causes of aircraft accidents before and after the event. Proceedings of the 22nd International Seminar of the International Society of Air Safety Investigators. (pp 39 - 45). Sterling, Virginia, USA: ISASI.

Rolfe, J. M. & Bekerian, D. A. (1985). Witnesses. Journal of the International Society of Air Safety Investigators. 85 (4). (pp. 23 - 25). Sterling, Virginia, USA: ISASI.

Rouse, W. B. & Rouse, S. H. (1983). Analysis and classification of human error. I.E.E.E. transactions on systems, man and cybernetics. SMC-13 (4). 539-549.

Saunders, G. H. (1975). Dynamics of helicopter flight. New York: Wiley.

Senders, J. W., & Moray, N. P. (1991). Human error: Cause, prediction and reduction. Hillsdale, NJ: Lawrence Erlbaum.

Skjenna, O. W. (1981). Cause factor: human. A treatise on rotary wing human factors. Ottawa: Minister of National Health and Welfare.

Transport Accident Investigation Commission. (1993). Pterodactyl Ascender 11 + 2 ZK-FKF. (Report No. 92-007). Wellington: Author.

Vette, A.G. (1997). Impact Erebus (Rev. ed). Auckland: Aviation Consultants.

Wickens, C. D. & Flach, J. M. (1988). Information processing. In E. L. Wiener and D. C. Nagel (Eds.). Human factors in aviation (pp. 111-155). London: Academic Press.

Wiegmann, D. A., & Shappell, S. A. (in press). Human factors analysis of post-accident data: applying theoretical taxonomies of human error. International Journal of Aviation Psychology.

Zotov, D. V. (1995a). Helicopter accident investigation. Palmerston North: Massey University.

Zotov, D. V. (1995b). Categories of pilot error in aircraft accidents. Unpublished manuscript, Massey University, Palmerston North, New Zealand.

Zotov, D. V. (1996). Reporting human factors accidents. The Journal of the International Society of Air Safety Investigators. 29 (3). 4-20.

Appendix A

ERROR CLASSIFICATION PROCEDURE

Section 1

1.1 Was the information necessary to recover from the situation or minimise the damage to the aircraft or its occupants, available to effect a timely intervention?

Yes Next question
No Go to Section 2
No Inf Go to Section 2

1.2 Did the pilot observe any of the information which would have allowed the recovery of the situation or the minimisation of the damage to the aircraft or its occupants?

Yes	Go to Section 2	
No	Information Error	Code 1
No Inf	Go to Section 2	

Section 2

2.1 Did the pilot attempt to diagnose the state of the system on the basis of the information available?

Yes	Next Question	
No	Failed to Diagnose system state	Code 2/1
No Inf	Go to Section 3	

2.2 Did the pilot complete the diagnosis of the system state?

Yes Next Question
No Incomplete Diagnosis Code 2/2
No Inf Go to Section 3

2.3 Did the pilot diagnose accurately the information available concerning the state of the aircraft?

Yes Go to Section 3
No Next Question
No Inf Go to Section 3

2.4 Was this a major contributing factor in the subsequent accident/incident?

Yes Misdiagnosis of system state Code 2/3
No Go to Section 3

Section 3

3.1 Did the pilot have a goal in mind when the accident/incident occurred?

Yes Next Question
No Failed to formulate Goal Code 3/1
No Inf Go to Section 4

3.2 Was the goal reasonable under the circumstances (e.g. weather, experience, etc)

Yes Go to Section 4
No Incorrect Goal Code 3/2
No Info Go to Section 4

Section 4

4.1 Did the pilot choose a strategy* with which to achieve the goal?

Yes Next Question
No Failed to adopt strategy Code 4/1
No Info Go to Section 5

4.2 Was the strategy chosen the most appropriate one under the circumstances?

Yes Go to Section 5
No Incorrect Strategy Code 4/2
No Info Go to Section 5

Section 5

5.1 Did the pilot attempt to execute a procedure** consistent with the strategy?

Yes Next Question
No Failed to execute procedure Code 5/1
No Inf Go to Section 6

5.2 Was the procedure the most appropriate under the circumstances?
Yes Next question
No Incorrect Procedure Code 5/2
No Inf Go to Section 6

Section 6

6.1 Was the procedure carried out as intended?

Yes	No Errors	Code 0
No	Action Error	Code 6

* Strategy: finding the means to satisfy a goal

** Procedure: a specification for conducting a set of predetermined subtasks or actions that are components of a higher-level task

(Note: the 'fine grain' coding, which subdivides the categories, gives extra detail which may be useful in subsequent studies).

3

Massey University Library Thesis Copyright Form

Classification of Pilot Error in Aircraft Accidents

(1) I give permission for my thesis to be made available to readers in Massey University Library under conditions determined by the Librarian.

(2) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.

(3) I agree that my thesis may be copied for Library use.

Signed
Date

The copyright of this thesis belongs to the author. Readers must sign their name in the space below to show that they recognise this. They are asked to add their permanent address.

NAME AND ADDRESS	DATE