DOT/FAA/AM-97/24

Office of Aviation Medicine
Washington, D.C. 20591

Automation in General Aviation: Two Studies of Pilot Responses to Autopilot Malfunctions

Dennis B. Beringer
Howard C. Harris, Jr.

Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma 73125

December 1997

Final Report

U.S. Department
of Transportation

Federal Aviation
Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-97/24		

4. Title and Subtitle	5. Report Date
Automation in General Aviation: Two Studies of Pilot Responses to Autopilot Malfunctions	December 1997
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Beringer, D.B., and Harris, H.C., Jr.	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
FAA Civil Aeromedical Institute P.O. Box 25082, Oklahoma City, OK 73125	11. Contract or Grant No.
12. Sponsoring Agency name and Address	13. Type of Report and Period Covered
Office of Aviation Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

16. Abstract

Study 1 examined four automation-related malfunctions (runaway pitch-trim up, roll servo failure, roll sensor failure, pitch drift up) and subsequent pilot responses. Study 2 examined four additional malfunctions; two more immediately obvious (runaway pitch-trim down, runaway roll servo) and two subtler (failed attitude indicator, pitch sensor drift down) than those in Study 1, and the effect of an auditory warning. Data collection was performed in the Civil Aeromedical Institute's Advanced General Aviation Research Simulator, configured as a Piper Malibu. Results suggest that maladaptive responses to some of these failures may, in a significant percentage of cases, lead to significant altitude loss, overstress of the airframe, disorientation of the pilot, or destruction of the aircraft. Percentages of successful recoveries, detection/correction times, and related indices of performance are discussed in the context of malfunction type, flight profile, and auditory alerts.

17. Key Words	18. Distribution Statement
Autopilot, Automation, Simulator Research, General Aviation, Instrument Flight, Flight Simulation, Psychology, Applied Psychology	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	27	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Automation in General Aviation: Two Studies of Pilot Responses to Autopilot Malfunctions

The autopilot is generally recognized as a useful tool in reducing pilot workload, particularly during single pilot instrument flight rule operations (Hoh, Smith, and Hinton, 1987). However, one does not have to search very deeply into popular press aviation publications to find accounts of actual or perceived problems associated with autopilot or flight management systems. The most visible and recollected ones are those which resulted in the loss of large commercial aircraft. One such example is the loss of China Airlines' Flight 140, April 26, 1994, on approach to Nagoya/Komaki airport, Nagoya, Japan (Katz, 1995). The flight-recorder data indicated that the aircraft, an Airbus A-300-600R, ultimately stalled and crashed after attaining a pitch-up attitude of approximately 52 degrees at 78 knots. The problem appeared to be the pilot's continued attempts to fly the airplane manually with the autopilot engaged in go-around mode. The captain, who had apparently inherited the approach from the first officer after an autothrottle but not autopilot disengagement, ultimately lost the struggle with the aircraft as the autopilot trimmed the aircraft nose up in response to the captain's continued attempts to force the nose down. Concerns in these Part 121 (Air Carrier) operations have received attention (Funk, Lyall, and Riley, 1993), and many of the problem areas (mode confusions, control authority issues, etc.) are common to both Part 23 (Normal, Utility, Acrobatic, and Commuter Airplanes) and Part 25 (Transport Category) aircraft. A number of recommendations have already been made for Part 121 operations, including those for design/certification and pilot training (FAA, 1996).

Problem

General aviation aircraft, however, far outnumber commercial air carrier aircraft in the United States and also appear to be a source of unfriendly encounters between pilots and autopilots. Wilson (1995) reports a personal encounter of a similar nature to that described for China Airlines but experienced in a Beech Queen Air. The autopilot had been engaged and appeared to be functioning properly. The pilot and passenger then engaged in conversation and, sometime thereafter, the aircraft pitched nose down. The pilot applied backpressure on the yoke with a resulting increase in the pitch-down tendency. The passenger/copilot also applied backpressure "to no avail." With airspeed and pitch down increasing, the pilot detected the motion of the trim wheel running to nose-down trim. Their first attempt to correct was to "turn off" the autopilot. When this failed to correct the trim problem, they "unplugged the monster." It is unclear from Wilson's narrative whether the latter two actions refer to use of the circuit breakers, but this would appear to be the intent. The pilot, in retrospect, reported limited experience with autopilots at the time and stated, "As we taxied out and went through the runup, things were fine. I ignored the autopilot as always."

These are not one or two isolated incidents. Katz (1995) reported that a National Transportation Safety Board (NTSB) examination involving one manufacturer's aircraft found 17 autopilot-related accidents and incidents between 1983 and the publication date. In the 7.5 years ending in June 1994, the FAA received 175 service difficulty reports on autopilots installed in the same make of aircraft. If this is representative of what one would find when examining other makes of aircraft, then the total is likely many times this number. One must also consider incidents that result in momentary loss of control of the airplane but are then corrected without adverse effect to aircraft or crew. The majority of these are likely unreported if the data we obtained from the pilots pilots participating in our experiments are representative indicators. Of these incidents and accidents, Katz notes:

> Many of these accidents could have been prevented if the autopilot system had been used correctly, if the

pilot had disconnected the system instead of trying to troubleshoot a problem or if the pilot hadn't assumed that a problem was temporary and later attempted to use the autopilot.

The NTSB notes that if an autopilot malfunctions or an airplane is improperly operated with the autopilot engaged, significant deviations from the flightpath, mistrimming of the aircraft or the need for excessive control forces may occur. These problems may result from a runaway electric trim or pilot attempts to oppose or overpower the autopilot pitch axis. In most situations when a pilot attempts to overpower the pitch axis for more than several seconds, the autopilot trim servo will move the elevator trim tab in a direction that will countermand the pilot's input. If the pilot continues to restrain the control yoke and the autopilot/electric trim doesn't automatically disconnect, the trim tab will continue to operate and yoke forces may become overwhelming. The Safety Board also believes that many pilots don't bother conducting preflight checks of autopilot for proper operation.

These opinions were further underscored by two accidents where pitch trim was implicated. In the first, a twin-engined aircraft crashed near Flagstaff, Arizona, during a circling VOR/DME approach. Although nothing was found to be wrong with the flight controls or engines, the elevator trim was found in the full nose-down position and it was determined that the trim annunciator light had been illuminated at the time of impact. In the second accident, a Bonanza pilot reported to ATC that he was unable to turn off the autopilot and was struggling with the aircraft. The pilot received final vectors to Chapel Hill, North Carolina, 45 minutes later and crashed on the approach. Examination of the aircraft showed the elevator trim to be in the full nose-down position, requiring approximately 45 pounds of force to hold level flight. It appears likely that the autopilot had indeed been disconnected or powered down, but that the out-of-trim condition was either not detected or a runaway trim servo, driving the trim tab to full deflection, was never disabled or even diagnosed.

Contributing Factors

A number of factors are likely to contribute to the chain of events ultimately leading to an autopilot-related accident. These may include, but are not limited to: insufficient pilot training, pilot lack of an underlying model of autopilot behavior, misdiagnosis of malfunction, organizational policies, pragmatic considerations, human performance limitations, and system designs that do not capitalize on human factors principles.

Insufficient training. There is presently no regulation stating that a pilot must receive training in the use of an autopilot before flying with one in an aircraft. Although such training is the rule in Part 121 operations for flight management systems, General Aviation is yet another story. Theoretically, one could fly any aircraft that one was checked out in, and if a model of that aircraft happened to have an autopilot, the pilot would be free to use it without specific instruction. The same is true for GPS and other systems that one could conceivably add to the aircraft. The tempering factors, one would expect, would be that a prudent pilot generally would learn everything possible about the airplane to be flown, particularly if it were owned or regularly flown by that pilot. Additionally, if the aircraft were leased, it would be expected that all potential lessees would be thoroughly checked out in aircraft systems operations prior to being allowed to lease the aircraft, usually for insurance purposes. This is often not the case, however.

Lacking conceptual model. It is also possible that pilots lack an underlying conceptual model of how the various components of the autopilot/autotrim system work in concert or in opposition. It has been argued that the ability to diagnose novel malfunctions (those not specifically encountered before) of a system is directly related to the availability of such a mental model of the system. In the case of general aviation, it is likely that many pilots will not have experienced autopilot failures prior to their first need to respond to one as pilot in command. Thus, the need to have a working knowledge of system structure and functional relationships is important to prevent the first encounter from being the last.

Misdiagnosis. The lack of an adequate conceptual model of the autopilot/autonav systems may also, as pointed out in the Chapel Hill accident example, result in a misdiagnosis of the malfunction, leading the pilot to nonproductive actions that may further aggravate the flight control problem.

Organizational policies / pragmatic concerns. The way in which the pilot responds to malfunctions may also be dictated by organizational policy, particularly if the organization is responsible for its own *ab initio* or continuing flight training. Some organizations prefer that pilots "work with" the autopilot rather than immediately disconnecting it in cases where a malfunction is apparently mild and does not pose an immediate and obvious threat to safe flight. There is also a pragmatic consideration when the pilot is also the aircraft owner. If a service technician is to be called upon to remedy an apparent autopilot malfunction following the termination of the flight, additional data on the aberrant behavior will be helpful in localizing the problem, potentially reducing the time required for diagnostics by the technician and, thus, cost.

Human performance limitations. Both perceptual and motor human performance limitations are likely to affect how a pilot responds to autopilot malfunctions. Detection of malfunctions is decidedly influenced by limitations in visual and aural perception, specifically where a stimulus to be detected is not in or near the line of sight or where the stimulus is not above threshold or is steady state. It has been noted that some auditory alarms go unnoticed by pilots who have high-frequency hearing loss due to a combination of aging and work-place exposure to high-amplitude narrow-band sounds.

Human factors and design issues. It is sometimes the case that installed systems simply do not conform to the standard human factors practices and principles. The instrument panel is a land of finite space, and not everything can be between zero and fifteen degrees below line of sight and located on the centerline of normal vision. This often results in systems that may be added on or optional equipment being located at the bottom of the radio stack or in the most convenient panel location available. If the unit contains displays that require frequent monitoring for continued safe operation, placement may make this impossible. It is also possible that warnings, be they visual or aural, may not conform to standards. One usual departure is the use of steady-state visual and aural warnings rather than alternating on/off/on warnings, which are more likely to attract the attention of the pilot.

Certification Standards

Present certification standards require that an autopilot system, in a hard-over failure where the control surface servo is driven at its maximum rate, cannot place the aircraft in greater than a 60-degree bank nor place undue loads (0 - 2 G's limits) on the airframe "within a reasonable period of time" (FAR 23.1329). This has been operationalized (DOT/FAA Advisory Circular 23.1329-2, 1991) as within the three seconds following the initial detection of the uncommanded bank. Similarly, this applies to pitch and pitch trim tests to the degree that the aircraft cannot stall, exceed limit speeds, or require excessive control force during recovery at the end of the three-second period. This supposedly provides three seconds in which the pilot can diagnose the problem and take corrective action (autopilot disconnect is assumed). A delay of one second was adopted for malfunctions on a coupled approach, on the theory that the pilot is likely to be attending the instruments more closely on approach than during cruise. Cooling and Herbers (1983) noted, in their discussion of human factors, that "...there are no studies available to support the FAA certification standard of a three second delay (enroute) or a one second delay (on approach) before initiation of recovery by the pilot from an autopilot malfunction." However, it has been suggested that the data were actually derived from an examination of airline pilots' responses collected during a study performed at Wright-Patterson AFB in the 1960s (ACE-110, 1996).

The focus of our research, in support of Aircraft Certification, was the responses of pilots to overt and subtle autopilot malfunctions and the factors influencing the speed and the selection of those pilot responses. Two studies were conducted, each examining four autopilot or autopilot-influencing system malfunctions, including those producing obvious and immediate effects and those producing more subtle and less direct effects. The intent was to determine how a sample representative of average General-Aviation pilots would respond to autopilot malfunctions and how those responses would compare with the times specified in the present certification procedures.

GENERAL METHOD

The same method was used in both studies with the exception that different autopilot malfunctions were substituted in Study 2. Thus, the following descriptions are applicable to both studies up to the actual characterization of the specific pilot sample and a few minor variations in the independent variables.

Design/Subjects

The experimental approach, a single-factor within-subject design using autopilot malfunction type (4) as the independent variable, was selected because high between-subject variability in response times to malfunctions was expected. Study 1 malfunction types were: "command over" roll (rate = 6 deg/sec), soft roll (sensor) (rate = 1 deg/sec), soft pitch (sensor) (rate = 0.2 deg/sec), and runaway pitch trim up. The last was selected for practical reasons to increase the likelihood of completing data collection. If not attended to, runaway pitch-trim down can create significant pitch-down attitudes and possible over-speed conditions, increasing the potential for a prematurely terminated or interrupted data run. Dependent variables recorded included flight performance indices (6 degree-of-freedom data plus airspeed, etc.), and states of critical switches with event/change times; autopilot disconnect, engage, pitch-trim and circuit breaker. Pilots were obtained from the local area who were instrument rated and had experience with complex aircraft and autopilot systems. These individuals were largely from the Oklahoma Pilots' Association, were contacted directly by the experimenters, and were compensated for their time. Ages ranged from 24 to 72 years (median = 42) and the sample contained 27 men and 2 women. No subject had less than 300 hours of flight experience.

Equipment/Procedures/Tasks

Data-collection sessions were conducted in the Advanced General Aviation Research Simulator (AGARS) (Appendix, Figure A1) in the Human Factors Research Laboratory, Civil Aeromedical Institute. This fixed-base simulator was configured as a Piper Malibu with Bendix/King avionics (KFC-150 autopilot); software approximated behavior of both, but exact flight equations were not available. High-fidelity primary flight displays were presented in the cockpit on three masked CRTs that replicated the Malibu panel layout and gave the appearance of electromechanical instrumentation. The out-the-window depiction spanned 150 degrees of horizontal visual arc and was a high-resolution textured representation of the Oklahoma City area.

The fixed-base nature of the simulator suggested that some unique circumstances might produce outcomes not generalizable to the aircraft environment. Specifically, responses to overt failures (i.e., roll servo), for pilots neither holding the yoke nor viewing the external scene, might be shortened by vestibular cues. Responses to subtle failures, as during the initial stages of runaway pitch trim where the pitch servo still has sufficient authority to counteract trim, are not likely to benefit. It was also anticipated that the relatively compelling visuo-vestibular effect of the highly-textured 150-degree external visual scene would be sufficient to detect when pilots were "heads up," particularly during roll perturbations.

Pilots participated in one 2- to 2.5-hour session. They were told that the study was to examine use of autopilots in routine flying and to gather opinion data on useful features. The first hour consisted of experiment-related paperwork and familiarization-training activities, including: reading excerpts from the autopilot (AP) manual, cockpit familiarization, and a half-hour familiarization flight using all AP modes. The second half of the session was used to collect performance data for the malfunction conditions. A simple round-robin instrument clearance was flown from Will Rogers World Airport to two local very-high-frequency ominrange (VOR) stations, and back, in Instrument Flight Rules (IFR) conditions between textured cloud layers (distinct visual horizon but no ground detail). Pilots were required to interact with Air Traffic Control, fly vectors, track inbound to two VOR stations, and fly a fully-coupled instrument-landing-system (ILS) approach, and were instructed to fly as much of the course as possible with the AP engaged. An additional task, as always, was to conduct visual surveillance of the surrounding airspace for traffic, and to this end, two intercepts with a Piper Navajo were constructed. One had the Navajo passing across the Malibu's nose at less than one mile and 1000 feet above, while the other had the Navajo passing, co-altitude, from right to left across the visual field at from approximately 13 miles to about eight miles distance.

Malfunctions were spaced such that sufficient time elapsed between failures (13-15 minutes) to prevent interference between episodes. Command roll and soft pitch were encountered in level flight, soft roll during descent, and half pitch trim during the ILS approach and half during ascent from 6000' to 7000' (see Figure A2 in the Appendix for experimental route and placement of malfunctions). Only the pitch-trim malfunction produced both auditory and visual warnings, consisting of a steady TRIM light and steady pure tone of 3.1 kHz at approximately 77 dB. The simulated system did not immediately disconnect during the runaway, representing a worst-case scenario (the KFC-150 AP does automatically disconnect, although some others do not), allowing the pitch servo to compensate for (and mask) the initial trim deflection. Data collection flights averaged 1.2 hours, followed by an AP-experience questionnaire and interview to determine each pilot's knowledge of AP and autotrim malfunction consequences and to gather task difficulty ratings.

STUDY 1 RESULTS

Response Times

Command roll (roll servo). Of all the failures, commanded-roll and pitch-trim failures were rated as easiest to diagnose (by 11 of 26 pilots). The commanded-roll failure emulated an AP-commanded roll that exceeded the target bank angle. Analyses for both roll malfunctions and the soft-pitch malfunction are based upon time from initial failure to disconnect of the AP by any means (yoke-mounted disconnect, panel disengage, circuit breaker). Times ranged from 1.8 seconds to 107.1 (means, medians, and ranges are summarized in Table 1). However, 69 % of the pilots disconnected within 13 seconds of the initial failure and half within 8 seconds. These "immediate" disconnects by 18 of the 29 pilots were defined by sequences where no other significant actions occurred between failure onset and AP disconnect. The distribution of these times is shown in Figure 1A. Using a response time of 8.7 seconds or less as a cutoff value, 93.7% of the sample of "immediate" responders were included.

Table 1. Study 1 response time mean, median, and range by failure and response category types.

Failure Type	Response Category	n	Response Time		Range	
			Mean	Med	Low	High
Command Roll	All (Disc)	29	16.5	8.5	1.8	107.1
	Immediate	18	5.9	5.9	1.7	11.8
	Manual Override	10	26.3	23.0	8.9	53.8
Soft Roll	Immediate	16	11.7	11.5	4.5	21.2
	Manual Override	13	37.5	26.0	13.2	85.1
Soft Pitch	Immediate	12	17.7	17.4	6.5	31.5
	Manual Override - 1	16	46.2	50.0	15.2	76.2
Pitch Trim Up	All (Disc)	25	10.5	6.9	0.2	39.2
	All (CB pull)	25	35.4	23.5	4.9	109.7
	All (CB lag)	25	25.0	15.7	0	102.3
	All (minus extremes)	23	22.7	15.7	5.1	71.3

Ten pilots chose to manually override the AP, whether by using the control-wheel steering option or by overpowering the roll servo without disconnecting the AP. Ninety percent had response times of 48.3 seconds or less (Figure 1B). Scores were log-transformed for post-hoc analyses to remove the usual skewness found in response-time data. Comparison of these log-transformed disconnect times for the two groups, with the highest and lowest extreme times removed, indicated a significant difference ($F[1,24] = 53.27, p<0.0001$) between the immediate disconnects (untransformed mean = 5.93 seconds) and the manual overrides (untransformed mean = 28.26 seconds).

Soft roll (roll sensor). The soft-roll failure was rated as third in difficulty to diagnose, but was rated easiest to correct (by 13 of 26 pilots). Following removal of one outlier (194 seconds), pilot performance was again categorized as immediate disconnect (16) or manual override (12). Those categorized as immediate disconnect responses averaged 11.72 seconds (range: 4.52 to 16.69) (Figure 2A), while those categorized as manual overrides averaged 37.45 seconds (range 13.16 to 85.14) (Figure 2B). Approximately 88% of all immediate disconnects occurred in less than 17 seconds, with 75% occurring in less than 14 seconds. Post-hoc comparison indicated the mean difference to be significant for both raw and log-transformed scores (log scores: $F[1,26] = 27.07, p<.00005$).

Soft pitch (pitch sensor). The soft-pitch failure was rated as most difficult to diagnose (by 12 of 26 pilots) and was rated third easiest to correct, missing a tie for second by one tally. Performances were again categorized as either immediate disconnect (12) or manual override (17), and the distributions are shown in Figures 3A and 3B. Three pilots never diagnosed the failures, manually flying the airplane without disconnecting the autopilot; their scores and one other outlier were removed, leaving 13. Immediate disconnects (Figure 3A) averaged 17.7 seconds (range: 6.5 to 31.5) and manual overrides (Figure 3B) averaged 46.19 (range: 15.2 to 76.2). Approximately 50% of immediate disconnects occurred in less than 16 seconds, with approximately 85% occurring in less than 24 seconds. Post-hoc comparison of the log-transformed data showed the distributions of the two types of responses to be significantly different ($F[1,22] = 20.69, p<.0005$).

Runaway pitch trim. This failure was different from the others in that only by pulling the pitch-trim circuit breaker would the problem be corrected. The interim solution was the AP disconnect/trim interrupt switch. Only three pilots chose the optimal response, depressing and holding the disconnect, then pulling the circuit breaker. Four others depressed and held the disconnect at various times during the recovery. The vast majority of initial responses were yoke AP disconnect (15), followed in frequency by panel-mounted

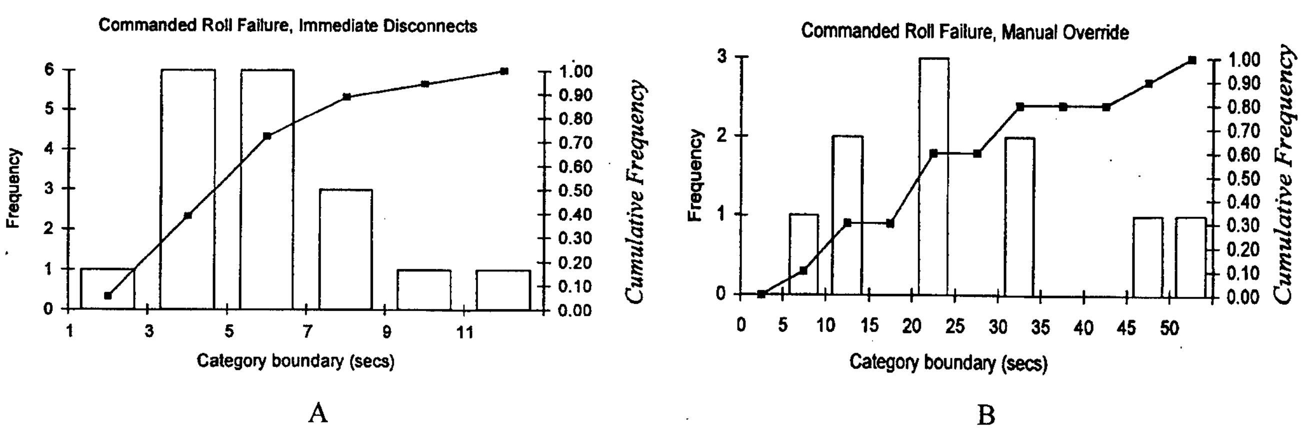


Figure 1. Commanded-roll response-time distribution and cumulative frequency plots for (A) immediate disconnects and (B) manual overrides.

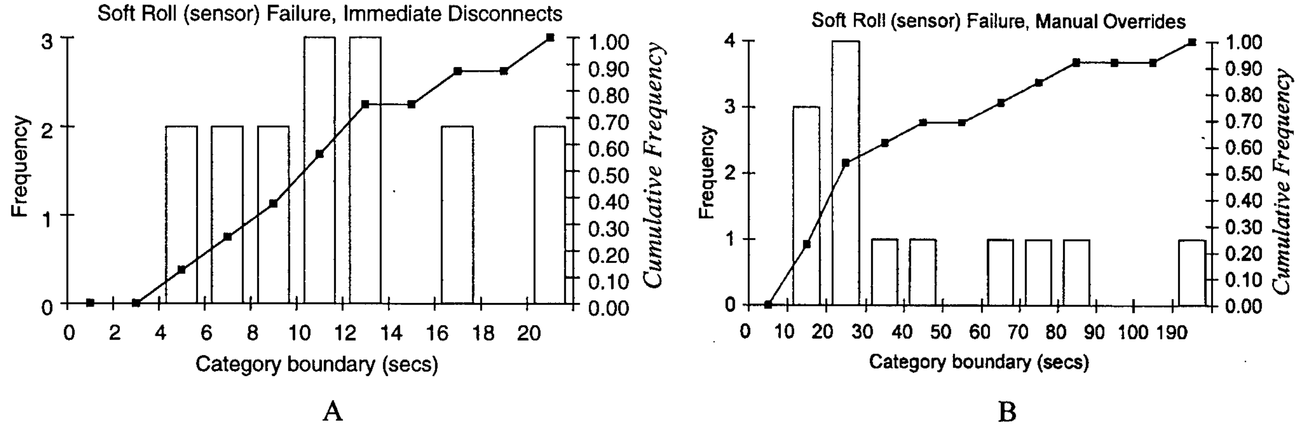


Figure 2. Soft-roll (sensor) response-time distributions and cumulative frequency plots for (A) immediate disconnects and (B) manual overrides.

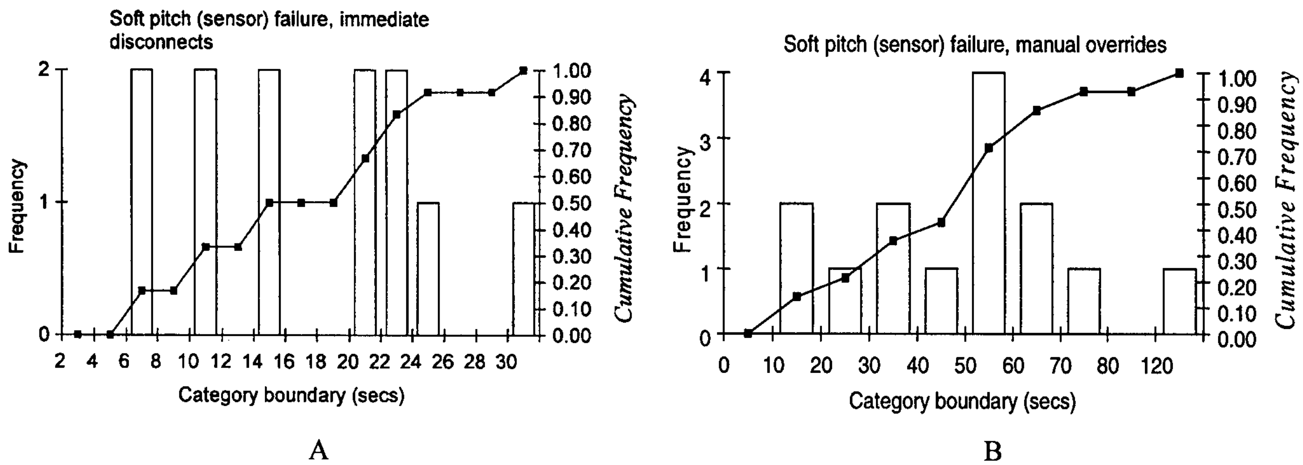


Figure 3. Soft-pitch (sensor) response-time distributions and cumulative frequency plots for (A) immediate disconnects and (B) manual overrides

AP-engage switch (5), mode manipulation (2), manual override (2), and pitch trim circuit breaker (1). Overall, 21 of the 25 pilots considered were classified as "immediate" responders, two were classified as manual overriders, and two as mode changers. It should also be noted that two pilots never heard the warning tone possibly due to high-frequency hearing loss, responding only to aircraft performance changes.

Two stages of response were of interest; first, the time required to detect a malfunction and initiate some action (AP disconnect, control-wheel steering, AP engage or circuit breaker) and second, the time lag between the initial action and the pulling of the pitch-trim circuit breaker. Average time to initial action for the usable 25 pilots was 10.46 seconds, with all except one response over 3 seconds. One can see from Figure 4A that 50% of the responses occurred in less than 7 seconds, with 65% of the cases in less than 9 seconds. Time to pull the pitch-trim circuit breaker averaged 35.4 seconds (range: 4.91 to 109.69) (Figure 4B), with an average lag of 22.69 seconds (high and low scores removed) between the initial response to the runaway pitch trim (disconnect or control movement) and the required remedy.

Initial examination of the *questionnaire and interview data* indicated that all pilots understood they could manually overpower the autopilot servos, and 22 were aware of the potential interaction between a

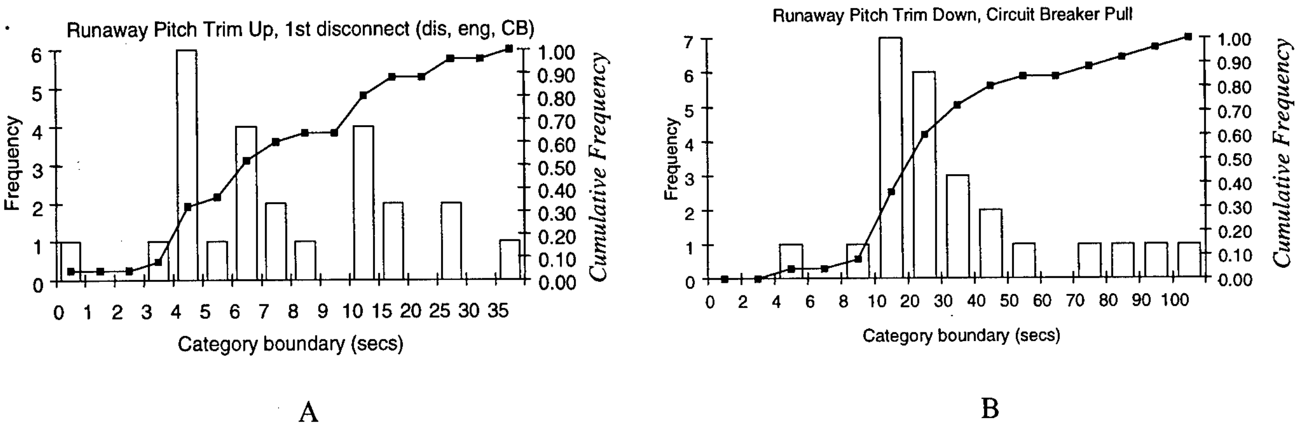


Figure 4. Runaway pitch-trim up response-time distributions and cumulative frequency plots for (A) first disconnect and (B) circuit breaker pull.

runaway pitch-trim motor and autopilot pitch-attitude (elevator servo) inputs. Four pilots had not considered the potential interaction previously but grasped the concept immediately during the interview. When asked what their strategy for dealing with autopilot malfunctions was, the group voiced two anchor strategies and a combination of the two as a third. The *immediate-disconnect* strategy was endorsed by nine individuals, while two others expressed a *procedural* approach that was closely related to the immediate disconnect strategy. Another five individuals suggested that they would *fly the aircraft through the malfunction* while attempting to diagnose the problem. A third group took a middle-of-the-road stance, saying that the strategy was malfunction dependent. These seven expressed their strategies as, "Fly through mild failures; disconnect for severe failures," or "diagnose while the unit is still engaged, then disconnect." Those individuals using a "fly-through" response for any part of the malfunction will be subsequently referred to as using a "manual-override" strategy.

Mode-of-flight effects. The mode of flight during which the failure is encountered is also of particular interest. Recall that the delay used during certification is to be *one* second during a coupled approach, as specifically delineated by the advisory circular, and that the experimental procedure was set up to examine the pitch-trim failure during both cruise climb and a coupled ILS approach. The aircraft is more likely to reach slow airspeeds in either of these conditions than when the failure is encountered in level cruise or cruise descent.

Independent-samples t tests indicated no significant mean difference between response times for these two flight modes. Levene's Test of variability, however, indicated a significant difference for the circuit-breaker lag ($F = 3.406, p<0.1$). The group experiencing the failure on climbout (SE of Mean = 7.37) was more variable in their responses than was the group receiving it on approach (SE of Mean = 5.07). When these scores were log transformed, as is usually advisable for response times, no significant effects of mean or variance differences were found. Although the first analysis could lend some credibility to the assumption that pilots were somewhat more attentive on approach, the lack of an effect for the log-transformed scores would tend to downplay this explanation. That the difference might represent an inherent difference between the post-hoc groups (climb, approach) was examined by performing comparable analyses of all other RT variables (commanded roll, soft pitch/sensor, soft roll/sensor). No significant mean or variance differences were found for either the raw or transformed scores, suggesting that these two groups of pilots were not significantly different in their performance on the experimental tasks.

Correlational Data

Point biserial correlations were calculated to examine the relationship between stated strategy, flight experience and response times (RTs). No systematic relationship was found between hours of flight experience and strategy use in the simulator. The expected relationships between RT and selected strategy were significant, as those pilots electing a manual-override strategy had, of necessity, longer overall RTs. Values for r ranged from -.69 to -.47 (negative due to response coding for analysis). A significant correlation was also found between occupation and roll sensor failure RT (rpb=.41), largely because 4 of 5 FAA pilots adopted a manual-override strategy for this failure and had longer RTs.

Pearson correlations were calculated relating RTs to time since last experienced autopilot failure, with significant ($p<.05$) values for soft pitch ($r=.48$), runaway trim ($r=.54$), and commanded roll ($r=.38$). Pilots who had recently experienced an autopilot failure were more likely to respond quickly than those who had not. Additionally, there were significant ($p<.01$) correlations between roll-sensor RTs and three training/experience measures: dual instruction received in the last 24 months ($r=.73$), simulated instrument time during the last 12 months ($r=.66$), and the number of hours of simulated instrument time in the last three months ($r=.62$). Interestingly, the group electing to use some form of manual-override strategy reported nearly twice as many hours in all three categories as were reported by the immediate-disconnect group. This arises from the fact that over half of the pilots in the manual-override group were required to fly in their occupation and to receive instruction as part of their continuing education.

Flight Performance Data

The Advisory Circular 23.1329-2 specifies that attitude and performance specification limits shall not be exceeded during recovery from excursions induced by an autopilot malfunction. Examination of pitch, bank, altitude, and indicated airspeed for each recovery indicated that only one individual exceeded 60 degrees of bank during one recovery, and for all other cases and all other malfunctions, the aircraft was in a flyable condition and did not exceed attitude or airspeed performance limitations. Thus, one can say that recoveries were timely enough to prevent the aircraft from assuming extreme attitudes or airspeeds (overspeed or stall).

STUDY 1 DISCUSSION

Present certification practice assumes that a malfunction will be either severe enough to produce supra-threshold cues or that an alert will warn the pilot, starting the three-second "recognition" period. Flight test personnel (FAA Aircraft Certification Service, 1996) have reported test malfunctions that have gone undetected until the test administrator or safety pilot pointed them out, sometimes after reaching criterion limits. These autopilots failed to obtain certification. Study 1 data indicated pilots required an average of 5.9 seconds to a clearly supra-threshold event, some requiring as long as 11.8 seconds. General certification practice for "obvious" malfunctions allows one second for detection. Combined with the three-second waiting period, this produces a four-second interval within which the pilot must detect and respond to the malfunction; less than the mean sample response. For the commanded-roll failure, one could accommodate 90% of this pilot sample using nine seconds as the interval upper bound. Using even seven seconds as the criterion, 70% of the sample would be accommodated. One should note that at the usual five deg/sec commanded roll rate, a 60-degree bank would not be exceeded for 12 seconds. A roll-servo hard failure at 15 deg/sec for this aircraft type, however, does so in four seconds.

It was not surprising that significantly longer intervals were required for pilot response to the more subtle failures. However, because the attitude indicator (ADI) continued to depict actual attitude during these malfunctions (in a true sensor failure, the ADI would not), detection times were probably shorter than would otherwise be expected. Given this ADI anomaly, the potential consequences of the pitch-trim down runaway, and the "moderate" roll rate in the commanded-roll failure, additional data collection was planned for runaway pitch trim down, as well as true attitude sensor failure and hard-over roll-servo failure (12-15 deg/sec roll rate) (see Study 2).

It is also worth noting the number of pilots who adopted the "wait-and-see" strategy. In these cases, the choice of strategy was a clear influence on the recovery time and, in some cases, on the "success" of the recovery. Recall the comments of Katz (1995) as an indictment of any procedure that does not use an immediate disconnect of the affected system as a definite threat to the safety of the pilot and the aircraft. Although no individuals actually placed the aircraft in a hazardous situation using the "fly-through" or "diagnose-then-disconnect" strategies in Study 1, these malfunctions were of types that were not likely to produce unrecoverable situations very quickly, specifically because the pitch trim failure was in the "up" direction. The failures in Study 2, however, are yet another matter and produced quite different results, to be detailed shortly.

One should also take note of the two pilots who reported having never heard the auditory warning. Although they represented a small proportion of the sample (6.9%), this finding does suggest that there are likely to be pilots who are at risk of a failure to perceive auditory cues due to the combined effects of high-frequency hearing loss, ambient noise, and the attenuating effects of headphones.

Initial recommendations that came out of Study 1 included:

- Increase the waiting period for "command-over" and "sensor-loss" failures to accommodate at least 75% of the general pilot population, using cumulative frequency curves on response time distributions.
- Consider eliminating separate treatment of approach and other flight modes given no detectable pilot response differences.
- Pursue additional failure annunciation or "fail-safe" modes from manufacturers.
- Continue use of attitude and performance limitations as ultimate criteria for acceptance.
- Examine the efficacy of cockpit auditory alarms and alerts when noise-attenuating headphones are in use.

It was recognized that the most hazardous malfunction, in terms of its ability to place the aircraft in a configuration from which it might be difficult to recover, was the runaway pitch-trim-down failure, described by Wilson (1995) and implicated in the Flagstaff and Chapel Hill accidents. Also included among the more hazardous "rapid-onset" failures was the runaway roll servo mentioned earlier, potentially producing a 15-degree/sec roll rate in this class of aircraft. Noting that only one of 29 participating pilots in Study 1 had to be "rescued" by freezing the simulator, the experimenters felt that the malfunctions presented were somewhat conservative in nature, compared with potentially more threatening system failures. On the opposite end of the continuum were the subtler failures, those having slow onset and progression rates or residing in systems upon which the autopilot depended for accurate data. Following some software revision to guarantee a secure continuation of the experimental session in the event that a pilot reached overspeed and/or failed to recover from a malfunction for any reason, Study 2 was initiated to explore the more hazardous and the more subtle malfunctions.

METHOD: STUDY 2 REVISIONS

Experimental Design

The basic experimental design was again a single-factor within-subject using autopilot malfunction type as the independent variable. The four malfunction types were selected to run the gamut from *largely covert* to *largely overt* in nature: runaway roll servo (roll rate = 12-15 deg/sec; *overt*), attitude indicator (ADI) failure (slow drift; autopilot tries to follow failed instrument; *covert*), soft pitch failure (rate = 0.2 deg/sec; *covert*), and runaway pitch-trim down (*initially covert becoming overt*). An embedded between-subject two-by-two factorial used the pitch-trim-down malfunction occurring with or without an auditory alert (an alteration from Study 1) and in one of two flight modes (cruise climb; final approach/ILS) as additional independent variables. We had noted in Study

1 that a number of pilots either could not hear the autopilot warning tone (determined by interrogation at the time) or could not recall hearing one (posttest interview). The additional condition was an attempt to determine if the auditory alarm had a significant effect for the specific failure associated with it (runaway pitch trim down). Dependent variables again included flight performance data and states of critical controls; autopilot disconnect/engage, circuit breakers, and pitch trim switches.

Subjects

Pilots who were instrument rated and had experience with complex aircraft and autopilot systems were again obtained from the local area. Pilot ages ranged from 20 to 57 years (median = 40) and the sample contained 22 men and two women. A number of the participants had been involved in Study 1, albeit nine months beforehand. They were intentionally included to increase participant familiarity with both the simulator and with the functioning of the simulated autopilot. In this way we hoped to have something better than a "worst-possible-case" scenario, and something a little closer to the familiarity one might expect with the aircraft most of these individuals were flying regularly. Previous flight experience (total hours) ranged from 290 to 10,000 hours (median = 2230).

Equipment/Procedures

The simulator, instrument flight plan, and overall procedures were identical to those used in Study 1. The session again concluded with an autopilot experience questionnaire and interview. Only the pitch trim malfunction produced both auditory (for half of the subjects) and visual warnings on the autopilot control panel. The presentation order for the new malfunctions can be found, again, in Figure A2 of the Appendix.

STUDY 2 RESULTS

Subsample Differences

Of immediate concern was how those pilots who had participated nine months earlier had performed in comparison with the fully naïve individuals. Examination of the dependent variables by subsample failed to reveal any systematic or reliable differences in performance between the two groups. Thus, subsequent analyses were performed on the full sample.

Runaway Roll Servo

The roll-servo failure emulated the servo-mechanism running the aileron to its stop (full deflection). The following data are the times from initial failure to first response and disconnect of the autopilot by any means (yoke-mounted disconnect, panel disengage, circuit breaker). First-response times ranged from 1.09 to 4.88 seconds (Mean = 3.17; Median = 3.11). A summary of all RT means by conditions appears in Table 2. Note that 90% of the pilots (Figure 5A) disconnected within 4.5 seconds of the initial failure and half within 3.5 seconds. Time to disconnect the AP ranged from 1.49 to 42.77 seconds (Mean = 7.29; Median = 3.11). Almost 80% of the pilots (Figure 5B) had disconnected in less than 5 seconds. Subsequent times to return to zero-degrees bank are shown in

Table 2. Study 2 response time mean, median, and range by failure type and response stage.

Failure Type	Response Stage	Response Time		Range	
		Mean	Med	Low	High
Roll Servo	First Response	3.17	3.11	1.09	4.9
	AP Disconnect	7.29	3.11	1.49	42.8
ADI failure	First Diagnosis	48.8	34.8	12.7	263.0
	Positive ID	58.8	39.6	13.8	264.6
	Return to level	22.1	21.7		
Pitch sensor Down	First Response	16.6	12.5	0.3	73.7
	AP Disconnect	24.8	15.4	5.9	73.7
Pitch Trim Down	Initial action	12.2	6.1		
	Circuit Breaker Pull	36.4	16.1	3.6	160.0

Appendix A, Figure A3. Associated flight-performance data are presented in a later section.

Attitude Indicator (ADI) Failure

When the attitude indicator failed, it drifted slowly to approximately a 25 to 30 degree right-bank indication when the aircraft was in level flight. The result was that the autopilot attempted to follow the failed instrument, placing the aircraft in a left bank. This was not a failure of the AP system but rather, a failure of the sensor feeding data to the system. We were particularly interested in how long pilots took to diagnose the problem. Initial diagnosis (recognition of the general problem) times ranged from 12.7 to 263 seconds (mean = 48.83; median = 34.82). Times to positive identification of the failed ADI ranged from 13.83 to 264.6 seconds (mean = 58.79; median = 39.63) See Appendix A, Figure A5A. Regarding return of the aircraft to level flight, first crossing of zero-degrees bank required an average of 22.11 seconds (median = 21.68). See Appendix A, Figure A5B. Thus, as would be expected, regaining flight control preceded complete diagnosis. This was aided by the visible, albeit faint, horizon between the cloud layers.

Soft Pitch (Pitch Sensor)

The pitch-sensor failure caused a slow deviation from level pitch while the ADI continued to show correct pitch indications, simulating loss of sensor data to the autopilot. First response to this failure ranged from 330 msec to 73.7 seconds (mean = 16.62; median = 12.51). See Appendix A, Figure A6A. AP disconnect times ranged from 5.91 to 73.7 seconds (mean = 24.8; median = 15.4). See Appendix A, Figure A6B. Although 60% of the pilots disconnected in less than 20 seconds, 33% fell between 30 and 60 seconds. This was due both to the comparative subtlety of the failure and to the ability of pilots to manually override the pitch servo without disconnecting.

Runaway Pitch-Trim Down

This failure was different from the others in that only the Pitch Trim circuit breaker would correct the problem. The interim solution was to hold the AP disconnect/trim interrupt switch. The majority of initial responses were yoke AP disconnects, later followed by pulling of the circuit breaker.

Both time to detect a malfunction/initiate action (using autopilot disconnect, control-wheel steering, panel-mounted autopilot engage switch or circuit breaker) and the lag between the initial action and pulling the pitch-trim circuit breaker were of interest. Average time to initial action was 12.2 seconds (median = 6.14). One can see in Figure 6A that 75% of the responses occurred in less than 10 seconds; 90% of the cases occurred in less than 15 seconds. Latencies to pulling the pitch trim circuit breaker averaged 36.4 seconds (median = 16.1; range: 3.6 to 16.0). It is clear from the distribution (Figure 6B) that two outliers (120 & 160) contributed to the inflated mean.

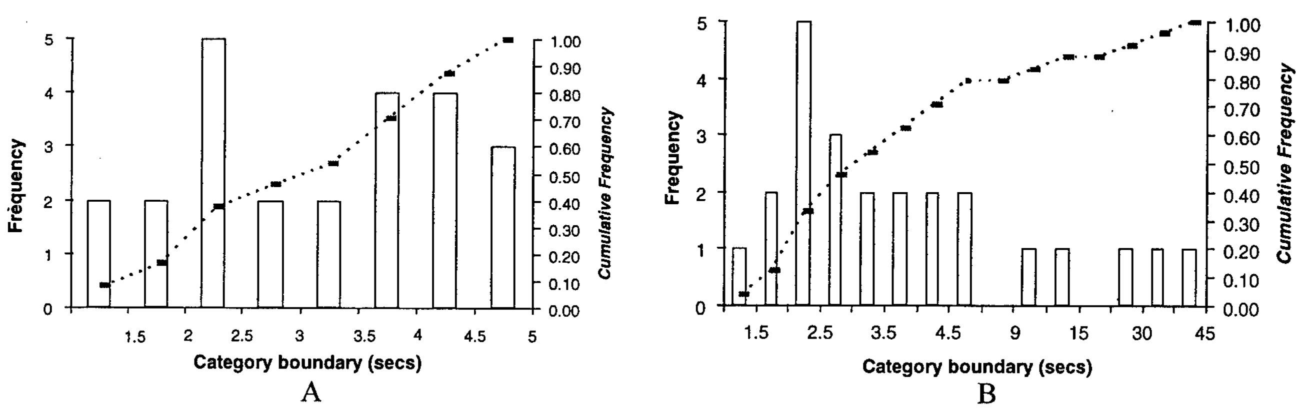


Figure 5. Runaway roll servo response-time distributions and cumulative frequency plots for (A) first-response and (B) AP disconnect.

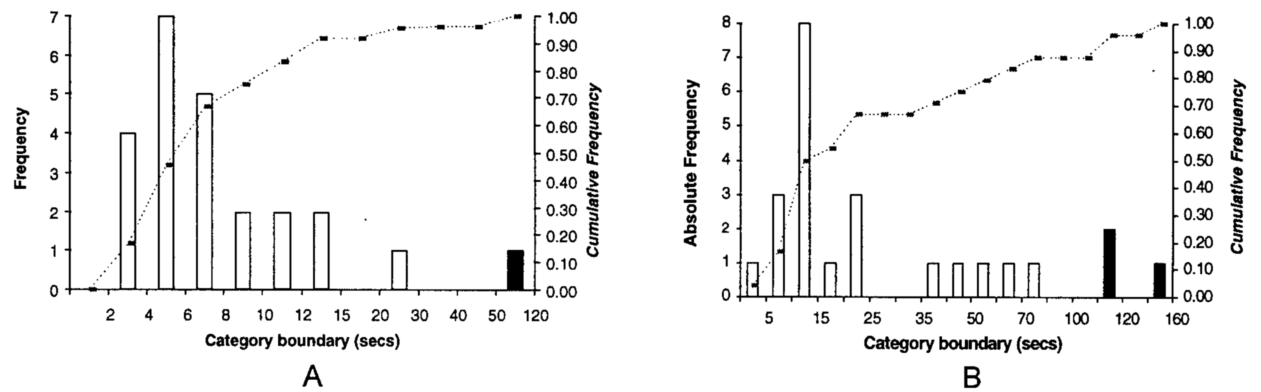


Figure 6. Runaway pitch-trim down response-time distributions and cumulative frequency plots for (A) first disconnect and (B) circuit breaker pull.

Ultimately, the most interesting questions about these data are how many pilots successfully recovered from the runaway pitch trim down malfunction and whether the auditory warning materially contributed to safe recoveries. Table 3 shows the distribution of potential ground contacts and overspeeds (simulator was frozen when high descent rates persisted within 100 feet of the ground or overspeed conditions were attained). Thirteen of the twenty-four participants encountered flight-terminating circumstances. Although the small sample size precludes statistical analysis, it appears that neither the mode of flight nor presence of an auditory alarm materially affected the distribution. This was also the case for time to first response (Figure A7).

Table 3. Study 2 distribution of potential ground contacts and overspeeds by flight mode and alarm presence.

	Alarm	No Alarm	*Total*
Climb	**3**	**4**	7
Approach	**4**	**2**	6
Total	7	6	13

Flight Performance Variations By Maneuver

It is also of interest to examine pilot performance relative to the other malfunctions. Table 4 depicts average maximum deviations in pitch, roll, airspeed, and altitude for each of the four malfunctions for those pilots who were judged to have recovered successfully. These were variations observed between the onset of the failure and the time recovery was judged to have occurred. The roll servo, being the more overt of the two roll failures, produced the lesser average maximum bank (38 degrees), whereas the more subtle ADI failure caused a 10-degree greater average bank excursion (48). Pitch deflections were about the same, however. For those pilots who successfully recovered from the runaway pitch trim, the average maximum pitch down was greater by 3 degrees for those who experienced the malfunction on approach.

It is also instructive to examine representative recoveries by studying the flight profiles. Figures 7 and 8 show two such recoveries, plotting values of altitude, airspeed, and pitch attitude by time during the

Table 4. Study 2 average maximum deviations in pitch, roll, altitude, and airspeed by malfunction type. Runaway Pitch Trim (RPT) is also categorized by flight realm.

Malfunction:	Roll servo	Roll sensor	Pitch sensor	RPT: climb	RPT: approach
Pitch/deg	-3	-4	-2	-9	-12
Roll/deg	-38	-48	0	0	-4
Alt. MSL	5886	4708	6942	6478	1292
change	*-96*	*-292*	*-58*	*-522*	
Air Speed	160	168	155	137	91
n	24	24	24	5	6

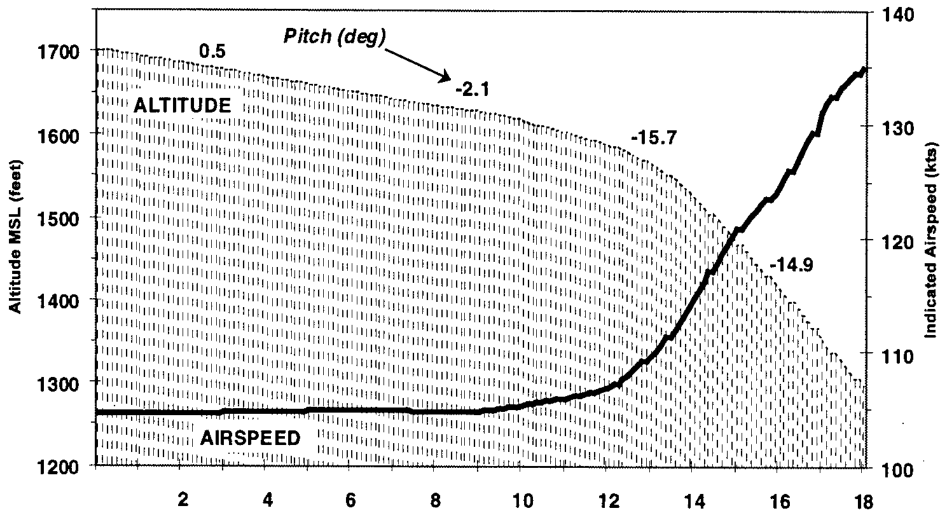


Figure 7. Example flight profile of one pilot's ILS approach depicting altitude MSL (shaded bars) and airspeed (line). Inset numbers represent pitch attitude.

progress of the malfunction. Each plot progresses from the onset of the runaway pitch-trim down to the conclusion of the malfunction. In Figure 7, the malfunction occurs for this pilot on the ILS approach and the data trace the aircraft from onset to trial termination. Note that pitch attitude begins at 0.5 degrees about 3.0 seconds into the malfunction and reaches a maximum of -15.7 degrees at approximately 13 seconds, which is about 5 seconds before termination.

Figure 8 depicts a runaway pitch trim encountered during climb from 6000' to 7000' (onset at 6500'). Pitch varies from +2.6 degs at onset to -2.8 degs after 1.6 seconds, progressing to -18.3 degs at 6.7 seconds and concluding at -30.3 degs just prior to the simulator being frozen (airspeed greater than 200 kts). Both of these profiles are typical of the performances of those pilots who did not recover from the malfunction.

Posttest Questionnaire/Interview

With reference to the most advanced license/rating obtained, this sample contained four Private, eight Commercial, and 12 Airline Transport Pilots (ATPs). Half of the pilots were either certified flight instructors or certified instrument instructors. The median number of years of flying experience was ten. When asked about the recency of their autopilot training, this group indicated a median of three years since last training, with one pilot having received a refresher session the week before the experiment and another pilot reporting that he received his training ten years prior to the experiment. The group indicated that their real-world autopilot flights were usually of one-hour duration, and 64% reported that their most recent autopilot flight had occurred more than six months prior to the experiment. Correlational analyses revealed no significant relationships between pilot experience variables and pilot performance variables.

When asked to report on the difficulty or ease of diagnosing and recovering from autopilot failures experienced during their experimental session, our subjects unanimously agreed that runaway pitch trim was the most difficult from which to recover. The most difficult failure to diagnose was a three-way tie: ADI, pitch sensor, and runaway pitch trim, with each failure receiving 27% of the votes. Pitch sensor was voted the easiest to diagnose by 46% of the subjects, with runaway pitch trim being cited by 36%. Pitch sensor was voted easiest to correct by 56% of the subjects.

All pilots understood that they could overpower the autopilot servos manually. A number were aware of the potential interaction between runaway pitch trim and autopilot pitch attitude (elevator servo) inputs, whereby the autopilot-driven elevator servo masks the initial stage of the pitch trim excursion.

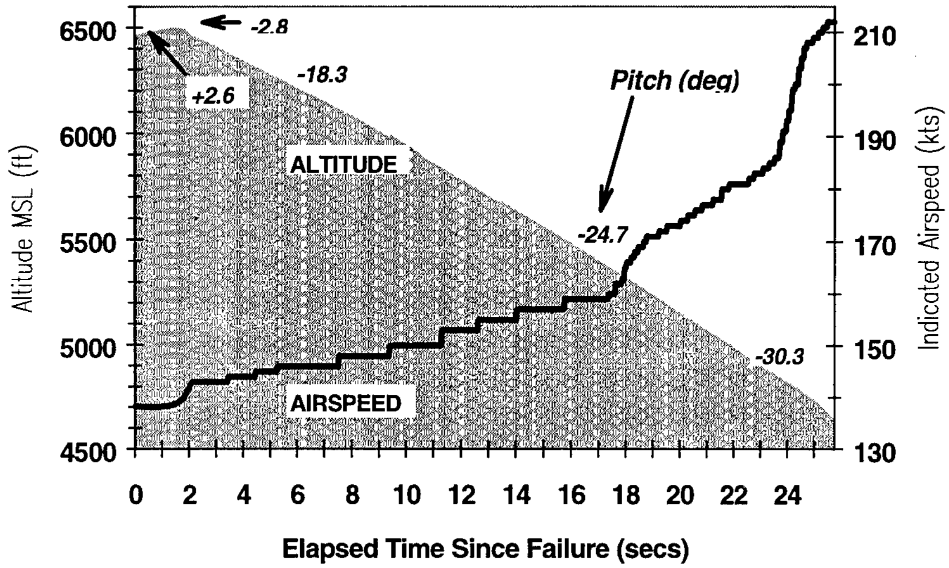


Figure 8. Example flight profile of one pilot's response to runaway pitch trim at altitude during climb. Inset numbers represent pitch attitude

GENERAL DISCUSSION AND CONCLUSIONS

Present certification assumes that a malfunction will be either severe enough to produce supra-threshold cues or that system auditory alerts will warn the pilot, thus starting the clock on the three-second "recognition" period. Flight test personnel (FAA Aircraft Certification Service, 1996) have reported instances where malfunctions have gone undetected until pointed out by the test administrator, sometimes after passing criterion limits. These autopilots failed to obtain certification.

Our data from Study 2 indicate that pilots responding to a supra-threshold failure, runaway roll servo, and who are intent upon an immediate response required an average of 7.29 seconds to respond with an autopilot disconnect, some requiring up to 42.8 seconds. Note that the median response time (3.11) fell within the 4 seconds used as a practical test criterion. One could accommodate 80% of the present pilot sample by specifying five seconds as the upper bound of the interval. However, an unattended roll-servo hard failure, at approximately 15 deg/sec for this class of aircraft, would exceed the current certification criteria in four seconds. In most cases, we observed opposite yoke input prior to or concurrent with the AP disconnect, such that bank criterion was not reached in the vast majority of cases.

In reference to the experimental findings in the context of a fixed-base simulator, the lack of any appreciable effect on interpretation appears to be supported by the fact that a comparison of the data for the two bank malfunctions showed that the subtle ADI failure required longer to detect and produced greater average maximum bank deviations than did the roll-servo failure. Also notable is that the slower roll rate for the ADI failure makes the difference in achieved bank even more significant. Pilot response during the initial stages of runaway pitch trim, where the pitch servo still has sufficient authority to counteract trim, is also unlikely to benefit from acceleration cues in the simulation. Due to the potential contribution of onset acceleration to the detection of the more overt malfunctions, motion-base simulator and/or aircraft validation of results is being pursued for the runaway servo and runaway pitch trim malfunctions.

It should be noted that the actual KAP-150 disconnects on a runaway trim, but our simulated KAP-150 did not. This allowed the pitch servo to compensate for (and mask) the initial trim deflection, as is possible in some other autopilot systems. Although the auditory trim malfunction warning provided an immediate cue, no detectable difference was present in performance between the two alerting groups. Failure of some pilots to hear the warning suggests a reevaluation of criteria for GA cockpit auditory warnings, with specific attention to the noise-exposed and aging populations.

Roles of Contributing Factors

It was apparent from the performances of many of the pilots and from the posttest interviews that additional *training* would greatly benefit the GA pilot population in responding to this particular class of malfunctions. Potential for a benefit may be inferred from: the slightly shorter response times found in Study 2 to malfunctions comparable to those in Study 1, from the correlations between recency of training experience and response times, and from the comments pilots made concerning their preferences for such training and the subsequent effects the "training" experience during the experiment had upon their subsequent flying. It was also clear that this training would benefit the pilots most if it contained both *procedures* for responding to identifiable malfunctions and a thorough explanation of the workings of the autopilot system and its interaction with and use of the elevator trim (*conceptual model* development). Such an effort should lead to a reduction in the frequency of *misdiagnoses.*

One must also find ways to work through *organizational policies* regarding procedures and help pilots differentiate between malfunctions that may be safe to "fly through" (i.e., failure of AP to hold heading) and those that should receive an immediate disconnect. *Cost* is still a highly motivating factor for most pilots, and gaining additional data for the service technician during a "fly through" may continue to influence individuals to allow a malfunction to continue and be observed rather than to be terminated using the autopilot disconnect or appropriate circuit breaker.

Finally, the *human performance* and *human factors* issues involve both the time required by the average pilot to respond adequately and, as a potential facilitator of that response, the means by which malfunctions are brought to the pilot's attention. Additional time needs to be provided, in some instances, for pilots to respond, particularly for the subtler malfunctions. This does not necessarily affect autopilot performance specifications, specifically because subtle failures are unlikely to cause the aircraft to exceed performance limitations within the presently specified three-second waiting period. However, should the failure be so subtle as to place the aircraft in an unacceptable attitude without the pilot's detection, present standards would disqualify that autopilot. Avoiding this disqualification depends upon either having the pilot detect and respond to the malfunction, either unaided or with the assistance of a warning device, or upon having a system that is either (a) so reliable that such malfunctions do not occur or, (b) that has automatic monitoring capabilities that sense, take action (disconnect), and inform the pilot of that action.

Present guidelines appear adequate for failures accompanied by high acceleration rates and those that require simple procedural responses. Findings for the auditory alarm presence/absence in these studies suggest that there are some detection problems associated with the more senior pilots, particularly in the frequencies at or above 3KHz, and an additional study is being conducted to provide recommendations for more detectable, differentiable, and attention-attracting alerts without any negative "startle" effects.

In summary, the potential recommendations coming out of this study include:

- Require initial and recurrency response-to-failure training; include in biennial flight review.
- Lengthen specified delay in pilot response during certification trials for subtle failures.
- Expand use of failure annunciation or "fail-safe" modes in autopilot devices, as in the KAP-150.
- Obtain baseline hearing threshold curves for pilot and nonpilot samples to determine the extent of hearing loss by age cohort, with possible recommendations for modifications to hearing assessment procedures.
- Evaluate effect of noise-attenuating and noise-canceling headsets on pilots' detection of presently used auditory warnings, with potential recommendations for integrated auditory warnings presentation through intercom/headset systems.

REFERENCES

ACE-110 (Small Airplane Directorate, FAA) (1996). Personal communication.

Beringer, D.B. (1996). Automation in general aviation: responses of pilots to autopilot and pitch trim malfunctions. *Proceedings, 40th Annual Meeting of the Human Factors and Ergonomics Society.* Santa Monica, CA: HFES, 86-90.

Beringer, D.B. and Harris, H.C., Jr. (1996). Pilot responses to automation failures in general aviation: autopilot and pitch trim malfunctions. FAA/CAMI, *Technical Memo TM-AAM-500-96-02.*

Cooling, J.E. & Herbers, P.V. (1983). Considerations in autopilot litigation. *Journal of Air Law and Commerce, 48*, 693-723.

FAA (1996). Human Factors Team Report on: The interfaces between flightcrews and modern flight deck systems. Washington, DC: Federal Aviation Administration, June 18, 1996.

FAA Aircraft Certification Service (1996). Personal communication with flight test personnel (ACE-115W).

FAR 23.1329; Advisory Circular 23.1329-2, Automatic Pilot System Installation in Part 23 Airplanes, 3/4/91.

Funk, K., Lyall, B. and Riley, V. (1993). A comparative analysis of flightdecks with varying levels of automation. Phase 1 Final Report, FAA Grant 93-G-039.

Hoh, R.H., Smith, J.C. and Hinton, D.A. (1987). The effects of display and autopilot functions on pilot workload for single pilot instrument flight rule (SPIFR) operations. Washington, D.C.: NASA Scientific and Technical Information Office, *NASA Contractor Report 4073.*

Katz, P. (1995). NTSB Debriefer: The dark side of "Otto pilot." In *Plane & Pilot*, (February), *31(2)*, 18-19.

Wilson, B.G. (1995). I learned about flying from that (#660): Unacquainted with the autopilot. In *Flying, (June), 122(6),* 122.

APPENDIX A

AGARS Block Diagram

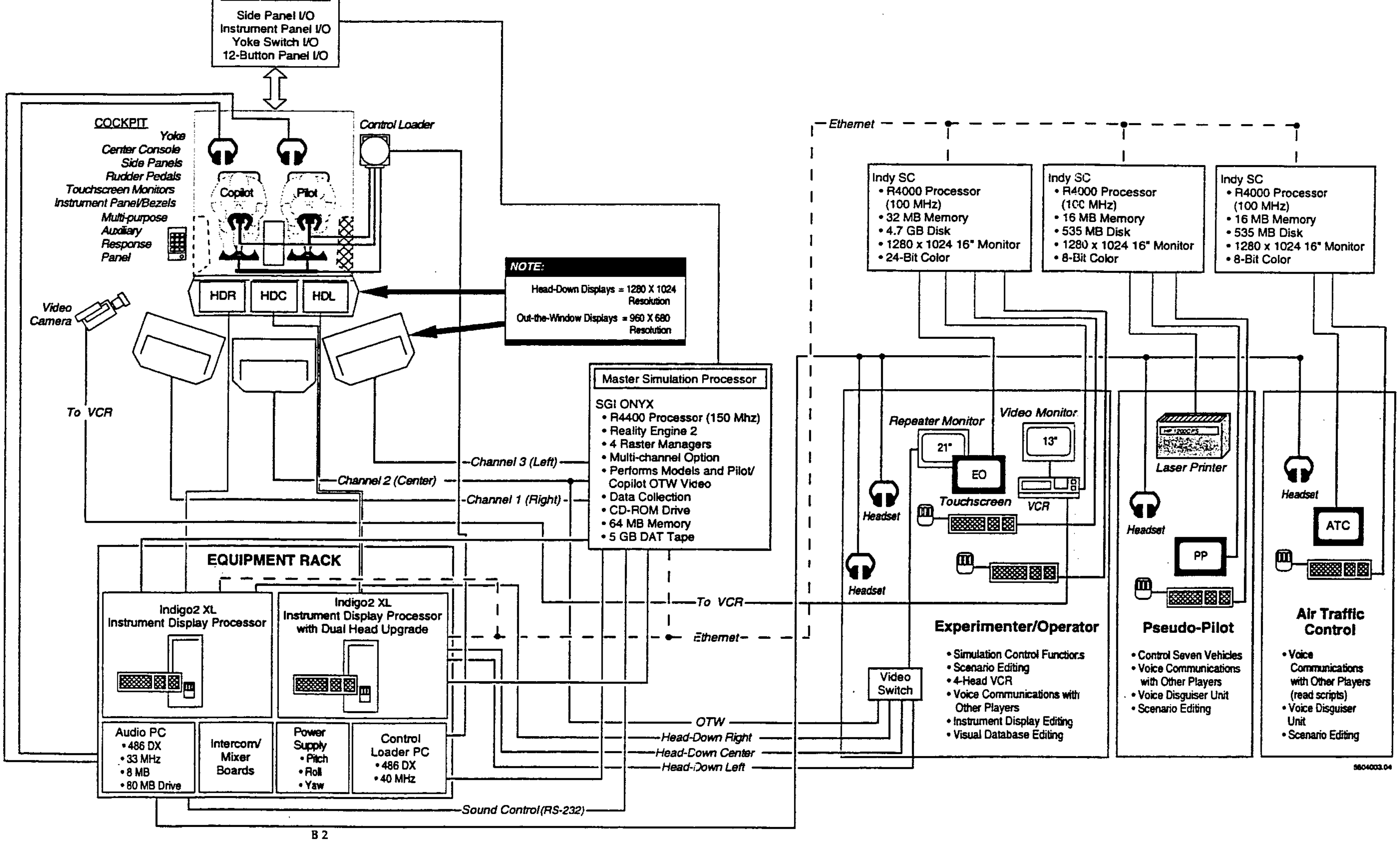


Figure A1. Block schematic diagram of the Advanced General Aviation Research Simulator.

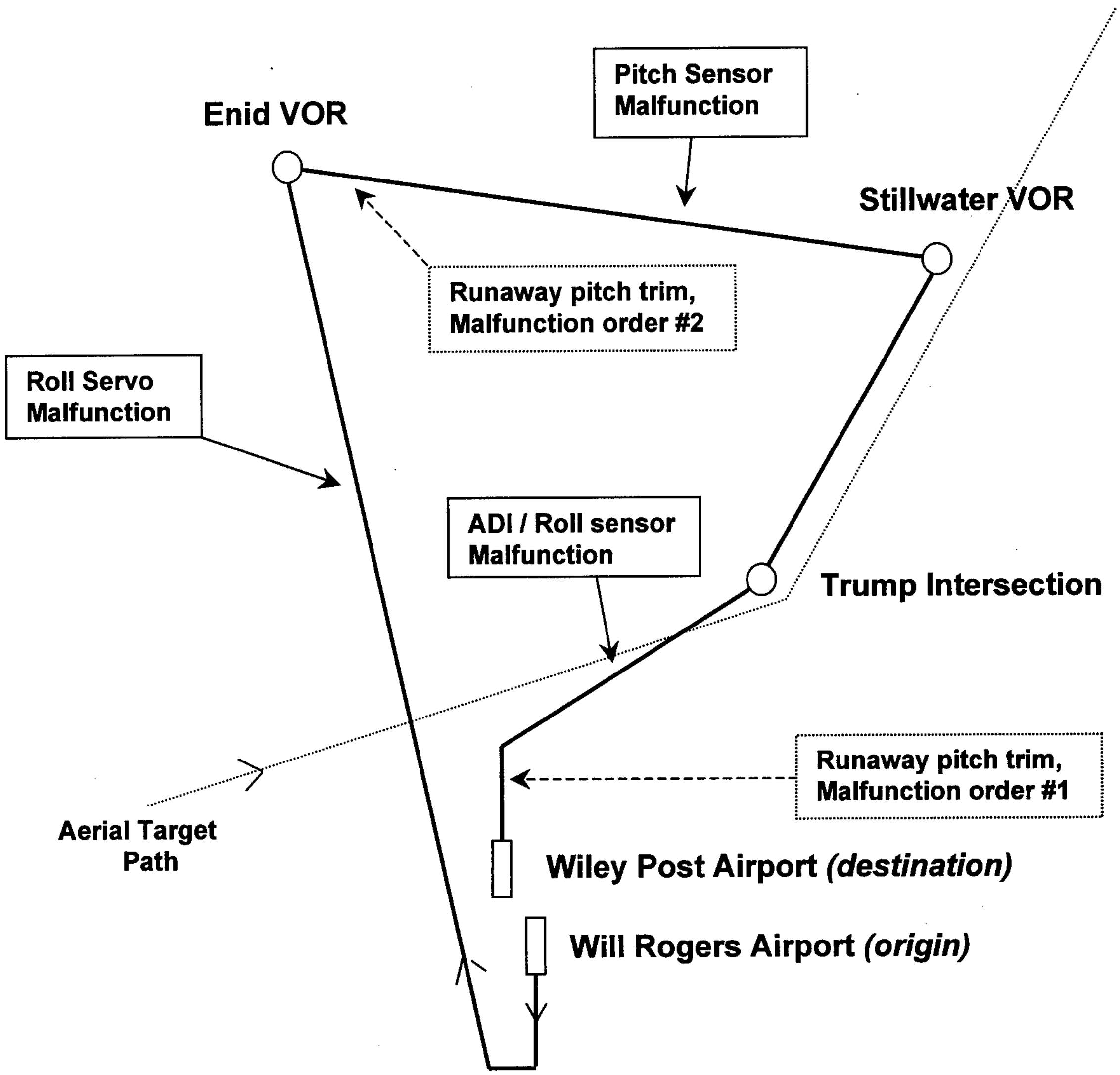


Figure A2. Experimental flight path with annotations showing malfunction event points along route.

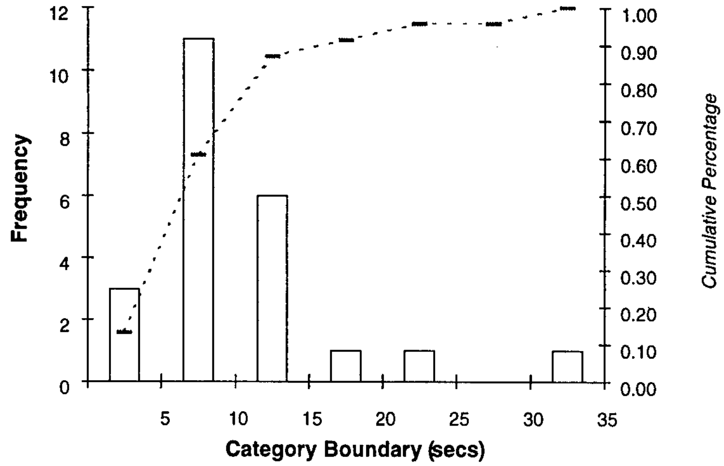
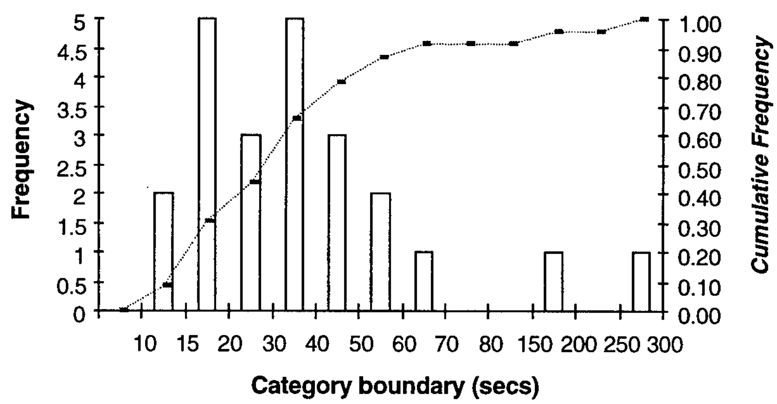
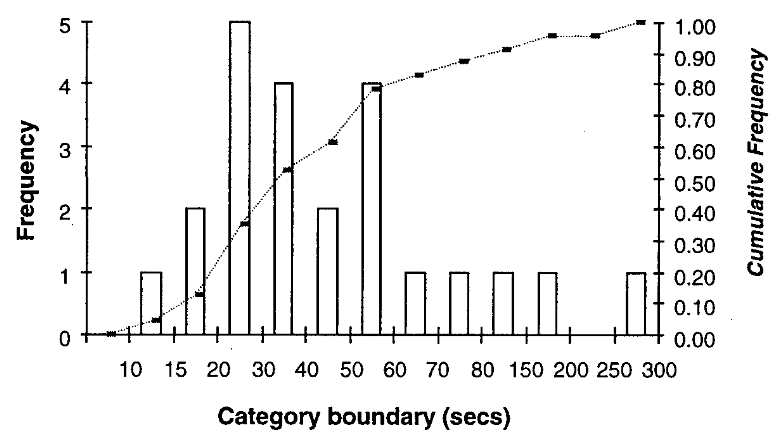


Figure A3. Runaway roll servo; distribution of recovery times to first zero-degree bank crossing and cumulative frequency plot.



A



B

Figure A4. ADI failure response-time distributions and cumulative frequency plots for (A) initial diagnosis and (B) positive identification.

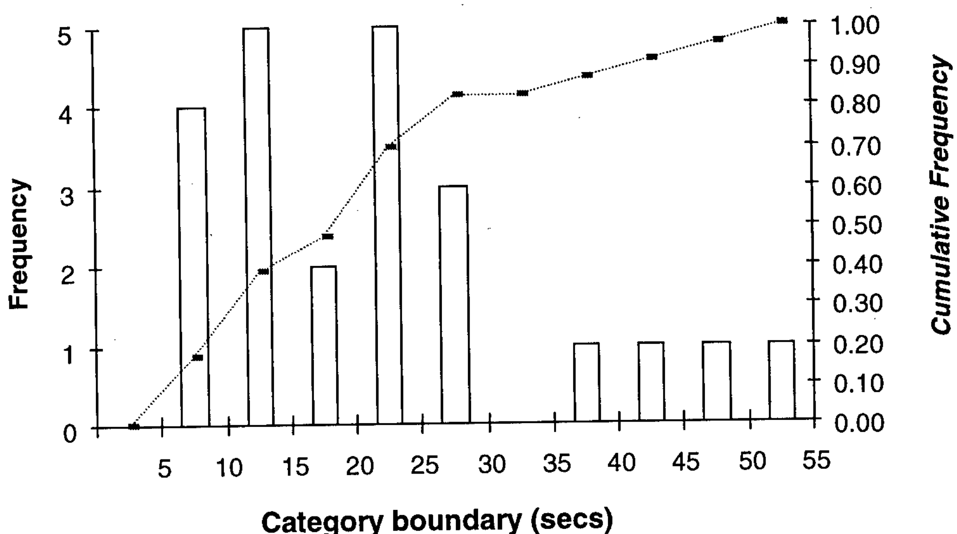


Figure A5. ADI failure; distribution of recovery times to first zero-degree bank crossing and cumulative frequency plot.

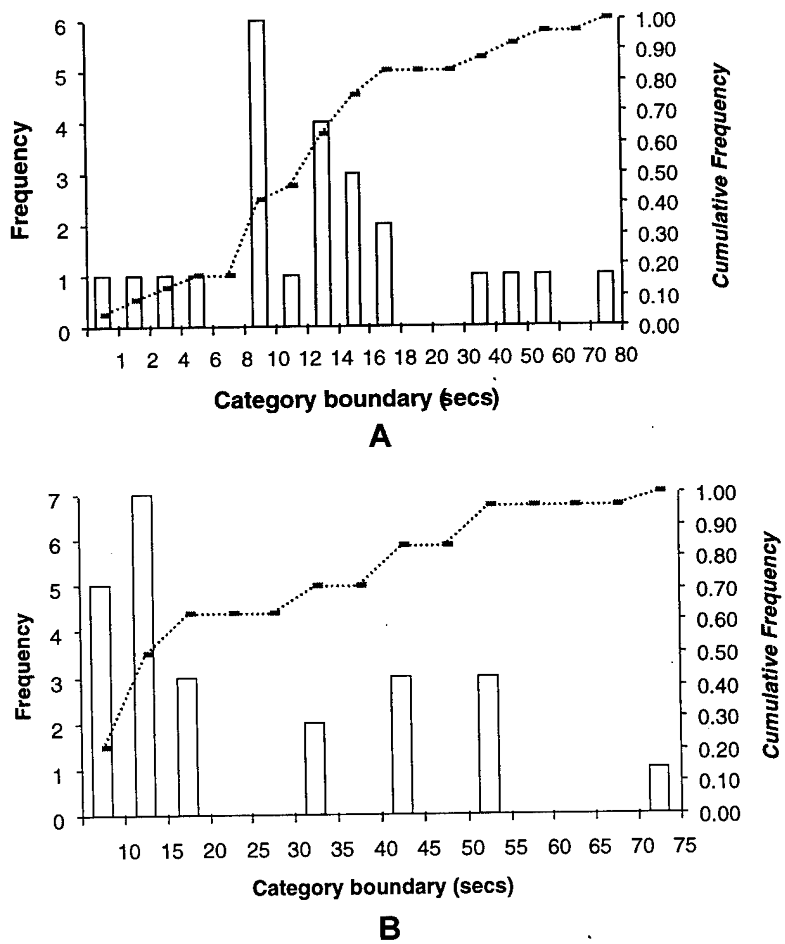


Figure A6. Pitch sensor failure response-time distributions and cumulative frequency plots for (A) first response and (B) AP disconnect.

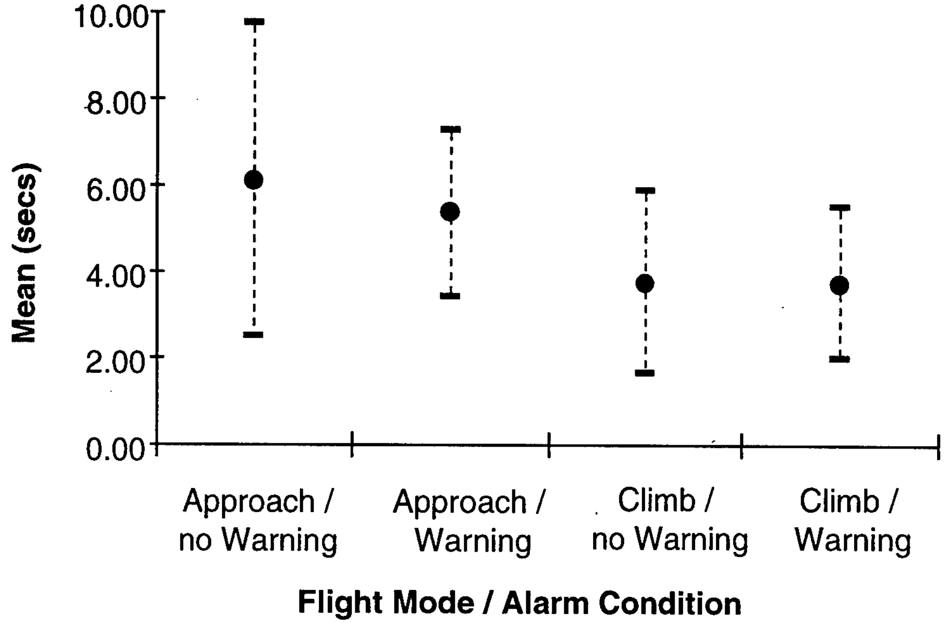


Figure A7. Mean and standard deviation for first-response time to runaway pitch trim down by flight mode and warning condition.

☆U.S. GOVERNMENT PRINTING OFFICE: 1997 -660-088/60048