CIVIL AVIATION AUTHORITY

CAA PAPER 99001

PILOT INTERVENTION TIMES IN HELICOPTER EMERGENCIES

CIVIL AVIATION AUTHORITY, LONDON, PRICE £5.00

CAA PAPER 99001

PILOT INTERVENTION TIMES IN HELICOPTER EMERGENCIES

PLSD/CHS/HS3/CR97020/1.0

J W Chappelow
P R Smith

CIVIL AVIATION AUTHORITY, LONDON, JANUARY 1999

ISBN 0 86039 750 5

Printed and distributed by
Westward Digital Limited, 37 Windsor Street, Cheltenham, England

Foreword

The research reported in this paper was funded by the Safety Regulation Group of the UK Civil Aviation Authority, the UK Department of Transport, and the UK Health and Safety Executive. The work was instigated at the DERA Centre for Human Sciences in response to Recommendation 4.4 of AAIB Aircraft Accident Report 4/83, (accident to Westland Wessex 60 G-ASWI 12 miles ENE of Bacton, Norfolk on 13 August 1981). Impetus was subsequently added by Recommendation 4.3 of AAIB Aircraft Accident Report 7/87, (accident to Twin Squirrel AS355 G-BKIH at Swalcliffe, near Banbury, Oxfordshire on 08 April 1986) and reocmmendation 92-26 of AAIB Accident Report EW/C92/2/4 (accident to Robinson R22M G-BPPC at Oldham in February 1992).

The CAA concurs with the conclusions of this work. Current civil requirements specify a 'corrective action time delay' of 1 second or 'normal pilot reaction time' (whichever is greater) for the case of engine failure in the cruise. This is commonly interpreted as a 1 second time delay in the cruise, however 0.3 seconds is typically used for other flight phases. The term 'corrective action time delay' is generally equivalent to the period between failure onset and the first movement of the appropriate control, referred to as detection time in this report. This time delay is clearly inconsistent with the results and should be reviewed in the light of the data presented. For AFCS equipment, current civil IFR requirements state that the AFCS must be designed so that it cannot create a hazardous deviation in flight path in the event of malfunction or failure, assuming corrective action begins within 'an appropriate period of time'. Advisory material sets out these appropriate delay times as between 0.8 and 3.9 seconds, depending on phase of flight. Although these times are more consistent with the results of this research, a review of the appropriate regulatory material would be highly desirable.

At the time of publication, the JAR/FAR Parts 27 and 29 engine failure intervention time criteria are under review by the rotorcraft Performance and Handling Qualities Harmonisation Working Group. This group consists of authorities and industry representatives and was formed to consider a number of flight related requirement issues and agree a harmonised set of proposed changes. The results of this research have been presented to the Group, and are contributing to the debate on proposed amendments to the requirements and associated advisory material.

Executive Summary

Total power failure, tail rotor control failure and tail rotor drive failure in helicopters are time critical emergencies. Automatic flight control system failures may also demand a prompt response in some circumstances. Detection times and response times to total power failures were measured during routine training sorties in Chinook, S61N and Super Puma simulators. The effects of distraction on reaction time were investigated by introducing a minor failure immediately before total power failure on some sorties in the Super Puma. Data were also collected during routine training sorties for responses to tail rotor drive failure and tail rotor control failure in the Super Puma simulator. The required response for all these emergencies was lowering of the collective lever. Detection times for automatic flight control system pitch runaways were measured in the S61N during dedicated experimental sorties. The response in this case was a cyclic input. An attempt was made to identify the principal alerting cues for each failure in each simulator.

Total power failure: The mean detection time was in the range 1.07 to 2.22s and the 90th percentile was in the range 2.07 to 2.95s. The mean response time was in the range 1.24 to 1.82s and the 90th percentile was in the range 1.95 to 3.85s. The mean total reaction time was in the range 2.3 to 4.13s and the 90th percentile was in the range 4.42 to 5.72s. There was no significant effect due to distraction: The mean detection time for these trials was 2.31s and the 90th percentile was 4.82s; the mean response time was 1.65s and the 90th percentile was 2.77s; the mean total reaction time was 4.16s and the 90th percentile was 5.1s.

Tail rotor control failure: The mean detection time was 0.9s and the 90th percentile was 1.22s. The mean response time was 1.53s and the 90th percentile was 3.78s. The mean total reaction time was 2.57s and the 90th percentile was 4.65s.

Tail rotor drive failure: The mean detection time was 1.53s and the 90th percentile was 2.94s. The mean response time was 2.74 s and the 90th percentile was 6.06 s. The mean total reaction time was 4.9s and the 90th percentile was 7.79s.

Automatic flight control system failure: The mean detection time was 0.4s in the hover and 1.59s in the cruise; the 90th percentile was 0.69s in the hover and 2.4s in the cruise. The mean effective total time required to control the failure was 1.62s in the hover and 2.24s in the cruise; the 90th percentile was 1.83s in the hover and 2.61s in the cruise.

The difference in detection times for automatic flying control system failures in the hover and the cruise (90th percentiles of 0.69s and 2.4s respectively) demonstrates the effect of the pilot paying close attention to a control variable which is immediately disturbed by the failure. A safe regulatory approach would in general have to cater for less focused attention.

Excluding the automatic flying control system failure in the hover (as a case of focused attention) leaves a range of 90th percentile detection times generally between 2 and 3s. Only the tail rotor control failure case provides a shorter 90th percentile detection time. Acceptance of a 10% failure rate would in general require an allowance of at least 2 and preferably 3s for detection time. Current regulations should be reviewed with this in mind.

List of Contents

1 INTRODUCTION

This work was undertaken for the Civil Aviation Authority under assignment 502HX-001 and 39CXC-001.

Few emergencies in modern aircraft require an immediate response. In multi-crew aircraft, challenge and response drills and crew consultation are the norm. Even in single crew aircraft, the priority is generally to fly the aircraft and achieve a safe attitude, altitude and airspeed, before dealing with the emergency. A small number of emergencies are, however, time critical. Total power failure, tail rotor control failure and tail rotor drive failure in helicopters are examples of such exceptions. An undue delay in reducing collective pitch demand in response to total power failure may result in an unrecoverable decay in rotor speed. Tail rotor control and drive failures also require a prompt reaction. Automatic flight control system (AFCS) runaways may also demand prompt intervention depending on the rapidity and extent of the excursion. A realistic estimate of pilot reaction time is an essential tool in the design and certification of helicopters.

Reaction times have long been of experimental interest, and a great deal is known about the factors affecting speed of reaction. Some of these factors have an obvious bearing on response to time critical emergencies in helicopters. For example: stimuli that occur only infrequently tend to elicit longer reaction times; reaction times increase as the number of possible stimuli and responses increases (in a logarithmic relationship); enhanced motivation speeds reactions; practice reduces reaction times, particularly when discrimination among alternatives is difficult; a stimulus with unique characteristics elicits faster reactions than one that shares characteristics with other possibilities; sharing attention among tasks lengthens reaction times (ref. 1). The relevance of all these factors is evident, but generalisation from laboratory to real-life tasks is inevitably in global terms rather than precise predictions. Pilots are generally well trained to identify the combination of visual, auditory and motion cues characteristic of time critical emergencies and their motivation cannot be in doubt. On the other hand, these are relatively improbable events, and may well occur in a period of high workload or distraction. It is not possible, therefore, to specify with confidence the range of response times to be expected in real operations purely on the basis of laboratory findings.

Unfortunately, in comparison with the extensive and detailed laboratory studies of reaction times, there has been little work examining reaction times in real life situations. One unpublished trial in support of an accident investigation did address the issue of total power failure in a helicopter using a S61N flight simulator (ref. 2). Total reaction times ranged from 1.5 to 5.5s, with a mean of 3.08s. The aim of the present trial was to provide more detailed information on which to base recommendations regarding certification requirements for helicopters. In order to achieve this aim, clear definitions of the components of total reaction time are required. Current regulations use a variety of terms which are not operationally defined. A necessary methodological step is, therefore, the development of suitable definitions.

2 METHOD

2.1 Simulators

Chinook Mk1, S61N and Super Puma simulators were used in the trial. All three had six axis motion platforms and dusk/night visual systems. Two separate trials were undertaken using the S61N simulator. The first addressed total power failures; the second was designed principally for AFCS runaways, but total power failure data were also collected.

2.2 Cues

A variety of visual, auditory and motion cues could alert the pilot to the emergencies being studied or provide diagnostic information. Video recordings were made in all three simulator cockpits during exploratory flights in order to evaluate the visual and audio cues presented during total power failures. Similar recordings were made for tail rotor failures in the Super Puma. For AFCS failures in the S61N, direct computer recordings of instrument indications were used. Initial height and speed, temperature and all up weight were chosen to be representative of typical training sorties. Note that ambient noise was recorded, not the audio cues available at the pilot's ear, which would, of course, be somewhat modified by headsets or helmets. Changes in the frequency content of the noise were of more interest than those in absolute levels, so this approach seemed appropriate at this stage. An attempt was made to measure linear accelerations on the flight deck during total power failures using accelerometers mounted between the pilots' seats on the centreline of the aircraft or directly beneath the pilot's seat.

2.3 Subjects

To be able to record response times that are as realistic as possible, it is preferable for subjects to be unaware that they are participating in an experiment, and this was the approach adopted with the exception of the second S61N trial. Permission to collect the data was given by those responsible for the pilots' training. Twenty-four RAF pilots were tested in the Chinook simulator, some more than once; a total of 35 total power failure exposures were recorded. Ten civil pilots provided response time data in the first S61N total power failure trial. The subjects for the Super Puma were also civil pilots; 49 cases of total power failure, 21 of tail rotor control failure, and 40 of tail rotor drive failure were recorded. In the second S61N trial 16 civil pilots were exposed to AFCS runaways and total power failures.

2.4 Procedure

For the Chinook and the first S61N total power failure trials and the tail rotor failures, the failures were initiated during routine training sorties by an instructor operating a switch. No attempt was made to restrict or control the conditions surrounding the failure so as not to compromise the normal training programme. The circumstances surrounding the Super Puma failures were, however, more uniform than those for the Chinook or the first S61N trial. Where possible a record was made of the main variables (altitude, airspeed etc.) at the onset of the simulated failure. A randomly selected sub-group of the total power failures in the Super Puma was preceded by an alternator failure. This was a minor emergency designed to act as a more or less standardised distraction. In the second S61N trial two AFCS runaways in pitch were presented in each sortie; the first in the hover, the second in the cruise when the pilot might be expected to be slightly less attentive. This was followed by a total power failure during an instrument procedure to position for an approach. Although these were dedicated, non-training sorties, the pilots were not aware of the specific form of the trial.

2.5 Data recording

With the exception of the second S61N trial, existing simulator software was used (modified if necessary) to collect data from two channels at as high a sampling rate as possible (30 Hz for the Chinook and Super Puma, 9 Hz for the S61N). In the case of the Chinook and S61N, where only total power failures were investigated, torque and collective position were recorded. The decline in torque was used to define the onset of the failure. In the second S61N trial an independent data logging system recorded torque, pitch, collective position, and longitudinal cyclic position at 33 Hz.

In the Super Puma, collective position was recorded, but there was no one performance variable suitable to define failure onset across three different emergencies (total power failure, tail rotor control failure, and tail rotor drive failure). It was necessary, initially at least, to take the start of data recording as synchronous with onset of the failure.

The initiation of a control response was taken to be the first noticeable deviation from steady state or, if the control was travelling in the wrong direction, the point of reversal was used. Collective responses were considered complete when the collective reached minimum or a steady state below the start position. The end point for the cyclic response to an AFCS runaway is essentially indeterminate.

2.6 Performance variables

Detection time was defined as the interval between failure onset and the initiation of a response. Response time was defined as the period from initiation to completion (where completion was identifiable). Total reaction time was defined as the detection time plus the response time. Rabbitt has argued cogently against simple partitioning of reaction times into serial processes, but for the present purpose this approach seems appropriate (ref. 3). Current regulations use a variety of terms, such as recognition time and decision time, which are not operationally defined, i.e. they are not associated with observable behaviours. Other terms, such as recovery time, are vague because the end point may not be easily defined or may depend on circumstances and aircraft dynamics. Thus, although the present definitions should allow regulations to be objectively defined, they have only a partial correspondence with current regulatory terminology.

3 RESULTS

3.1 Caveats

The results for each type of failure are presented separately. The statistical details are in Annex A. Some general comments are appropriate first.

3.1.1 *Response characteristics*

Responses for total power failures and tail rotor failures were expected to conform to a fairly well defined pattern which allowed the start and end of each control movement to be identified with some confidence. In a small proportion of cases, due to hesitations or control reversals, this confidence level was less than ideal. These data were excluded from the detailed analyses.

The end point of the cyclic response to an AFCS runaway was not definable since pilots generally had a variety of options open to them once the failure was under control. Response time and total reaction time were, therefore, notional variables. It was possible, however, to identify the point at which the pitch change due to the AFCS failure halted or began to reverse. The difference between this time and the time of the failure onset was taken as the effective total reaction time.

3.1.2 *Motion cues*

Several attempts were made during the initial trials to assess potential motion cues by recording linear accelerations in each of the different simulators. Even after problems due to noise had been eliminated, there was no evidence of significant linear accelerations (>0.1g) during the onset of total power failures or tail rotor failures. Motion cues during AFCS runaways were subjectively assessed as being even less provocative.

3.1.3 *Repetition*

Some sorties, mainly in the Super Puma simulator, involved repeated exposure to the same emergencies in the course of a single training session. Not unexpectedly, analysis showed a significant difference between first and subsequent exposures. For the purposes of this report only first exposures have, in general, been included in the analyses. An exception was made in the analysis of outcome (i.e. crashed or recovered) data because so few data points were available.

During the second S61N trial, AFCS runaways were presented during relatively demanding and undemanding phases of flight. Unfortunately, due to logistical constraints, the more demanding phase always preceded the less demanding, and the last failure in the sortie was always a total power failure.

3.1.4 *Failure onset – Super Puma*

The lack of a performance variable to define failure onset in the Super Puma recordings proved problematic when a significant number of unrealistically short response times were found. This finding raised doubts as to the wisdom of assuming that the start of data recording in the simulator more or less coincided with the first symptoms of the failure. By attaching a separate data logging system to the simulator computer it was possible to compare recordings of control activity produced by the simulator software with more comprehensive data records. This comparison suggested that, on average, data recording lagged some 0.33s behind failure onset. This correction factor has been added to the Super Puma data in the analysis that follows.

3.2 Total power failure

3.2.1 *Cues*

3.2.1.1 Visual cues

The video recordings taken during the exploratory flights with no pilot intervention were analysed frame by frame for visual cues. The time lines below describe the sequence of the more noticeable cues.

Table 1 Time line for visual cues in the Chinook

Time (s)	*Event*
0.0	Torque starts to decline
0.5	Rotor RPM begins to decline from 100%
0.8	Torque reaches minimum position at 0%
3.8	RPM reaches minimum position at 84%; generators trip off line; situation unrecoverable

Table 2 Time line for visual cues in the S61N

Time (s)	*Event*
0.0	Torque starts to decline
0.6	Rotor RPM starts to decline from 100%
2.1	Torque reaches minimum at 0%
2.2	Airspeed starts to decline from 104 kt Aircraft starts to roll to the left
3.8	RPM needles split
4.6	Roll reaches maximum at 20°
7.8	Pitch up reaches maximum at 15°
10.0	Rotor RPM reaches minimum at 85%; situation unrecoverable
11.1	Airspeed reaches minimum at 70-80 kt

Table 3 Time line for visual cues in the Super Puma

Time (s)	*Event*
0.0	Engine RPM starts to decline
3.3	Airspeed starts to decline from 130 kt
4.3	Engine RPM at minimum
5.0	Aircraft starts to roll to the right and yaw to the right
5.4	Aircraft at 90° roll; situation unrecoverable

3.2.1.2 Auditory cues

The video recordings were also analysed for the auditory cues to total power failure. In all three aircraft a change in noise was detectable as soon as the emergency was initiated.

In the Chinook, six unrelated frequencies changed over a 2.88s period from the moment the failure occurred, and seemed to represent the forward transmission slowing down.

Table 4 Summary of frequency changes in the Chinook

initial frequency (Hz)	*final frequency (Hz)*
5225	4425
3513	3000
2625	2187
1187	994
656	560
310	90

In the S61N, the auditory cues were not as pronounced as those in the Chinook. A frequency ripple lasting 8.3s started at 7525 Hz and declined to 2112 Hz. In the Super Puma, four frequencies (7.4 kHz, 7.1 kHz, 6.1 kHz and 5.9 kHz) decayed to about 3 kHz over 4s, then disappeared into the background noise. A tone of about 8.2 kHz appeared at failure onset then decayed like the others. A 5.4 kHz tone declined in amplitude from about 2.5s after the start of the failure. The Super Puma also produced an audio warning for low rotor speed (less than 245 r.p.m.) in the pilots' headset.

3.2.2 *Initial conditions*

The initial conditions are summarised in the table below for the Chinook, S61N and Super Puma. Note that the failures in the second S61N and the Super Puma trials were given in a more standardised set of circumstances, so minima and maxima are not recorded.

Table 5 Summary of initial conditions

		Chinook	*S61N trial 1*	*S61N trial 2*	*Super Puma*
altitude (ft)	min.	200	500		
	max.	6000	2000		
	mean	2724	1086	3000	2000
indicated airspeed (kt)	min.	90	80		
	max.	135	100		
	mean	115	86	not recorded	130
manoeuvring (climb/turn)	yes	13	4	16	0
	no	17	3	0	49

Results

The Super Puma failures were all initiated during the downwind leg of the circuit at 15° of collective pitch. Some were preceded by a distracting alternator failure.

3.2.3 *Reaction times*

Mean detection time, response time and total reaction time for the three aircraft types are summarised in Tables 6 to 8 together with the minimum and maximum values, and 90th percentiles. In the case of the Super Puma data it was sometimes difficult to determine with any confidence where the control movement ended. As a result, rather fewer response time and total reaction time results are quoted than detection time results. Analysis of the two S61N trials showed a significant difference in response times, but detection times and total reaction times were not significantly different. The data for the two trials are combined here.

Table 6 Summary of detection time data

Aircraft type	*No. of pilots*	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Chinook	24	0.37	3.62	1.07	2.07
S61N	23	0.66	3.09	1.67	2.53
Super Puma (undistracted)	29	1.02	3.14	2.22	2.95

Table 7 Summary of response time data

Aircraft type	*No. of pilots*	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Chinook	24	0.54	2.93	1.24	2.28
S61N	23	0.93	3.00	1.61	3.85
Super Puma (undistracted)	14	1.11	2.44	1.82	1.95

Table 8 Summary of total reaction time data

Aircraft type	*No. of pilots*	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Chinook	24	1.00	4.72	2.30	4.42
S61N	23	1.86	6.18	3.33	5.72
Super Puma (undistracted)	14	2.74	5.08	4.13	4.73

3.2.4 *Initial conditions and helicopter type*

The first analysis included Chinook data and data from the first S61N trial only (ref. 4). Initial conditions did not appear to have a significant effect on detection or response times. They were not included in subsequent analyses. An analysis including all three aircraft types, and based on helicopter type as the single independent factor, did show significant differences in detection, response and total reaction times. These times were all shorter for the Chinook than for the other two types.

3.2.5 *Distraction*

A separate analysis on the Super Puma data (see Annex A) showed no significant effect on detection, response or total reaction times due to the presence of a distraction. It is notable, however, that the distracted case produced a 90th percentile detection time nearly 2s longer than the undistracted case. This suggests that a small proportion of the pilots suffering the prior distraction may have been rather more severely affected than the majority.

3.2.6 *Outcome*

There were insufficient data on outcome (recovered/crashed) to permit analysis for the S61N and Super Puma. Of the 30 Chinook exposures for which outcome data existed, 26 pilots successfully recovered from total power failure, and four failed to recover. Failures from which the pilots recovered involved significantly shorter response times than those which resulted in crashes, but detection times were not significantly different. One of the four detected and responded to the emergency in good time, but allowed an unusual position to develop which prevented proper recovery. His data are not, therefore, relevant to the issue under consideration. The remaining three were not distinguished from the rest of the sample by initial conditions. The pattern of their responses did, however, appear qualitatively different from the vast majority of the remainder in that the collective movement was not a simple, rapid reduction in demand, but incorporated reversals or hesitations.

3.2.7 *Discussion*

It took pilots, on average, about 1s to detect total power failure in the Chinook simulator and a similar time to respond to it. S61N and Super Puma pilots took significantly longer for detection and there was a trend towards slightly longer times for the response. Mean total reaction times were, as a result, 1 to 1.8s longer. The S61N clearly presents a less critical problem in total power failure than the Chinook or Super Puma. The situation becomes unrecoverable in 10s rather than less than 4s for the Chinook or about 5s for the Super Puma. This fact may have some bearing on the difference between the Chinook and S61N results. General differences between military and civil training may also explain some of the difference between the Chinook and the other two helicopter types.

Given that visual cues probably provided diagnostic rather than alerting information, and that motion cues did not appear to be particularly effective (in the simulators, at least), auditory cues were probably the most significant alerting stimuli. There is some reason to believe that auditory cues in the S61N were less attention getting than those in the Chinook, at least. This difference may have made a small contribution to the difference in detection times. The results do not suggest a major advantage due to the low rotor speed warning system in the Super Puma.

The results indicate the broad range of response times to be expected, and suggest that average total reaction times in the range 2 to 4s should be expected in most circumstances. It is clear, however, that aircraft designers and certifiers should consider not mean reaction times but the tail of the distribution, i.e. those pilots whose total reaction times are longer than average without being abnormal. The 90th percentiles fall in the range 4.4 to 5.7s (2.07 to 2.95s for detection time). If the 90th percentile is an acceptable upper bound for reaction time, then a value of about 6s for total reaction time and 3s for detection time would seem to be the appropriate criteria to adopt for design and certification purposes.

In the Super Puma it was possible to compare responses with and without a deliberate distraction, but not to demonstrate a statistically significant effect, although the trend was in the expected direction and there is some indication that a few pilots may have suffered a significant effect. The nature and timing of distractions are clearly immensely important in determining their effect. The context must also have a bearing, and the context of a simulator sortie probably enhances overall alertness and tends to diminish the impact of minor distractions.

Notwithstanding the lack of evidence for a major advantage due to the low rotor speed warning system in the Super Puma, it is conceivable that reaction times may be shortened if warning systems were introduced, particularly in aircraft where the alerting cues are not very noticeable. The results suggest there could be some advantage in attempting to mimic or enhance the auditory cues to which pilots are already sensitive. Such an ecologically valid audio warning could have several advantages: It could require less interpretation than a conventional audio warning, and could support faster diagnosis than a voice warning. It is now well recognised that the number of audio warnings a pilot can distinguish is limited; the proposed system would be in a clearly separate class, and may not need to be counted in this number. Other technical approaches are possible (ref. 5).

3.3 Tail rotor control failure (Super Puma only)

3.3.1 *Initial conditions*

The failure was initiated on a transit between 500 and 1000 ft with 15.5° collective pitch. A failure of this sort was expected some time during the sortie, but the exact timing was unknown.

3.3.2 *Cues*

Table 9 shows the main visual cues available during the emergency with no pilot input. Auditory cues were not evident.

Table 9 Time line for visual cues

Time (s)	*Event*
0.0	Aircraft starts to yaw to the left
0.4	Aircraft starts to roll; airspeed starts to decline from 120 kt
1.2	Heading changed by 50°
1.5	Roll angle reaches 90°; situation unrecoverable

3.3.3 *Reaction times*

Collective position was recorded. Some response times could not be confidently determined. The effect of rejecting these data was probably to eliminate longer rather than shorter responses. This artificial censoring would have the effect of making the data distribution less skewed. As a result the 90th percentiles for response time and total reaction time are probably slight underestimates. In addition, there are obviously fewer response and total reaction time data points than detection time data points.

Table 10 Summary of collective reaction data

Time	*No. of pilots*	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Detection	14	0.63	1.51	0.90	1.22
Response	10	0.89	4.01	1.53	3.78
Total reaction	10	1.51	5.13	2.57	4.65

3.3.4 *Discussion*

In some cases tail rotor control failure can be a relatively benign problem involving freezing the tail rotor pitch at a moderate setting. The case presented here was more dramatic and became uncontrollable in about 1.5s. Collective response was rather fast in comparison with the responses to total power failure in the Super Puma. The reason probably lies in the strong alerting cues provided by high yaw and roll rates, which would be forcefully represented by the simulator visual system, and possibly also adequately signalled by the motion platform as well, even though substantial linear accelerations were not evident.

3.4 Tail rotor drive failure (Super Puma only)

3.4.1 *Initial conditions*

Tail rotor drive failures were initiated during level flight on finals at 2000 ft. The failure was unexpected.

3.4.2 *Cues*

Table 11 shows the main visual cues available during the emergency with no pilot input. Auditory cues were not evident.

Table 11 Time line for visual cues

Time (s)	*Event*
0.0	Aircraft yaws to the left (about 80° s-1)
1.6	Aircraft starts to roll left
2.0	Airspeed starts to decline from 120 kt
2.3	Aircraft at 45° roll; starts to pitch down
3.6	Aircraft at 60-70° pitch down

3.4.3 *Reaction times*

Collective position was recorded. Again there was some unavoidable censoring of response time data because some response times could not be confidently determined. As a result, there are obviously fewer response and total reaction time data points than detection time data points.

Table 12 Summary of collective reaction data

Time	*No. of pilots*	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Detection	35	0.58	3.21	1.53	2.94
Response	14	1.11	6.52	2.74	6.06
Total reaction	14	3.26	8.26	4.90	7.79

3.4.4 *Discussion*

The results present an interesting comparison with the tail rotor control failure case. The situation does not apparently deteriorate quite so rapidly. Collective response was also slightly slower. The maximum and 90th percentile are notably larger than for any of the other emergencies, although the mean is similar. There is no ready explanation for this fact.

3.5 **Automatic flight control system failures**

3.5.1 *Initial conditions*

Two failures were presented, both resulting in an uncommanded pitch up. The first occurred in the hover just after take off. The second occurred in the cruise at 3000ft when the pilot's attention was probably not so closely focused on the flight instruments. The simulator was configured for maximum all up weight and maximum aft centre of gravity. Full cloud cover at 800ft was simulated.

3.5.2 *Cues*

AFCS runaways were accompanied by no significant motion cues or changes in audio cues. Visual cues, in the form of a change in pitch attitude detectable by reference to the attitude indicator or the horizon displayed on the dusk/night visual system, were the primary indication of the failure. Pitch rates typically achieved $5.55^{\circ}.s^{-1}$ within 0.53s of the failure onset.

3.5.3 *Reaction times*

Cyclic position was recorded. Only detection time – the time between failure onset and the first indication of an appropriate cyclic response – could be measured with confidence. Table 13 summarises these results. Response times and total reaction times were not evaluated. However, the interval between the first cyclic response and the reversal of the pitch up was measured as an indication of effective total time to control the failure. Table 14 presents these results; in two cases in the hover the result was a crash due to the failure occurring at too low a start height.

Table 13 Summary of detection time data

Condition	*No. of pilots*	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Hover	16	0.21	0.72	0.40	0.69
Cruise	16	0.60	4.65	1.59	2.40

Table 14 Summary of effective total reaction time data

Condition	*No. of pilots*	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Hover	14	1.17	2.40	1.62	1.83
Cruise	16	1.71	3.12	2.24	2.61

3.5.4 *Discussion*

Shorter detection times should be expected in a continuous control task, and the results from the failures in the hover conform to that expectation. Failures in the cruise resulted in significantly longer detection times, even though these failures occurred after those in the hover, a fact which may reasonably be supposed to have alerted at least some pilots to the purpose and form of the trial. The cruise naturally demanded less continuous monitoring of flight variables than the hover, and this general reduction in attentiveness seems to have had a more noticeable effect on detection time than the artificial distraction introduced for some total power failures in the Super Puma.

In generalising from these results, two things should be borne in mind. The mean detection time in the hover was close to typical values obtained in simple reaction time tasks. Similar values should be expected whenever the pilot is closely attending to a continuous control task, and the failure provokes a disturbance in that task. Second, in some aircraft the effect of an AFCS runaway could be more dramatic than that observed in the S61N. In such cases detection times may well be shorter than those observed here in the cruise. There is little scope for improvement when the pilot is closely attending to the task. Whether the improvement in detection times caused by more rapid excursions would offset the increased peril implied by the increase in rapidity is an open question, the answer depending on individual aircraft characteristics.

4 OVERALL DISCUSSION

Limitations of the data, particularly the unavoidable censoring of longer response times (mainly in the tail rotor failure data) should be borne in mind when drawing conclusions. Nevertheless, the results do indicate the broad range of performance to be expected

Detection times for cyclic responses to AFCS runaways in the hover were smaller than those for collective responses to total power failures and tail rotor failures. Clearly, a disturbance in a task at the focus of the pilot's attention is likely to provoke a minimal detection time. AFCS runaways in the cruise and all the other failures represent a more general case, where the pilot's attention is not so focused on the critical element. The cyclic and collective response data for these cases indicate that mean detection times of 1 to 2.5s can reasonably be expected.

More variability is evident in the collective response time data and, hence, in the total reaction time data. Depending, probably, on the urgency of the emergency, mean total reaction times generally lie in the range 2 – 4s.

Although these mean values represent what can reasonably be expected, design criteria should take account of normal, below average performance. The 90th percentile has been adopted here as an arguably reasonable criterion. Adopting it would imply accepting a 10% probability of failure to recover from the emergency. Whether or not a 10% failure rate is acceptable depends most importantly on the probability of a mechanical failure occurring – a topic beyond the scope of this paper. The 90th percentile for total reaction time for collective inputs appears to be about 4.5 – 5.5s, the tail rotor drive failure results being the only notable exception (7.8s). The 90th percentile for detection time lies in the range 2-3s, assuming no distraction; the tail rotor control failure results provide the exception in this case (1.2s).

There may be some scope for reducing reaction times by introducing warning systems, particularly if they give advance warning or enhance existing cues, as others have suggested (ref. 5).

The effects of distraction were explored in a minor way by this study. Despite the lack of a statistically significant result in the Super Puma trial, there are clear indications from the AFCS trial that changes in attentional focus can have an impact on detection times. The suspicion must remain that in normal operations the scope for important effects due to distraction and shifts in attention must be quite large. Further investigation of this issue would require an initial exploration of the types of distraction commonly to be expected.

The type of emergency and the aircraft type were both found to have an influence on reaction times. Both influences can be interpreted to a degree in terms of the urgency of remedial action required. This is clearly an economical and face valid explanation. Other interpretations are possible; helicopter type must be confounded with other variables like experience, type of operations, training regime etc.. These issues could be addressed by further study on the same pattern. To be effective, such a programme would need to be comprehensive in the range of aircraft types, operators etc. that it covered, and collect a large volume of data. It would also be advisable to use specialised equipment for the data collection to allow a comprehensive range of variables to be sampled.

The present results have implications for helicopter certification requirements. The relationship with current requirements is complicated by the lack of a common, and operationally defined terminology. British and American civil requirements refer to 'recognition time', and Ministry of Defence standards refer to 'decision time'. Although recognition of the failure and decision on remedial action are necessary phases in the pilot's reaction, these processes are not available for inspection. Without an objective definition in terms of a measurable feature of behaviour, such terms have no utility. In the hope of introducing some clarity, 'detection time' has been defined as the interval between the onset of the failure and the initiation of the control response. Although this interval arguably accommodates several processes, it has two virtues: it is objectively defined, and it represents the minimum delay before remedial action starts isolated from any delays caused by technical features of individual control systems. The latter fall within the interval defined in this report as 'response time' which also, in principle, has objective start and end points even though they may sometimes be difficult to discern in individual records.

Although the distinction between 'active and attentive' and 'passive' has obvious validity, objective measurement of such states is problematic. The present experiments drew attention to the difficulty of introducing a meaningful distraction in a consistent manner. Only the distinction between 'hands on' and 'hands off' seems viable.

It may be possible to simplify the regulatory position by adopting the objective definitions of delays used here, and by addressing two broad classes of failure. Accepting, for example, a 10% failure rate (i.e. 90th percentile performance) would result in the following expectations (all assume the pilot is 'hands on'):

(a) When the failure is immediately detectable because the deviating variable is within the pilot's primary focus of attention, expect detection times of less than 1s and total reaction times of less than 2s (e.g. AFCS runaways in the hover).

(b) When the failure is less immediately obvious and may become evident through secondary effects, expect detection times of less than 3s and total reaction times of about 6s.

5 CONCLUSIONS

- If a failure to recover rate of 10% is acceptable, a safe regulatory approach would allow up to 1s detection time and 2s total reaction time for failures likely to be immediately evident in a control variable at the focus of the pilot's attention (e.g. AFCS runaways in the hover). Up to 3s detection time should be allowed for failures likely to be less immediately obvious (e.g. total power failure and tail rotor drive failure in the aircraft represented in these trials). Total reaction times of about 6s should be expected, assuming no distraction. The broad similarity of results for collective response across aircraft types and emergencies suggests these results may be treated with some confidence.

- Current regulations should be reviewed in the light of the present findings and the need for operationally defined components of reaction time.

6 ACKNOWLEDGEMENTS

The authors would like to thank Brintel Helicopters Limited, Farnborough; Brintel Helicopters, Aberdeen; Helikopter Service, Stavanger; and the Chinook pilots from RAF Odiham for their co-operation and time. We are also grateful to Tony Hazell for his analysis of the auditory cues.

7 REFERENCES

1 Boff, K. R. and Lincoln, J. E. (eds). Engineering Data Compendium Human Perception and Performance Vol III. Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, 1988.

2 Green, R. G., Chappelow, J. W., and Skinner, R. J.. Response Times to Emergencies. In IAM Letter Report No 25/91, Pilot Intervention Times Preliminary Report.

3 Rabbitt, P. M. A. Times for the analysis of stimuli and for the selection of responses. British Medical Bulletin 27:259-265, 1971.

4 Smith, P. R. and Chappelow J. W., Pilot Intervention Times: Interim Report. RAF Institute of Aviation Medicine Report No 766, February 1994.

5 D. R. Haddon and T. Hughes, Enhanced warning and intervention strategies for the protection of rotor speed following power failure. CAA Paper 95009, October 1995

List of abbreviations

AFCS Automatic flight control system

Annex A Statistical Analysis

The total power failure data were log transformed to meet the assumptions of analysis of variance (ANOVA). Under these transformations a test of Studentised residuals showed no significant outliers in the whole data set. The transformed data also showed no deviations from a Normal distribution in terms of skewness and kurtosis.

An initial analysis on the Chinook and first S61N total power failure data sets examined the effects of height, airspeed, manoeuvre, and aircraft type on detection, response and total reaction times. No significant effects were found. There being no particular reason *a priori* to expect major effects from variables other than helicopter type, the other variables were omitted from later analyses.

HELICOPTER TYPE

ANOVAs were conducted for detection, response and total reaction times for total power failure examining the factor helicopter (H). There were three levels of H: Chinook, S61N and Super Puma (undistracted cases only).

Detection time

There was a significant effect due to helicopter type: $F = 19.09$, $df = 2,73$, $p < 0.001$. The Chinook times were lower than those for the S61N and the Super Puma ($p < 0.001$), and the S61N times were lower than those for the Super Puma ($p < 0.05$).

Table A1 Estimated means for detection time (s)

Chinook	*S61N*	*Super Puma*
1.07	1.67	2.22

Response time

There was no significant difference between the helicopter types: $F = 2.97$, $df = 2,58$, $p < 0.06$.

Table A2 Estimated means for response time (s)

Chinook	*S61N*	*Super Puma*
1.24	1.61	1.82

Total reaction time

There was a significant difference due to helicopter type: $F = 10.45$, $df = 2,58$, $p < 0.001$. The Chinook time was lower than that for the S61N ($p < 0.01$) and the Super Puma ($p < 0.001$).

Table A3 Estimated means for total reaction time (s)

Chinook	*S61N*	*Super Puma*
2.30	3.33	4.13

DISTRACTION (Super Puma total power failure only)

ANOVAs were conducted for detection, response and total reaction times for the factor distraction. Minima, maxima, and 90th percentiles are reported here as well as the estimated means for completeness. (Note that the means for the undistracted case from this analysis are inevitably slightly different to those in the preceding analysis due to the bias correction factor employed in the back transformation.)

Detection time

There was no significant difference due to distraction: $F = 1.29$, $df = 1,40$, $p > 0.05$.

Table A4 Super Puma detection time

	No of pilots	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Distracted	13	1.22	5.90	2.31	4.82
Undistracted	29	1.02	3.14	2.01	2.95

Response time

There was no significant difference due to distraction: $F = 0.03$, $df = 1,20$, $p > 0.05$.

Table A5 Super Puma response time

	No of pilots	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Distracted	8	0.44	2.81	1.65	2.77
Undistracted	14	1.11	2.44	1.61	1.95

Total reaction time

There was no significant difference due to distraction: $F = 0.99$, $df = 1,20$, $p > 0.05$.

Table A6 Super Puma total reaction time (s)

	No of pilots	*Min. (s)*	*Max. (s)*	*Mean (s)*	*90th percentile*
Distracted	8	3.09	5.26	4.16	5.10
Undistracted	14	2.74	5.08	3.82	4.73

OUTCOME (Chinook total power failure only)

ANOVAs were conducted for detection, response and total reaction times for the factor outcome (O). There were two levels of O: crashed and successful recovery. First and subsequent exposures were included in this analysis.

Detection time

There were no significant effects.

Response time

There was a significant effect of outcome: $F = 14.35$, $df = 1,28$, $p < 0.001$.

Table A7 Estimated means for response time (s)

Crashed	*Recovered*
1.71	0.90

Total reaction time

There was a significant effect of outcome: $F = 6.42$, $df = 1,28$, $p < 0.05$.

Table A8 Estimated means for total reaction time (s)

Crashed	*Recovered*
3.21	1.95

AFCS RUNAWAYS (S61N only)

Detection time

There was no evidence of differences between individual subjects and subject was used as a random effect in the analysis of variance. There was a significant effect of phase of flight (hover/cruise): $F = 67.65$, $df = 1,15$, $p < 0.001$. Detection time was higher in the cruise condition ($p < 0.001$).

Table A9 Estimated means for detection time.

Hover	*Cruise*
0.40	1.59