NASA Technical Memorandum 78482

A Simulator Study of the Interaction of Pilot Workload With Errors, Vigilance, and Decisions

H. P. Ruffell Smith

JANUARY 1979

NASA

NASA Technical Memorandum 78482

A Simulator Study of the Interaction of Pilot Workload With Errors, Vigilance, and Decisions

H. P. Ruffell Smith
Ames Research Center
Moffett Field, California

NASA
National Aeronautics
and Space Administration

Scientific and Technical
Information Office

1979

TABLE OF CONTENTS

A SIMULATOR STUDY OF THE INTERACTION OF PILOT WORKLOAD WITH ERRORS, VIGILANCE, AND DECISIONS

H. P. Ruffell Smith*

Ames Research Center

SUMMARY

This research comprised a full mission simulation of a civil air transport scenario that had two levels of workload. Twenty fully qualified three-man crews took part in the study. The actions of the crews and the basic aircraft parameters were observed and heart rates were recorded.

Reduction of these data permitted the enumeration of errors, vigilance, decisions, and their association with heart rate to be investigated.

The results showed that the number of errors was very variable among crews but the mean increased in the higher workload case. The increase in errors was not related to rise in heart rate but was associated with vigilance times as well as the days since the last flight.

The recorded data also made it possible to investigate decision time and decision order. These also varied among crews and seemed related to the ability of captains to manage the resources available to them on the flight deck.

Error rates and heart rates were essentially the same as those found in actual flight operations, indicating the quality of the simulation. It is suggested that similar levels of full mission simulation could benefit training and accident investigation.

INTRODUCTION

In 1974 the author made proposals to NASA Ames Research Center for a full mission simulator experiment to study the performance and error rate of a representative sample of aircrew currently operating scheduled routes. It was suggested that performance might be downgraded during high workload situations, particularly those related to abnormal operation or equipment malfunction.

The National Research Council agreed to support the author for a study to be carried out at the NASA Ames Research Center, and preparations for a trial of this nature began in July 1975. Early in 1976 a full mission simulation experiment was conducted that exposed 20 three-man civil airline crews to low and high workload situations. This paper describes how their performance changed in respect to errors, vigilance, and decisionmaking. As it seemed

---

*National Research Council Associate.

probable that arousal would be correlated with some aspects of performance, it was decided to record the heart rate of the three crew members for the entire time they were in the simulator.

An experiment of this size cannot be undertaken without the cooperation of many people and organizations. Mr. John Griffin, a student at Foothill College, edited the background recordings of ATC communications and integrated these with the scenarios. He also helped with the reduction of the audio and heart-rate data and with the compilation of errors. Mr. George Cooper, who recently retired as Chief Test Pilot of NASA Ames Research Center, gave expert help in devising the scenario and in running the study. I am indebted to the National Research Council for their support and I am most grateful for the continuing interest and guidance of my scientific adviser, without whom this study would not have been undertaken.

METHOD

Full mission simulation requires a degree of realism available only in the most sophisticated airline training simulator and the subjects must be aircrews current in the relevant type of aircraft and operation. Relating this study to a contemporary wide-body transport aircraft seemed appropriate because they have been used by the airlines for several years and are likely to be one of the mainstays of civil aviation for many years to come.

The scenario cannot be written without detailed knowledge of the operation of the airline, the air traffic control system and the systems and capabilities of the chosen aircraft. The observers must also have the same kind of knowledge. Recording facilities must be available for their comments, as well as a continuous printout of values for aircraft performance, navigation, and communication.

Simulator

The simulator used in this study, manufactured by Singer-Link, had a motion platform with six degrees of freedom. The flight deck was configured for long-haul operation, and there were three crew stations for captain, copilot and flight engineer (these positions and their occupants are annotated as P1, P2, and P3, respectively, throughout this report). In addition to the three positions for the crew on the flight deck, the simulator cab also included positions for pilot and engineer instructors who also, in normal training function, acted as simulator operators. There were facilities for inserting aircraft, runway, and environmental conditions into the simulation program. Aircraft system faults could also be instigated without the knowledge of the crew members. The simulator was equipped with a visual attachment that allowed the pilot to transfer from instrument to visual flight during the approach. The visual system could be programmed by the simulator operator to mimic changes in cloud base.

The crew members could converse either with or without electrical "intercom." In the latter case, their conversation was picked up by a separate microphone in the roof of the flight deck. Air traffic control (ATC) communications for the particular flight were simulated by the operator, who also acted as the airline dispatcher when the crew communicated with the operations department. The simulator was modified so that all air-to-ground communications could be recorded.

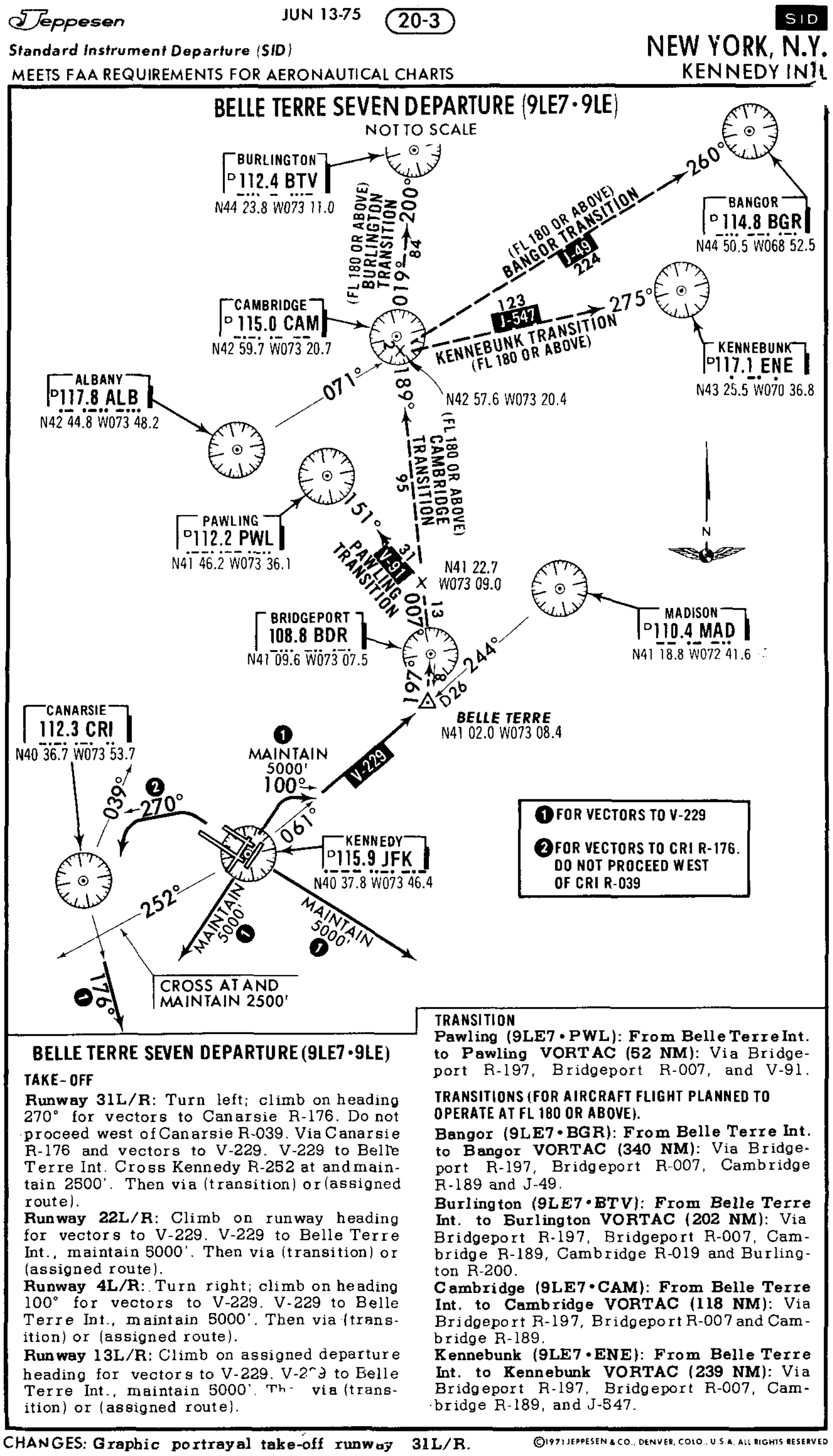
To increase realism and provide a standard level of distraction for each section of the flight, background ATC communications with other aircraft were injected into the crew's intercom. This was accomplished by recording messages during actual flight operations. Twelve different tapes were made for each of the two sectors of the scenario, covering the routes from Dulles to Kennedy airports and the route from Kennedy to the point of return dictated by the scenario for the flight to London. The recordings included ground control and tower messages for each airfield, clearance delivery, Automatic Terminal Information Service (ATIS), relevant departure and approach controls, as well as enroute sections of the flights.

Scenario

Simulator time was available in blocks of 4.5 hr, which the scenario had to be constrained to fit. It was necessary to choose routes unfamiliar to the participating crews that would provide both high and low workload. The simulated operation was that of a charter service from Dulles (Washington) to Heathrow (London), stopping briefly at Kennedy (New York) for fueling and uplifting of serendipity cargo. The scenario could thus be divided into two separate sectors.

The first sector, requiring about 45 min flying time, was designed to be conducted using standard operating air traffic and navigational procedures. During this sector neither weather nor runway conditions for takeoff or landing would provide any difficulty for a regular airline crew, and, although one of the autopilots was unserviceable, no further mechanical failures occurred. Nevertheless, during the cruise phase of the first sector, the crews were warned about thunderstorm activity by ATC and were subsequently diverted around this by radar vectors. No action was required by the crew except to activate engine nacelle ice protection and to warn the passengers to fasten their seat belts. It was hoped that this initial sector would provide a comparatively low workload situation, allow crews to become familiar with each other, and forgetful of the ECG electrodes and simulator environment. At the end of the first sector, a quick turn-around, lasting about 30 min, was simulated while extra fuel and cargo were loaded.

The second sector was planned to produce a high workload situation requiring decisionmaking in the context of complicated interacting factors. The aircraft was very heavy (almost at maximum gross weight for takeoff), the runway conditions and wind were marginal for this weight, the standard instrument departure (SID) from New York was complicated, and the published instructions for it unclear (fig. 1). There was also a last minute change in the



Jeppesen JUN 13-75 (20-3) SID

Standard Instrument Departure (SID)

MEETS FAA REQUIREMENTS FOR AERONAUTICAL CHARTS

NEW YORK, N.Y.
KENNEDY INTL

BELLE TERRE SEVEN DEPARTURE (9LE7•9LE)

NOT TO SCALE

❶ FOR VECTORS TO V-229

❷ FOR VECTORS TO CRI R-176. DO NOT PROCEED WEST OF CRI R-039

BELLE TERRE SEVEN DEPARTURE (9LE7•9LE)

TAKE-OFF

Runway 31L/R: Turn left; climb on heading 270° for vectors to Canarsie R-176. Do not proceed west of Canarsie R-039. Via Canarsie R-176 and vectors to V-229. V-229 to Belle Terre Int. Cross Kennedy R-252 at and maintain 2500'. Then via (transition) or (assigned route).

Runway 22L/R: Climb on runway heading for vectors to V-229. V-229 to Belle Terre Int., maintain 5000'. Then via (transition) or (assigned route).

Runway 4L/R: Turn right; climb on heading 100° for vectors to V-229. V-229 to Belle Terre Int., maintain 5000'. Then via (transition) or (assigned route).

Runway 13L/R: Climb on assigned departure heading for vectors to V-229. V-229 to Belle Terre Int., maintain 5000'. Then via (transition) or (assigned route).

TRANSITION

Pawling (9LE7•PWL): From Belle Terre Int. to Pawling VORTAC (52 NM): Via Bridgeport R-197, Bridgeport R-007, and V-91.

TRANSITIONS (FOR AIRCRAFT FLIGHT PLANNED TO OPERATE AT FL 180 OR ABOVE).

Bangor (9LE7•BGR): From Belle Terre Int. to Bangor VORTAC (340 NM): Via Bridgeport R-197, Bridgeport R-007, Cambridge R-189 and J-49.

Burlington (9LE7•BTV): From Belle Terre Int. to Burlington VORTAC (202 NM): Via Bridgeport R-197, Bridgeport R-007, Cambridge R-189, Cambridge R-019 and Burlington R-200.

Cambridge (9LE7•CAM): From Belle Terre Int. to Cambridge VORTAC (118 NM): Via Bridgeport R-197, Bridgeport R-007 and Cambridge R-189.

Kennebunk (9LE7•ENE): From Belle Terre Int. to Kennebunk VORTAC (239 NM): Via Bridgeport R-197, Bridgeport R-007, Cambridge R-189, and J-547.

CHANGES: **Graphic portrayal take-off runway 31L/R.**

Figure 1.- Belle Terre Seven Departure, Kennedy International Airport.

clearance that modified the SID. This departure was followed by a long and uneventful climb for 25 to 30 min to an altitude of 31,000 ft. At the top of the climb, after cruise speed had been attained and the equivalent power selected, slow clogging of the oil filter for No. 2 engine was activated. This clogging affected the differential pressure across the filter and was shown on a combined oil pressure gauge on the systems panel monitored by P3. At a certain value, an annotated warning light on the central warning panel, situated to the right of P1's flight instruments, was illuminated. The allowable pressure differential was exceeded and the engine had to be shut down; thus an opportunity was provided for assessing vigilance and decision times.

The clogged oil filter presents the crew with several interacting problems that need information retrieval and consequent decisions. Because of range and lower redundancy, it is immediately obvious that the flight to London must be abandoned. Furthermore, the 31,000 ft altitude cannot be maintained on the remaining engines without unacceptably overboosting them. This combination of factors creates a need for interaction with air traffic control as to the nearest airport with suitable weather and runway conditions and how navigation to this destination is to be achieved. With the fuel load for London, the aircraft is much overweight for a landing in the northeastern United States over which it is flying. There is a right time and place when enough fuel must be dumped to allow for a safe landing on the chosen alternate airfield, while retaining enough as a reserve for a further diversion should the chosen runway be unavailable.

By manipulating the weather and runway conditions, the scenario forces the decision to return to a particular runway at Kennedy Airport, the others being precluded by the strength of the crosswind and by obstruction from repair work in progress or damaged aircraft; to a certain extent these conditions standardize the tasks and workload. At a particular point during this return, while the third pilot (P3) is engaged in dumping the fuel, the No. 3 hydraulic system is slowly depleted to zero. This further failure provides an opportunity for measuring the vigilance of a busy P3 and further affects the operation of the aircraft. The only important consequence is that the second and only remaining autopilot fails, thus necessitating hand flying of the aircraft and further increasing the workload.

On nearing New York, the crew is required by air traffic control to fly a specific holding pattern at a time when they are concerned with the legality of the marginal weather at New York and the short length of the runway in use. Finally, a hand-flown ILS approach must be made with asymmetric power to a relatively slippery runway with strong crosswind, followed by a landing that is safe only with the correct amount of automatic braking and immediate application of symmetrical reverse thrust.

The second sector thus provides a situation that starts with difficulty and is followed by a period of low workload that tends to lull the crew into complacency. Then there is a sudden onset of circumstances requiring ever-increasing information retrieval, decisionmaking, and skill. Extra harrassment is provided by the simulation of a cabin crew member relaying complaints and requests from upset and anxious passengers. The whole of this scenario was based on situations that had happened in real life, and none of the participating crews considered that the simulated situations were in any way unrealistic.

Subjects

The participating aircrew were all volunteers from the same carrier and were in current flying practice in the aircraft on scheduled airline routes. Twenty crews, each consisting of a captain, a first officer and a flight engineer, were selected from a total of 140 crew members who volunteered on the basis of their availability at the times the simulator sessions had been made available for the study. Although the subjects were volunteers and not selected at random, they formed about 30% of the strength of the aircrew at the domicile from which most of them came.

The first two complete crews were used to make sure that all the recording methods could be integrated and observers positioned so as not to be obtrusive. These preliminary runs also gave the instructor/simulator operator practice in standardizing ATC and communication procedures. Only 18 crews were available for the "data" runs.

The ages of the crew members varied between 54 and 59 for captains, 38 to 50 for copilots, and 45 to 58 for flight engineers. The total years of flying were between 33 and 41 for the captains, 16 and 32 for the copilots, and 20 and 37 for the flight engineers. Total flying time was usually about 500 to 750 hr for every year. The experience of the group on this type of aircraft was more variable, that of the captains between 6 months and 6 years, that of the first officers between 6 months and 4 years, and that of the flight engineers between 1 and 6 years. All the different crew members had accumulated about 600 hr for every year they had been flying the aircraft (table 1).

The information recorded about the participating crew members included days since last flight, days away from home in the last 5 days, and hours out of bed prior to reporting for the simulator run (table 2).

Experimenters

A team of people ran the experiment: an instructor/simulator operator, two observers, and a technician to change the background recordings for the ATC communications. The instructor/simulator operator was a senior flight instructor employed by the participating airline with a lifetime experience of operating in several airlines throughout the world, as well as many years of teaching in simulators. He helped devise the scenario, encouraged the crews to volunteer, and arranged their participation to fit in with their airline work schedule. During the runs he acted as dispatcher before the simulator flights, programmed the scenario in a standardized fashion and gave the benefit of his observation of the performance of the crews at the debriefing.

One observer was the author, who is a physician with a background of human factors in aviation and an experienced pilot familiar with observing

TABLE 1.- AGE AND FLYING EXPERIENCE

Run	Age			Years flying			Total flying hours (in thousands)			Months of flying (in this type aircraft)			Flying hours (in this type aircraft)		
	P1[a]	P2[b]	P3[c]	P1	P2	P3	P1	P2	P3	P1	P2	P3	P1	P2	P3
1	56	38	54	32	20	35	12	6	24	12	7	72	100	310	2500
2	56	43	49	33	18	22	23	5	20	36	30	48	2000	1300	2500
3	55	43	48	37	25	22	24	12	16	66	72	48	3500	3000	2800
4	54	50	57	35	32	35	21	15	24	12	30	48	600	1500	3200
5	58	38	52	38	18	28	34	13	17	72	12	60	5000	700	3000
6	55	48	58	37	21	34	24	12	11	12	48	72	600	2400	400
7	57	46	45	34	24	20	23	12	8	10	48	12	500	2400	400
8	55	50	49	37	32	20	26	18	13	36	36	12	2000	1800	600
9	55	40	51	36	16	28	29	10	16	48	6	66	3000	250	3500
10	54	47	56	36	31	37	24	10	35	36	36	72	1800	1800	4800
11	57	40	57	35	20	34	22	9	30	48	9	66	2500	420	4500
12	57	42	53	35	20	33	25	10	22	6	14	72	350	750	3000
13	55	41	56	35	22	33	28	9	25	12	10	54	500	450	3300
14	54	49	52	35	31	32	26	10	20	12	36	36	600	2000	1800
15	57	45	50	36	25	22	23	20	18	12	48	36	600	2600	2400
16	55	38	48	34	17	23	25	10	19	60	12	66	2500	550	4000
17	59	46	52	41	23	27	26	8	15	72	36	48	4200	1500	2400
18	56	38	54	35	16	26	21	8	22	48	9	60	3000	400	4000

[a] Captains

[b] First officers

[c] Second officers/flight engineers

TABLE 2.- DAYS SINCE LAST FLIGHT, NIGHTS AWAY FROM HOME, HOURS OUT OF BED, AND TIME OF DAY OF SIMULATOR SESSIONS

Run	Days since last flight			Nights away from home in last five			Hours out of bed prior to experiment			Simulator session no. (time of day) (*a*)
	P1	P2	P3	P1	P2	P3	P1	P2	P3	
1	1	3	3	0	2	0	10	12	12	4
2	10	13	5	0	0	0	12	12	12	4
3	3	3	11	2	2	0	4.5	6	8	3
4	7	17	1	0	0	4	3	2	2.5	2
5	4	1	20	0	4	0	13	12	13	4
6	34	11	5	0	0	0	11	11	12	4
7	32	7	34	0	0	0	2	2	1	1
8	1	7	2	5	0	0	10	13	11	1
9	3	4	41	1	0	2	12	11	12	4
10	2	3	5	3	1	1	8	8	8	3
11	11	70	5	0	0	0	2	2	2.5	2
12	3	7	6	2	0	0	5	2	2.5	2
13	4	2	6	5[b]	3	1	2	2	2	1
14	4	2	5	0	1	0	9	6	6	3
15	4	2	3	1	3	2	9	13	9.5	4
16	8	13	10	4	0	0	13	13	12	4
17	3	5	6	3	0	0	9	7	7	3
18	14	43	75	0	0	0	6	4	3	2

[a] Time of delay

Session no.	Report	In simulator		Frequency
1	06.00	07.00	11.15	X3
2	10.15	11.15	16.00	X4
3	15.00	16.00	20.15	X4
4	19.15	20.15	00.30	X7
				18

[b] Three of these in motel for training.

airline flightcrews under many different conditions, both in flight and in simulators. Prior to the experimental part of this study, he attended a 4-wk course of instruction at the ground school of the participating airline to learn about the significant characteristics of the aircraft and its systems. He witnessed the conversion training of aircrew, both in the simulator and subsequently in the air, and undertook observational trips with regular line-crews on the same airline. This preparation was essential, both to mold the scenario and to understand the standard operating procedures and notice any departures from them. During the experiment he gave special attention to the activities of the captains and first officers (P1's and P2's).

The other observer was a psychologist with a pilot's license, a background in human factors and experience in observing the performance of aircrew in flight and in simulators. He gave special attention to the behavior of the flight engineers (P3's). He was also responsible for coordinating the recording equipment and its installation in the simulator and for integrating modifications with its maintenance staff.

The technician was a work-study student from NASA Ames Research Center who was studying aeronautics at a local college and who is a veteran USAF flight crewman (an air-refueling boom operator). He also holds a private pilot's license. The technician's main task was to synchronize the background communication tapes for the various ATC sectors.

Recordings

Five categories of data were recorded: the comments made by the observers and the simulator operator about the performance of the crew members; the computations made by the crew prior to and during the flight; the communications between crew members and the simulated air traffic control; the aircraft parameters related to the task; and continuous electrocardiograms for the three crew members.

Observers' comments— The two observers were in the simulator cab with the crew and were able to comment on their activities. The observers also took note of any aircraft parameters they thought particularly relevant to the current situation and of the time since the start of takeoff. These comments were recorded on two separate direct channels of a continually running seven-track tape recorder.

As much as possible, observers maintained a continuous "running commentary" during the time that the crew were in the cab, so that preflight as well as in-flight activities were recorded, most attention being paid to activities during flight. The activities recorded included: the pilot currently in control of the aircraft, which automatic control modes were selected, the time and sequence of operation of controls affecting aircraft configuration, and the identity and radio frequency of the navigation facilities selected, together with the required track and heading. Notice was taken of the time spent looking at charts and approach plates, the use of normal and abnormal checklists, and the retrieval of data from aircraft and flight operating manuals. Note was also made of the management of aircraft systems by the flight engineer (P3), especially the way he dealt with requirements, such as fuel dumping, and how these interfered with his integration with other crew members.

At the end of each session, the observers and the simulator operator held a debriefing session. During this session an attempt was made to recall any special features of performance in the two sectors, such as significant errors and related factors that might be included in "flying style" and idiosyncrasies. This information was recorded longhand on proformas.

Crew paperwork— All the paperwork relating to each run was collected and later perused for errors. This included the takeoff computation sheet, the form for the second officer's (P3) fuel computation, and the notes prepared for the takeoff and approach indicated airspeeds.

Electrocardiograms— A single lead ECG was taken for each crew member during the whole time he was seated in the cab. Three Hewlett Packard prejelled disposable electrodes had previously been attached at the front and sides of his chest. Leads from the electrodes were attached to preamplifiers in pockets behind the seats. The amplified signals went via screened leads to three FM channels of the seven-track tape recorder at the rear of the cab.

Flight deck communications— Most of the verbal communications between crew members were picked up on a separate microphone situated in the roof of the simulator. The signals from this were recorded on a separate direct channel of the seven-track tape recorder.

The simulator circuits carrying communications made by crew members with ground facilities were tapped, and the signals were also fed into another direct channel of the tape recorder.

Aircraft Parameters— The aircraft parameters considered appropriate for assessing crew performance were obtained from the digital computer controlling the simulator and recorded on a high-speed line printer. There was insufficient time prior to the experiment to make arrangements for this information also to be recorded in digital form on magnetic tape. The values displayed were: indicated airspeed, pressure altitude, instantaneous vertical speed, heading, angle of attack, lateral deviation from the required navigational path in angle and distance, similar deviation in angle and distance from the desired instrument landing system (ILS) glide slope, the position of the landing gear, the position of the flaps in increments of one degree and that of the spoilers, the engine pressure ratio for No. 1 engine (representing power setting), the frequencies set in No. 1 navigation receiver, No. 1 communication receiver, and No. 1 automatic direction finder (ADF), the elapsed time since takeoff, and the total fuel remaining.

The line printer was started as power was applied for takeoff and stopped at the end of the landing run. The values were printed every second below 1800 ft and every five sec above that height.

Procedures

All of the volunteer crew members had previously been made aware of the purpose of the study and had agreed to the recording of their performance as well as their heart rates. They did not know any details of the scenario before they reported at the training facility. They were met by an experimenter who first impressed upon them the need for secrecy as to the content of the scenario until the trials were over, pointing out that even general knowledge of it by subsequent subjects would invalidate the study. Details of their age, experience and other facts as previously mentioned were then recorded. The ECG electrodes were applied, assurance being given that the physiological data, together with any records of performance which would be made during the course of the run, would be guaranteed anonymity.

The three crew members then moved to a briefing room where they met the project coordinator, who acted as the dispatcher. He told them the flight for which they had been scheduled and provided them with the standard paperwork for the flight. This included the departure time, weather for the whole route, the two intended destinations, the routes to be followed, the weight and balance of the aircraft, and the runway available for the departure. All this information was in the same form as they would normally have on a regular flight, having been copied from actual flight documents that had been issued to flights departing from Dulles and Kennedy Airports. The time available for preparation for flight was about three-quarters of an hour.

Usually the crew members discussed the general plans together. After this the second officer went first into the cab to do the preflight checks while the captain and the first officer continued to study the interaction of gross takeoff weight (GTOW), power settings and runway conditions, and to work out the details of the navigation. They were all in the simulator cab some 20 min before scheduled time of departure; as soon as they had taken their seats an experimenter connected their ECG leads to the preamplifiers stored in the pockets in the back of the seats. The instructor/simulator operator then took his position behind P1. The two observers then boarded together with the operator for the sequencing of the background tapes. The ECG's were checked on an oscilloscope and recording of them started at once. After the loudness levels of the various audio inputs had been checked and adjusted, they were connected to the recorder and the integrity of all the inputs was checked downstream from the recording heads of the multichannel tape recorder.

The three ECG and four audio channels were switched on and the recorder started as soon as all the crew had settled in their seats in the cab. The line printout of aircraft parameters was not started, however, until the application of power for takeoff. Continuing their standard preflight procedures, the crew interacted with the simulated air traffic controller and ground handler for push-back, taxi, takeoff, and departure clearances. Preparation for the flight and its entire conduct were carried out exactly as they would be in real life.

Preceding and during the simulated flight, the observers commented continuously on the operation of controls for airframe, engines, ancillary equipment, and fuel. Special notice was taken of errors in procedures (e.g., the use of checklists) and of specific errors (e.g., mistakes in setting up navigation and communication frequencies). Also reported was the time taken to locate information in documents in relation to their complexity, format, and available illumination.

At the end of the first sector the ECG leads were disconnected at the preamplifiers and the crew allowed to leave the cab to obtain light refreshment. Reentering the cab after a period of about 20 min, they carried out the same checks as they would on a short turnaround on the line. The crew then proceeded with the second sector in the same way as for the first, all navigation and communication being carried out according to published routing and ATC instructions as during actual flight operations.

At the end of the second sector, the crew members left the simulator and their electrodes were removed. They were then debriefed for their complaints and comments in the presence of the instructor/simulator operator and the observers. To avoid any possibility of confrontation or ill feeling no attempt was made to discuss performance, faults in flying technique or decisionmaking. After the aircrew left the experimenters and the instructor/simulator operator made a manuscript record of all the significant factors that they could remember relating to the performance of the crews in the two sectors. This was done as a precaution in case there had been any failures in the recording of observer comments or aircraft parameters.

DATA REDUCTION

Audio and ECG Channels

The four audio and three ECG channels contained on a single 1-in. tape were played back together. During the playback, each of the four audio channels was rerecorded on a separate single channel cassette recorder; all four were started simultaneously with the seven-channel recorder. The signals from the three ECG channels were fed to a Gould-Bio-Tach that converted the raw ECG into a beat-to-beat heart-rate signal. The three heart-rate signals were then recorded on a Brush seven-channel pen recorder.

The recording paper was run at a speed of 10 cm per min. Provision was made for switching to raw ECG data should this be required, though this was not done routinely. The pens were adjusted so that full scale deflection was from 0 to 200 beats per min. Simultaneously with the recording of heart rate, two experimenters listened to both the No. 1 observer's channel and that for the ATC communications. Both experimenters listened to one channel in each ear.

To economize in the reduction of less significant data, the transcription of the preflight heart rate and the audio channels was limited to 10 min preceding the application of power for takeoff.

Simultaneously with the recording of the heart rate, the experimenters transcribed other data onto the four remaining columns of the pen recorder paper. One of these was used for the aircraft parameters, (e.g., speeds, heights, passage of navigation fixes, ATC clearances, etc.), while the other three columns were used for the actions taken by each of the three crew members.

Most of the crew's actions could be divided into continuing and discrete ones. The continuing actions included studying maps and approach plates, looking for information in aircraft manuals, and conferring with another crew member. The discrete actions were exemplified by lever operation (e.g., gear handle, flaps, spoilers, etc.), the setting of values required for navigation and communication frequency selection, and headings and radials for the flight path control.

The rerecorded audio cassettes, dubbed from each of the four audio channels, were also available for transcription to allow a more detailed assessment of such things as the circumstances immediately preceding an incident or error; if these were thought to be significant, the cassettes were transcribed separately using the playback of an office dictating machine. When these pen recorded strip charts with their written comments were completed, each was examined for relevant data. Simultaneously, the printout of the aircraft parameters was placed in juxtaposition with the strip; thus, there was continuous information available throughout the flight that could be used to indicate arousal, vigilance, decision times, and errors, either those noted by observers or those deduced by examining the aircraft values. This task was performed separately for each of the 36 sectors and represents a total of 80 hr of recordings.

Aircraft Parameters

Identifying the application of power at takeoff in both recordings ensured that the lines of figures on the printout of the aircraft parameters could be accurately associated in time with the pen recorder strip. The printout had a time column that started again from zero every 10 min; using these, it was possible to write in the total elapsed time at each 5-min interval. Thus, each line of printout could be identified and related to the pen recorder strip chart to an accuracy of 5 sec for heights above 1800 ft and to 1 sec for heights below that. In this way aircraft responses could be matched with the action of the various crew members and their simultaneous heart rates.

RESULTS AND DISCUSSION

The immense amount of detailed data collected about each of the sectors makes it possible to study almost every aspect of each crew members' tasks and how he was affected by a wide variety of circumstances. For instance, it would be possible to work out the amount of time it took any of the subjects to locate a specific piece of information in aircraft manuals, charts or approach plates. It would be possible to apportion the time used for navigation, communication with air traffic control and how the need for this interfered with specific aspects of flight management and control. This study, however, was a specific attempt to investigate the kind and number of errors and how these related to overall workload and arousal. The data are now in an easily usable format and can be employed for any of these other purposes with an economical use of time. Nevertheless, at present only the data concerning errors, vigilance times, decision times, and heart rate have been worked up.

Errors

Information about the errors made by the crew members was obtained in several ways. Some came from the notes made by the instructor/operator and by the observers during their debriefing at the end of each session, while some crew members, although not questioned, reported errors that had not been noticed by the observers. Additional errors of all classes were subsequently found during examination of the transcript of the observers' comments onto the strip chart. Finally, there were errors, particularly those relating to skill and flying technique, that could be identified on the printout of aircraft parameters. The errors were arbitrarily classified as relating to: navigation, communications, systems operation, flying, tactical decisions, crew integration, flying skill, use of autopilot, and others.

The navigational errors included selecting the wrong frequency for the required NDB or VOR beacon, selecting the wrong radial or heading, failing to select a navigational facility at the correct time, and misreading charts. (See page 38, appendix A.)

The communication errors all concerned the use of radio telephone (R/T) for ATC purposes. They were: selecting the wrong frequency, not replying to a clear message, misunderstanding the message, use of the wrong call sign, and forgetting to comply with a message that a crew member had previously acknowledged.

The systems errors were mainly: mishandling of engines, hydraulic and fuel systems, misreading and missetting of instruments, and failure to use ice protection. (See page 40, appendix A.)

The majority of flying errors involved engine handling, neglect of speed limits, and flying at the wrong altitude. Sometimes P1 and P2 did not formally hand over the control of the aircraft, and these occurrences were included under this heading.

Most of the tactical decision errors seen in this scenario concerned the amount of fuel to be dumped to achieve a desired landing gross weight that would be compatible with the adverse runway conditions forced on the flightcrew for landing in the second sector. Errors were also made concerning the interaction between the landing gross weight, the flap setting, and the degree of automatic braking selected.

The errors of crew integration included episodes where P1, failing to realize that P2 or P3 was overloaded, asked them to retrieve information, further disrupting their performance. (See page 38, appendix A.) There were also failures to use the automatic pilot when this could have relieved the workload.

The errors relating to flying skill were those when the pilot flying was unable to control the aircraft to the parameters that he desired, as distinct from "flying errors" where the pilot was successfully controlling the aircraft to wrong parameters. These included: failing to maintain the desired speed by 20 knots with the aircraft clean or 10 knots in the approach configuration overbanking in a turn, failing to hold the desired altitude within 250 ft, engine handling with many large and abrupt changes of power, rapid changes of angle of attack, and general rough handling of flying controls.

Occasions when the appropriate automatic flight control was not used were classed as errors when, in the opinion of the observer, such use would have materially improved crew integration or reduced workload. There were also many occasions on which the actual autopilot controls were incorrectly sequenced so that the required mode did not function.

The classification "other" included a mixture of events such as not briefing passengers about turbulence, difficulty in organizing charts and approach plates, and problems with handling spectacles.

The numbers of the various classes of errors made in each sector are shown in table 3. The table shows that there were many more errors made in the high workload second sector, a ratio of about 2.5 to 1. This second sector, although about twice as long (mean 97 min vs 47 min), by design contained a period of some 20 min during the climb-out from New York in which only a low level of activity was required, hence errors would not be expected. This low workload period was needed to set the scene for the measurement of vigilance for the subsequent abnormalities.

In view of the foregoing, the numbers of errors were thought to be a better indication of the differences between the two sectors than error rate as a function of time.

The main difference between the number of errors made in the two sectors concerns tactical decisions, systems operation, and navigation. This is to be expected because the circumstances of the first sector required few decisions, whereas many had to be made in the second. Similarly, in the second sector the engine shutdown and consequent dumping procedures provided the opportunity for more errors in the operation of systems. The need to turn back in the second sector involved navigation that had not been planned before flight; this induced the known difficulties consequent on diversion to another airport.

TABLE 3.- NUMBER AND CATEGORY OF ERRORS FOR LOW AND HIGH WORKLOAD SECTORS

Run	Navigation		Communication		Systems operation		Flying		Tactical decisions		Crew integration		Flying skill		Autopilot		Other		Total		Total for run
	Sector																				
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	
1		4	1			2	1	1	1	7							1		4	14	18
2	1	3		1		3				9								1	1	17	18
3		2		4		1	2	1	1	4			2		1				6	12	18
4		1		1				3	1	8		1	1	5	1	1	1	1	4	22	26
5	1	1		1	1	4		3	3	5					1	2	1	3	7	20	27
6	1	3	5	1	6	1		3	1	4	1		3	3	2			1	19	16	35
7		2	1	1		5	2			5	1	2	1	3	1	2		1	6	21	27
8	3	3	1	1	1	3	2			4			1	1				2	8	14	22
9		1	1	1	1	2	1	3	1	4			1	2	2			1	7	15	22
10	1	1			1		1	2	1	4				1	1	1			5	9	14
11		2		1	1	5		2	1	6		1	3	3					5	20	25
12	2	1				3	3		1	4				2		2		2	6	14	20
13	1	3	4	1	1	4	1	2	1	5		1	2	3	2			1	12	20	32
14	1	1	3	3		5		3		9					1	2	1	1	6	24	30
15						1	2	2	1	5					1			1	4	9	13
16		4	2	2			2	3	2	4		1		3		1		2	6	20	26
17	1	3		1	1	7		1	2	8		3		5				3	4	31	35
18	2	5		3	1	4	3	3	2	10		2	5	1	1		2	1	16	30	46
Totals	14	40	18	22	14	50	20	32	19	105	2	11	19	32	14	11	6	21	126	324	450

The Significance of the Errors

Most of the errors significant for safe flight or for airline economics can be found under the headings of navigation, systems operation flying, and tactical decisions. All the errors in these categories were referred to a senior member of the training staff of the participating airline. Those that he expected to cause comment in the training situation are listed below:

Navigation

First Sectors

Omitted turning point.

VOR/ADF switch set to ADF at takeoff.

Both the VOR/ADF switches set to ADF for 21 min after takeoff.

Tuned wrong outer marker at JFK.

Used wrong VOR to identify Swanpoint intersection.

Most fixes and turning points missed by 10 mi.

Second Sectors

Wrong frequency selected for Bridgeport VOR.

Unable to locate the VOR on the cluttered chart. Had to go to approach plate.

Missed Cambridge VOR by 18 mi. Also made a serious heading mistake due to an error in subtraction.

Used the 061 radial of Bridgeport VOR instead of Kennedy VOR.

Forgot to hold at Bohemia.

Commenced climb before clearance from ATC.

Went through Bohemia hold by 4 min (24 mi).

Went through the Bohemia hold by 23 mi.

Inserted the wrong way-point in inertial navigational system when calculating distance and deduced a wrong position.

P1 misunderstood the standard instrument departure.

Systems Operation

First Sectors

No ice protection used at takeoff, therefore malfunction of EPR gauges; P3 throttled back without referring to other instruments. Aircraft slowed to stick-shake speed.

P2 used the wrong flight director computer.

Ice protection not used for takeoff and therefore EPR gauges malfunctioned. P3 throttled back.

P2 used the wrong air data computer for flight director.

Altimeter setting error of 150 ft.

Second Sectors

Engine was overboosted to 1.89 EPR.

Engines overboosted (EGT warning lights).

P3 did not see the hydraulic leak for 5 min and then only when the "empty" light went on.

P1 had the storm lights on and P3 had the warning lights set on dim. Consequently the engine oil low pressure lights were not seen.

P3 only dumped the fuel from the inboard tanks and prolonged the dump unnecessarily.

One engine overboosted several times to 1.54 EPR.

Incorrect fuel dump resulting in unbalanced main tanks.

P3 made an error of 100,000 lb when dumping fuel.

Engines overboosted several times, max EPR 2.07.

Engines overboosted to 1.51 EPR. Engine not throttled back for 5 min after oil clog noticed.

P2 crossed over air data computers to try to solve a flight director discrepancy. This was illogical.

Flying Errors

First Sectors

Activated ground proximity warning system twice during approach.

Indicated airspeed 326 knots at 9,000 ft. (250 knot limit below 10,000 ft).

Second Sectors

Slow in the holding pattern with an angle of attack of 6.7° with the aircraft clean; indicated airspeed was 15 knots below minimum for this configuration.

Indicated airspeed 20 knots below book speed for height and all up weight.

Continued the dump in a 45° banked turn.

500 ft below the cleared altitude of 4,000 ft.

Flew at Mach number 0.86, which equals the limiting V_{mo}.

Lost 1,000 ft after shutting down No. 2 engine. Also had indicated airspeed of 215 knots with the aircraft clean during the hold at Bohemia (20 knots low).

Misunderstood the departure clearance. Exceeded cleared altitude of 2,500 ft.

Tactical Decisions

First Sectors

P1 failed to take over during unstable approach.

Second Sectors

Dumping started very late, 24 min after engine shutdown.

Correct landing gross weight 564,000 lb. Seven crews decided 585,000 lb.

P1 selected maximum automatic braking.

Decided 30 flap for approach without checking T page (by two crews).

Decision to return to JFK made without considering other nearer alternates by nine of the eighteen crews.

P1 accepted altitude change to 35,000 ft by ATC. Impossible at that weight.

The scenarios of both the low and high workload sectors induced errors that were comparable in number and kind to those which had been previously observed in flight. The average number of total errors made in the high workload sector was comparable to those reported for a European airline in high workload sectors in 1971 (see table 4). This similarity supports the credibility of the scenarios and simulation techniques used in this study.

Usually errors that persisted long enough to produce serious effects were no different from those that were picked up as soon as they were made. This persistence highlights the importance of effective monitoring by all crew members. Because of these considerations, it seemed worthwhile to try to find out how various classes of error came about. There was a great amount of detailed information available for all the circumstances preceding the significant errors, and an attempt was made to relate some of the more significant errors to the circumstances that preceded them.

Where possible, it was thought advisable to reexamine the flight deck circumstances for a period of some 15 min before the error was made and to continue this until the error was noted and corrected or was no longer of any significance. Thus the length of postevent study was variable. All the different kinds of records were used; the ATC communication and the cockpit voice recordings were located and transcribed, as were the observer's comments. This procedure was needed because in some cases all the available information had not been noted during the initial simultaneous transcription

TABLE 4.- ERROR RATES IN HIGH WORKLOAD

Sector	Runs/ Flights	Mean errors	Standard deviation	Avg. flight time (min)	Error rate/min
NASA simulation					
JFK and return to JFK	18	18.2	6.2	97	0.19
European Shorthaul Airline					
London Glasgow	17	18	9.2	70	0.26
Glasgow London	18	19	8.8	72	0.26
London Frankfurt	6	21	7.5	79	0.27
Frankfurt London	5	21	4.65	85	0.25

that has been previously described. The computer printout for the aircraft parameters was also examined in detail for the same period.

From the various sources, it was possible accurately to reconstruct the flight path to show the effect of navigational errors, to quantify the erosion of safety margins in such things as angle of attack and speed limits for aircraft configurations, as well as to verify the violations of air traffic control speed limits and the overstressing of power plants. In some cases, therefore, it was possible not only to see how the errors came to be made but also how they affected the conduct of the flight.

In many cases, especially where the errors were large, the observers were able to recall many of the significant circumstances. An example of this was an occasion when P1 and P2 were discussing the location of a navigational facility and spread out charts so that not only was their attention concentrated on this problem, but also the chart in use was placed so that it physically obscured the P1s flight instrument panel. Almost certainly this was the reason the VOR to which they were homing was overflown by many miles.

The most significant errors concerned the landing gross weight (LGW), navigation, and engine handling. The LGW is important because, if it is too high it can affect the maneuver margins on the approach and also the ability to stop safely in the available runway length with the prevailing wind and surface conditions. On the other hand, if too much fuel has been dumped, the choice of alternate airfields may be limited.

Navigational errors have special importance in areas with heavy traffic such as the New York Metroplex that forms a large part of the area in which the significant parts of this scenario take place. On three occasions the inbound leg of the hold at Bohemia was prolonged, once for 23 mi. Under actual flight operations, an error of this magnitude would be unlikely because it would almost certainly be seen on the radar by the relevant air traffic controller.

Engine handling can affect safety, for example, if power is reduced immediately after takeoff. This happened on three occasions because of faulty indications due to icing of the PR probes. Furthermore, overboosting of engines may limit their life or lead to subsequent failures. Several incidents of this kind were seen.

Examples of workups of each of these categories of error are shown in appendix A. Although the number of errors whose origins and effects have been reconstructed in this way is too small for statistical purposes, it does show how increased use might be made of the data available in crash-survivable recorders for the investigation of accidents. These data can be used to improve the reconstruction of cockpit activities during a simulation of the circumstances preceding an accident. The use of current line crews provided with the flight plan and other paper work for the trip and the inclusion of the ATC communication and navigation workload might increase the validity of the investigation so that the causes of obscure accidents might be brought to light more easily.

Relevance of Error Recording for Training

The kind of scenario and recording techniques used in this study demonstrated to the volunteer aircrews and training personnel how easy it is for errors to be made in high workload situations. This has implications for training. Many of the discrete errors and wrong decisions were related to overloading one particular crew member, particularly when he was engaged in reciting and complying with checklists for the procedures connected with abnormal operation. It was also seen how in some cases compliance with these procedures could interfere with the monitoring cover built into standard operating procedures.

The realism of the circumstances, coupled with ability to refer instantly to the printout of the aircraft parameters at the end of a run, seemed to have a marked impact on the participating crew. This effect might be heightened if recordings of the cockpit conversation and communication with ATC could also be replayed during retrospective assessments.

Vigilance

Vigilance is an important constituent of airmanship and the scenario was designed so that the crew's vigilance could be assessed. By means of the transcripts of the observer's comments and the "cockpit voice" channels that were written on the strip chart, it was possible to measure the time taken to notice the rise in filter pressure on the oil pressure gauge or the subsequent oil pressure warning light for No. 2 engine, following the simulation of the filter clog. The time taken to notice the depletion of the fluid in No. 3 hydraulic system was measured in the same way. With care, these times could be assessed to ±1 sec.

The times for the two situations show a large difference between the shortest and longest (table 5). Except for Run 10, in which the audio recording malfunctioned, the time for the light to be seen by the crew varied from 3 to 155 sec. P1 was first to notice this light on 13 out of the 17 runs for which this information was available, probably because the warning panel is near P1's flight instruments.

The time taken to notice the depletion of the hydraulic fluid in No. 3 system was also variable; the hydraulic light was first seen by P3 on 13 of the 17 runs, by P1 on 3 and once by P1 and P3 simultaneously. This can be expected because the hydraulic gauge is toward the aft end of the systems monitoring panel, forming part of the P3's work station. It cannot be seen by P2 and only with difficulty by P1.

An attempt was made to relate the time taken to notice the problems with the preceding state of arousal indicated by the percentage of rise in heart-rate of each crew member or to the sum of the increases of all three. No such relationship could be found (appendix B).

Decision Times

Because of the clogged oil filter for No. 2 engine, many decisions had to be made by the captain. Five representative examples that could easily be timed were selected for measurement. They were: (a) to shut down the engine, (b) to return to Kennedy, (c) to dump fuel, (d) to start the dump, and (e) to make the turn onto the first navigation facility on the way back. All the times were taken from when the oil pressure warning light was first noticed by any member of the crew. The intervals were obtained from the recordings of the cockpit voice and communications and where relevant, were verified by reference to the printout of aircraft parameters (table 6).

There is a wide variability in the time and order to make decisions (table 7). The difference in sequence may reflect the mental processes used by the individual captains to assess the priorities of the situation.

An attempt was made to associate the times for the first three of the decisions with arousal as indicated by heart rate. These readings were taken from the readout of heart rate on the strip chart 20 sec prior to the time the decisions were recorded. These data were converted to percentage rise over the lowest recorded reading for that subject. A statistical study of these results (appendix B) revealed no relationship except that decision time was affected by whether P1 was "flying" or not "flying," and this parameter always influences heart rate.

TABLE 5.- TIME TAKEN TO NOTICE PROBLEMS

Run	Time to notice oil pressure light, sec	Seen by	Time to notice hydraulic depletion before and after warning light, sec		Seen by
			Before	After	
1	17	P1	---(No report)	---	P3
2	3	P1	---(No report)	---	P3
3	---	P1	-30		P3, P1 (same time)
4	55	P1		30	P3
5	10	P1	-40		P3
6	22	P3		6	P1
7	35	P1		15	P1
8	7	P1		12	P3
9	20	P3		67	P3
10			---(No audio)	---	
11	2	P1	---(No audio)	---	P3
12	15	P3	-30		P3
13	96	P1		75	P3
14	30	P1		130	P1
15	5	P1		10	P3
16	5	P1		50	P3
17	5	P1		45	P3
18	155	P3	---	---	P3
	n = 16 Mean 30.0 (sec) Range 2 to 155	P1 × 13 P2 × 0 P3 × 4	n = 13 Mean 26.2 (sec) Range -40 to 130		P1 × 3 P2 × 0 P3 × 13 (P1 + P3) × 1

TABLE 6.- DECISION TIMES AFTER OIL "P" DIFFERENTIAL SEEN IN SECOND SECTORS

(In seconds)

Run	Shut down engine	Return NYC	Decide dump	Start dump	Return Providence
1	255	275	562	782	417
2	201	255	257	980	415
3	124	80	311	469	174
4	94	309	315	1032	272
5	115	135	295	608	222
6	161	212	400	947	291
7	180	255	610	1164	262
8	214	405	255	879	575
9	184	206	324	504	244
10	505	1025	835	1015	445
11	285	270	365	603	301
12	135	260	685	1031	270
13	272	301	531	876	341
14	163	213	568	750	230
15	138	134	134	832	156
16	153	210	161	1566	285
17	326	383	453	981	366
18	265	100	534	914	245
Mean	209	279	422	885	306

TABLE 7.- CLOGGED OIL FILTER, NO. 2 ENGINE — ORDER OF DECISIONS

	Shut down engine	Return to NYC	Decide to dump fuel	Start the dump	Turn to Providence VOR
Run	A	B	C	D	E
1, 5, 6, 7, 9 12, 13, 14	1	2	4	5	3
3 and 11	2	1	4	5	3
8 and 16	1	3	2	5	4
2	1	2	3	5	4
4	1	4	3	5	2
10	2	5	3	4	1
15	2	1	3	5	4
17	1	3	4	5	2
18	3	1	4	5	2

Heart Rate

Sampling— The heart-rate data were sampled visually from the pen recorder tracings every 100 sec for each of the crew members throughout both sectors of every run. In the uneventful first segments, these readings were averaged for the whole flight except for the 2-min periods immediately following takeoff and preceding landings. In the high workload second sectors, the readings were averaged for that portion of the flight prior to the oil filter clog of No. 2 engine and again for that portion of the flight following the clog.

Spot checks of heart rates were also taken from the tracings 20 sec prior to the decisions to shut down the engine, to return to New York, and to dump fuel. The highest values of heart rate immediately following takeoff and during the approach were found for both sectors of each run, as were the highest values immediately after landing.

Significance of the heart rate— To compare heart rates among subjects, it was decided that all the heart-rate values should be expressed as a percentage rise over the lowest recorded during their "preflight" time in the simulator. The average percentage rises in heart rate for all of the subjects in every phase of the 18 runs are shown in figure 2. There is little difference between the mean percentage rise during takeoff for the low and high workload sectors; for approach and landing, however, the mean rise in the second sector is higher; this is almost certainly due to the need for a manual, instrument, three-engine approach in a strong crosswind. It will also be seen that in the second sector, during the uneventful climbout before the onset of problems, the percentage rise in the heart rate is lower than at any other time. These results are also similar to those that have been seen in flight during other studies.

Associations were sought between the mean percentage rise in heart rate for P1, P2 and P3 individually as well as for the crew as a whole. A complete list of the calculations made with regard to possible associations between heart rate, errors, vigilance and decision times is shown in appendix B.

There was a strong association between an increase in the heart rate of both P1 and P2 when they were the "flying" pilot for that sector. This was true for all phases of takeoff, cruise, approach, and landing (fig. 2). This finding was comparable with observations made in flight by the author and others during scheduled airline operation (refs. 1 and 2).

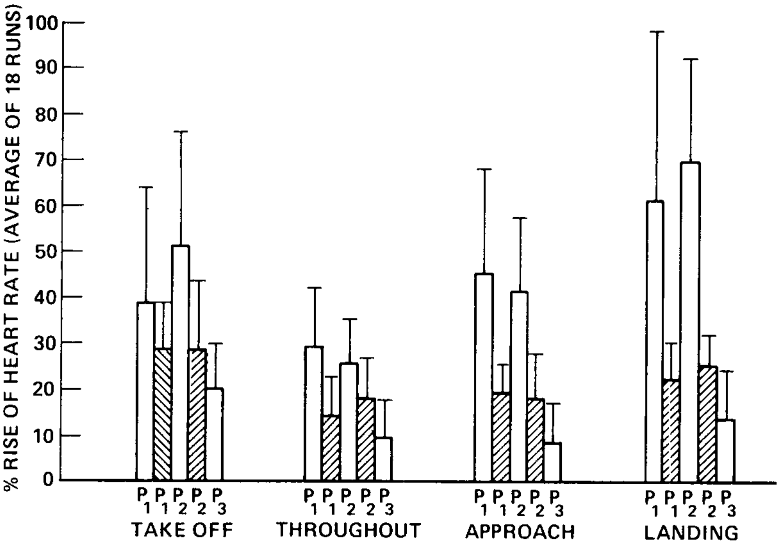
The lack of effect of arousal, as indicated by heart rate, on vigilance or the number of errors was unexpected. This finding may support the contention that heart rate by itself should not be used as a measure of workload in complex situations.

Performance

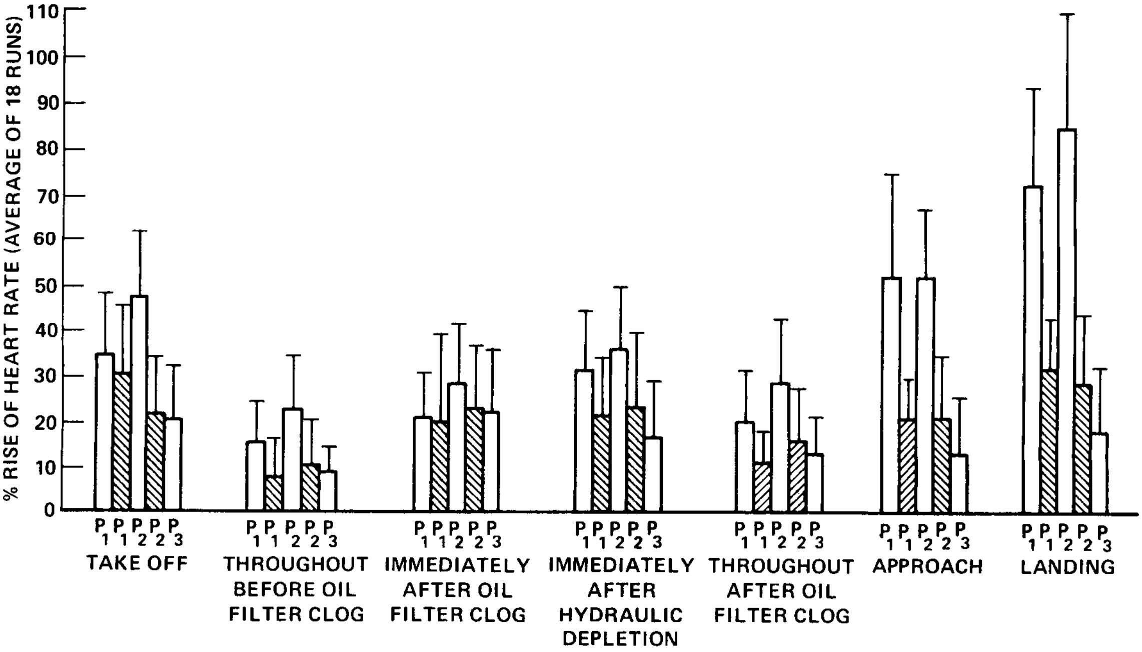
The wide variety of aircraft parameters available on the printout provides an opportunity for measuring the flying performance of the pilots in relation to flying skill, navigation, and compliance with air traffic control instructions. When the experiment was first set up it was considered desirable that provision be made to measure deviations in azimuth and from the glide slope while pilots performed a manual instrument approach with reference to ILS. This indicator of skill was chosen because it had previously been used by Billings et al. (refs. 3 and 4) to show the effects of various stresses (e.g., high blood alcohol) on the ability of pilots to perform their tasks.

Information is available on the printouts for an assessment of this kind to be made, but it was decided not to do this because the data are not available on magnetic tape in digital form and the printout figures would require extensive manipulation before meaningful scores could be worked out. Furthermore, it had been noticed during the simulator runs that both P1's and P2's were able to carry out precise approaches on the ILS and that their performance was not at all affected by the foregoing stress and high workload. Nevertheless, the data are available should verification of the author's opinion be needed in the future.

The impression was gained that performance was adversely affected by defects of memory. Some captains had difficulty with short-term memory for



(a) Sector 1.



(b) Sector 2.

Figure 2.- Percentage rise in heart rate versus lowest recorded preflight (P1 and P2 flying and not flying — hatched bars denote "not flying").

clearances and other ATC messages. Those individuals who recognized that they had this problem wrote down all the significant figures and locations before acknowledging the transmission.

Long-term memory problems were also encountered. For instance, several crew members had difficulty in recalling information gained during training about aircraft systems. This problem was more prevalent in those who had been flying the particular aircraft type for a long time.

Resource Management

The need for extremely competent airline captains is recognized by the rigor of their selection that includes a requirement for long experience in commercial flying. They are rewarded by this position of great authority. The main requirements for successful captaincy have, however, not been spelled out nor is this the place for such a complete list. For the first time this experiment provided opportunities to observe 18 different captains as they responded to the demands of the same abnormal flight conditions. There seemed to be large variations in respect to leadership, resource management and decisionmaking. Though these attributes are difficult to quantify, it was possible to observe and record differences between crews with respect to all of them.

In some of the runs leadership quality seemed lacking and on occasion P2 appeared to have usurped the role of P1.

Resource management relates to both the human and material adjuncts available to the captain. This ability is more easily quantified; for instance, it is possible to observe the ways in which information for solving the problems posed by the scenario is obtained. Large differences were seen, varying from the meticulous confirmation of remembered information by reference to documents, to the use of preconceived values that were not checked. In some runs this kind of behavior led to the need for late reappraisal and changes in decisions, thus prolonging flight time and increasing workload at critical stages of flight.

Sometimes, while the crew was discussing the best course of action among themselves, it was difficult to determine whether P1 or P2 was in control of the aircraft even at times when no autopilot was available for use. This behavior was most often noticed when either P1 or P2 was searching either a chart or approach plate for navigational fixes and the associated radio frequencies and sought help in this from the flying pilot.

Another facet of poor utilization of human resources was the failure to anticipate the overloading of a crew member by certain combinations of circumstances. An example of this was the lack of recognition by some P1's of the implications of the depletion of the fluid in No. 3 hydraulic system. With this they lost the remaining autopilot. This increased the workload in a way some P1's did not realize, especially when they were the pilot "flying" for that sector and when some of them continued to hand-fly the aircraft themselves as well as to make decisions. When P1's were aware of the situation and delegated their P2's to controlling the aircraft in its flight path, immediate benefits were evident. Great differences were also evident in the way in which decisions were arrived at by P1's following the clog of the oil filter of No. 2 engine because then he was able to give full attention to assimilating the information from documents, ATC, and other crew members and to use these data to make unhurried decisions.

Some of the captains appreciated the abnormal situation quickly, and within seconds had realized the need to shut down the engine, which would require permission from ATC for a lower altitude. They immediately realized the need to land at the nearest suitable airfield and began to assemble the information about weather and runway conditions that would enable them to work out the correct landing gross weight and how much fuel they needed to dump.

On the other hand, there were occasions when the decision to shut down the engine was delayed for several minutes and after this, vain attempts were made to maintain altitude at the expense of speed or overboosting the remaining engines. In some cases the landing gross weight was subsumed without reference to information concerning the particular landing runway.

Interrelationships

Many of the data were examined statistically in an attempt to find relationships between various sets of results. A detailed account of this study is in appendix B. In summary, the following sets of data were examined:

Those concerned with changes in the condition of subjects in relation to the time and place of the experiment:

1. Time of day of simulation session versus total errors in Sectors 1 and 2.

2. Hours out of bed before start of simulator run versus total errors in Sectors 1 and 2.

3. Days since last flight versus total errors in Sectors 1 and 2.

Those concerned with effect of age of pilots:

1. Age versus percentage rise of heart rate at takeoff for P1 and P2 when flying and not flying for Sectors 1 and 2.

2. Age versus percentage rise of heart rate at approach; P1 and P2, flying and not flying for Sectors 1 and 2.

3. Age versus percentage rise of heart rate during landing; P1 and P2, flying and not flying for Sectors 1 and 2.

Those concerned with effects of arousal as indicated by percentage rise of heart rate in relation to:

1. Number of errors versus percentage rise in heart rate throughout Sectors 1 and 2.

2. Time taken to notice clogging of oil filter versus rise in mean heart rate for the previous part of Sector 2.

3. Time to notice depletion of No. 3 hydraulic system versus percentage rise in mean heart rate for the period immediately after the clogged oil filter in Sector No. 2.

4. Time to make decisions versus percentage rise in heart rate after noticing oil pressure light for No. 2 engine. (The decisions being (1) to shut down No. 2 engine, (2) to return to New York, and (3) to decide to dump fuel.)

The most significant relationship found for any of the results was that concerning the percentage rise in heart rate for the pilots at the controls, which was always greater than that for the pilot not at the controls. This held good for all phases of flight in both the low and high workload sectors.

The percentage rise in heart rate of the P3 during the period in Sector 2 prior to the onset of the engine problem also seemed related to errors made in the second sector, the number of errors increasing with the rise in heart rate.

There is also a relationship between the number of errors made in the first sector and the number of days since the last flight of P1 and P3; this does not hold true for P2. There is less correlation between the number of errors made in the second sector and the number of days since the last flight of P3. It might be speculated that a short recency is more important to the P1's and P3's, who were older than the P2's. The effect of the number of days since last flight by the P1 was not influenced by whether he was the flying or nonflying pilot.

The time for the first decision (i.e., to shut down the engine in Sector 2) was regressed on the percentage rise in heart rate of P1 20 sec before the decision both when flying and not flying. The decision took longer if P1 was the flying pilot, but heart rate had no association with the decision time. No such relationship was found for either of the subsequent decisions.

It was also found that the time taken to notice the oil clog predicts the number of errors in the high workload sectors. This behavioral measure of vigilance seems to be a better indicator of performance than does heart rate.

Ergonomics

Although this study was not directed towards the more mechanistic facets of flight deck operation, it was inevitable that these should be noticed if there were deficiencies that impinged on the performance of the flight crews. Several examples of poor design in regard to ergonomics were seen concerning equipment, documents, and illumination.

Flight deck layout and instrumentation– Lack of attention to human factors was evident in the layout of both the pilot's and the engineer's instrument panels in the association between controls and indicators and in the excessive reach needed to operate some controls. The position of the central warning and caution panel and its possible effect on response times has already been alluded to. Deficiencies were also seen in the human factor aspects of controllers for radio, communication and navigation equipment. Examples of these were seen in the position and format of the readouts and controls for the frequency selection and the position and format of keyboards and readouts for inertial navigation.

Important defects related to control of engine power. In the forward half of their travel, the four power levers (throttles) are too far away and too much offset from the midline of pilots with shorter arms, thus making precise differential control of power difficult for some. Furthermore, the "feel" of the levers is such that, as they approach their forward position, small movements produce comparatively large changes of power.

The indication most used for setting the amount of power is the engine pressure ratio (EPR). In the flight deck the gauges measuring the EPR have small, complex dials 2 in. in diameter. There are two three-digit counters, one for setting the "bug" or "lubber mark" on the scale for the calculated maximum for the prevailing flight conditions and another showing the instantaneous EPR values. There is a pointer and scale showing the same values. The figures on the counters are small and difficult to see, especially by many captains who are pres-byopic; this difficulty is accentuated in low levels of illumination. Though the pointer is bold and its angular position can be seen easily, the scale is short, due to the small diameter and because the pointer operates over no more than about three-fifths of the circumference. The figures associated with the scale marks are even smaller than those in the digital readouts.

The exhaust gas temperature (EGT) indicators were also difficult to read because of short scales and pointers. Tunnelling by the bezel rings obstructed the scales of the hydraulic fluid contents gauges when viewed at an angle, and this may have contributed to the failure of many P3's to notice the depletion of the contents.

The airspeed indicator (ASI) fitted to this model has a single pointer and a single turn scale showing indicated airspeed (IAS). On the same dial there is a digital readout for computed airspeed (CAS).[1] The scale for IAS is expanded in that portion used for takeoff and approach but compressed between 200 and 300 knots. In this range the scale is short and for accurate information the gaze must be switched to the readout of the CAS because the change in the angle of the pointer is not a sufficient cue. The instances of high and low IAS noticed in this range may be due in part to these difficulties. The movable "lubber marks" or "bugs" fitted to the periphery of the bezel rings of the ASI that are used for reference speeds for takeoff and landing have excessive parallax.

Other indicators might be found wanting if different scenarios related to them as specifically as the tasks in this one did to ASI, EPR, EGT, hydraulic quantity and oil pressure gauges.

The central warning panel was positioned to the left of the pilot's engine instrument panel near its lower edge. It was not readily seen by P2 and was often obscured from P1's gaze by any unfolded charts he was using for navigation. The warning panel was similarly difficult for P3 to see because it would be obscured by P1's body and throttle arm.

The combination of awkward controls and poor instrumentation almost certainly contributed to the rough handling of power that was seen on several occasions. During actual flight operations, handling of power can affect passenger comfort and peace of mind and in some circumstances may contribute to a reduction of engine life consequent upon overboosting. In some approach conditions, especially when there is a severe windshear, defective power handling may affect flight safety.

Illumination— Lighting in this flight deck was inadequate for several tasks. Visibility of the overhead panels was particularly poor for the P3's who often had to use their flashlights. The illumination provided for the reading of charts and other documents was poor and the overhead lights seemed inadequate; on several occasions the P1's were seen to use the thunderstorm lighting for the forward instrument panel to read charts. To do so, they had to lean forward, holding the chart in front of the control column to catch the light coming from under the glareshield, thus obstructing their view of much of the panel.

Documentation— A very large number of documents is required for the conduct of a civil transport flight. These documents refer to several classes of information and may be separated according to their permanence. There are four loose leaf volumes of operating manuals, two relating to the aircraft and two to the general operating policy of the airline; changes in these documents are relatively infrequent. Next in permanence are the charts and approach plates contained in the route and airport manual. These are stowed in loose leaf binders from which the relevant sheets are removed for a particular stage

---

[1] CAS is IAS less the static source position error and will read 0-5 knots higher than the airspeed pointer. The difference will be small at low altitudes and at speeds less than 200 knots.

of the flight and are updated at monthly intervals. The third class of documents, which relates only to the particular flight in hand, consists of the printouts of flight plans, copies of weight and balance sheets, and weather maps.

The main difficulty with the aircraft and operations manuals lies in attempting rapid location of information that does not seem logical in its layout. For example, the tables about stall speeds for different gross weights and flap settings are listed under the main heading of "Minimum Equipment," although this title is merely the first item in a list that is covered by the title "Limitations." The format of the table for stall speed is itself confusing; the values are modified by factors contained in a footnote that is not in a logical order.

Not only were there problems in the nomenclature of the indices, but often, when several items were needed at the same time, they could not easily be cross-referenced. On several occasions P3's were seen to be keeping their fingers between as many as three separate pages so as to avoid repeated reference to the index. There is also a problem in stowing these rather bulky volumes so that they are easily accessible.

In order to combat some of these difficulties, several crew members had quick reference cards they had made up themselves. Although this may overcome some of the deficiencies of the manuals, it introduces another hazard in the use of nonstandard information that may be inaccurate and out of date.

The route and airport manual is too bulky to use without removing the sheets needed for the trip from the cover, and stowing these sheets then presents a problem. Many ways of organizing these flimsy sheets were seen, perhaps the best consisted of plastic envelopes bound together. In this way, the required sheets were easy to hold and see. The worst situation noticed was that generated by a P1 during the return to New York: he had 10 of these flimsy sheets in his hands while he tried to associate the runway dimensions and conditions with the optimum landing gross weight and weather limits for different diversion airfields.

The weight and balance sheet is well planned and easily understood. The copies of the flight plan and the weather maps, however, seem to be on poor quality paper and their low color contrast makes them difficult to use.

Without the aircraft and operations manuals, the amount of paperwork needed for this two-sector scenario (Dulles, JFK and London (LHR) has a single side area of some 20 m^2 (fig. 3)).

CONCLUSIONS AND RECOMMENDATIONS

The observers, the training captain coordinator, and the participating aircrew were all convinced of the realism of the simulation of the scenario. Similar effects to those of actual flight operations were produced by high workload, unfamiliar navigation and unfamiliar air traffic control routing.

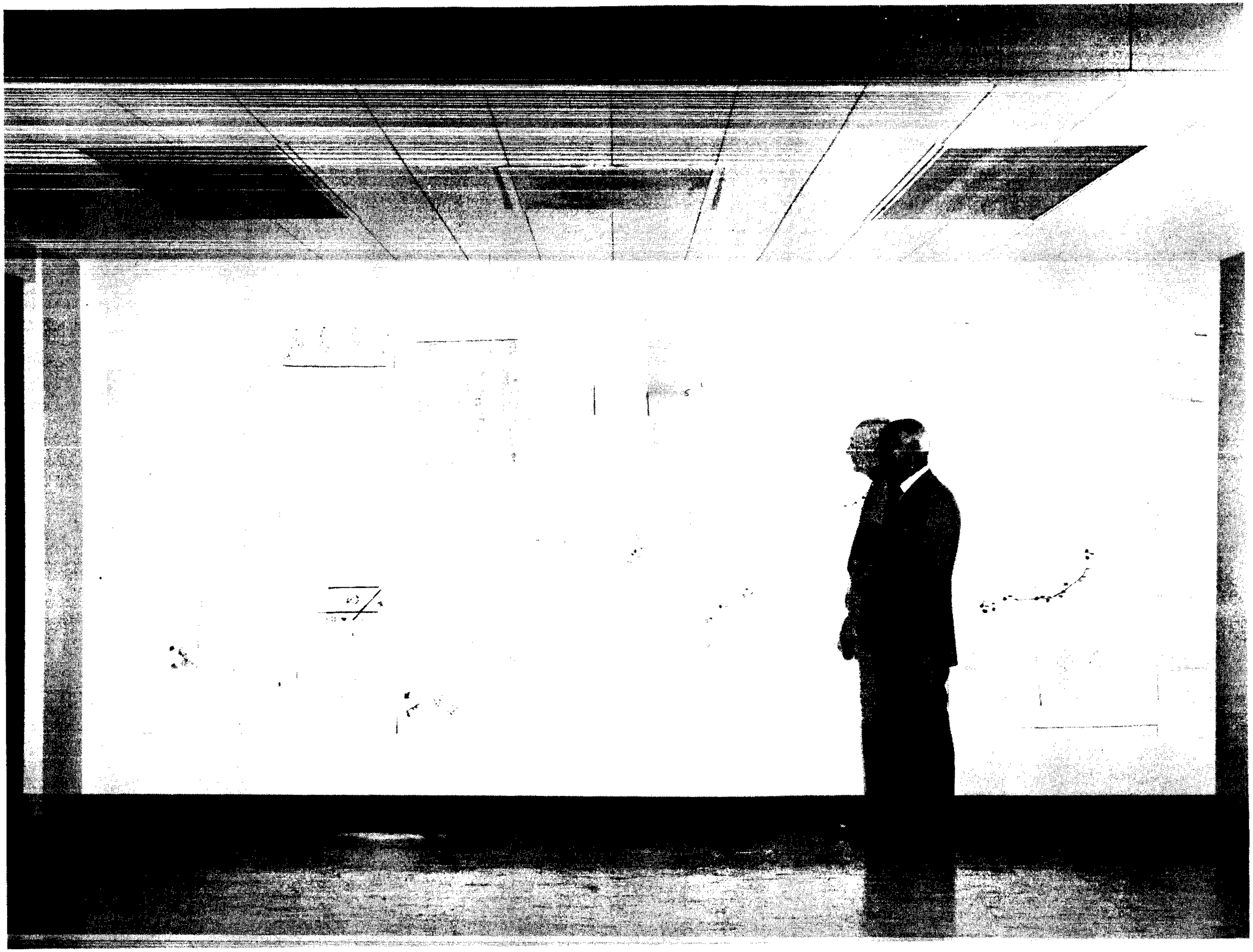


Figure 3.- Documents other than manuals required for conducting the scenario.

Arousal, as measured by heart rate, mirrored observations made in flight and the rates and types of errors were also similar.

The results indicate that high workload can lead to decreased performance of flight crews. This decrease is manifested by errors in the operation of systems and mistakes in navigation that are associated with prolonged response times to aircraft abnormalities.

Some of the difficulties are induced by deficiencies in the design of flight decks and instrumentation, others by those of documents and charts. Many of the problems, however, relate to the management of human and mechanical resources. The variability between crews in reacting to the same problems suggests that those who perform less well might be helped by special training.

Consequently, it is recommended that aircraft and equipment manufacturers be reminded of the importance of designing flight decks and instrumentation to the best human factors practice, that increased effort should be given to redesigning documents and charts, and that special training in resource management and captaincy be developed and validated. Such training should include the use of full mission simulation of scenarios that are representative of actual situations. Special emphasis should be given to those situations where rapid decisions and safe solutions for operating problems are required.

The techniques developed to produce realistic situations for the full mission simulation achieved in this study might also be used in the investigation of aircraft accidents. Reenactment of the flight, with all known factors included and with as much fidelity as possible, might help to establish the chain of events preceding the final critical error. (The reenactment should be done by linecrews with equivalent experience.) Such an exercise revealed significant factors that helped to elucidate the accident involving British European Airways Trident 1 after takeoff from London in 1972 (ref. 5).

The same techniques might be beneficial in developing and validating standard operating procedures to achieve optimum integration of flightcrews and to avoid conflicting instructions and activities.

Ames Research Center
National Aeronautics and Space Administration
Moffett Field, Calif., 94035, August 24, 1978

APPENDIX A

NARRATIVES OF SOME TYPICAL ERRORS

This appendix consists of descriptions of certain specific errors that were chosen for further investigation because of their importance. They relate to landing gross weight, navigation and engine handling.

Example of Error in Landing Gross Weight

The following is a description of the circumstances that led to P3 dumping 77,000 lb too little fuel and how this mistake was allowed to persist. The landing was made at a gross weight in excess of that required for the available stopping distance.

After the decision to dump fuel had been made and without consultation of documents relating to the length and conditions of the runway in use, the figure of 84,000 lb, the accepted maximum weight for any runway, was arrived at by P1 and P2 without discussion.

P3 then calculated a dump time of 4 min 30 sec; this was accepted by P1 without comment, although it is approximately one-third of the time needed. Without prompting, P3 then recalculated the dump required and arrived at a time of 12 min for the dump.

Instead of dumping enough fuel for this landing weight to be achieved, having made proper allowance for the enroute burn, P3 then only ran the dump for 3 min, perhaps because he reverted to his original erroneous figure or because he misread the gross weight indicator. Unsatisfied, he again started to recalculate but was interrupted by the depletion of No. 3 hydraulic system.

During the next 8 min P3 was subjected to a high workload but then noticed that the gross weight was much too high and decided to refigure the fuel. During that time he was subjected to further interruptions and did nothing more about the fuel until P1 noticed that the gross weight indicator read 647,000 lb and decided to make an over gross weight landing. A minute and a half later, P3 rechecked the fuel as part of the landing check list and became concerned about the gross weight. He spent a minute and a half rechecking calculations and announced that the aircraft gross weight computer must be in error. Two min later the simulator was "landed" at 172 knots with only 25° flap with 1,000 ft/min descent at about 77,000 lb over the correct weight.

During the 32 min between the time the decision was made to dump fuel and the landing, there were 15 interruptions to P3's specific tasks concerning tailoring the amount of fuel to be dumped in relation to the conditions and length of the landing runway. These interruptions consisted of:

P1 requests

- Three engine drift down speed?
- How long to dump?
- How is dump going? (during check list before start of dump)
- Head wind component for landing?
- Three engine cruise speed?

Check lists instigated by P1

- Completing engine shutdown check
- One engine inoperative check list (started twice because of interruptions)

Other

- Pressurization control — cabin pressure altimeter
- Steward's list of passengers needing reticketing for transmission to Ground Operations (three requests and one compliance)
- Hydraulic quantity depletion warning light

In this way P3 was never able to complete and verify his fuel calculations and dump times before he was interrupted, either as a routine part of standard operating procedures (SOP's) or by a request from P1 or the senior steward. P3 thus became overloaded and his work fragmented. P1 failed to recognize the situation and so did nothing to resolve it (see page 15).

In the face of later evidence, the error seemed due to a persistence of an original misconception. Rationalization took place to the extent that the gross weight computer was considered at fault rather than P3's calculations. During this time the heart rate of P3 was increased by 25%; this indicated considerable arousal that may not have been optimal for the task in hand.

Examples of Navigational Error

Examination of the data reveals how, in three runs, P1 failed to start the right-hand holding pattern at Bohemia on the way back to New York. This should have been done at a point 10 mi short of Deer Park VOR (see page 14).

> Departing Riverhead VOR, with P1 flying the aircraft on very nearly the correct heading and having apparently misread the chart, they were intending to enter the hold at 10 mi by distance measuring equipment (DME) past Riverhead rather than the same distance before Deer Park, the next way point.
>
> Departing Riverhead 77 min after takeoff, P2 communicated with Air Traffic Control (ATC) and obtained clearance to descend to 15,000 ft. P1 notified ATC 15 sec later that he was on the radial from Riverhead for the holding pattern, reduced power, and started to lose height. During this time P2 was studying charts and P3 was reading the "Hydraulic Abnormal" check list for the approach.
>
> At 79 min, the distance from Riverhead had become 17 mi by DME, and P1 started to discuss the hydraulic problem with P2; he

had apparently forgotten the ATC instruction to hold or was experiencing a time-compression effect due to diversion of his attention.

At 79 min 30 sec, P1 remarked that they were now 22 instead of 10 mi past Riverhead, but took no action. He continued discussing with P2 the effects of the wetness of the landing runway. He decided there was need to dump more fuel to achieve a landing gross weight of 564,000 lb, and P3 started to make the required fuel calculations.

At 81 min 30 sec, and 33 mi by DME from Riverhead, the situation was resolved by the simulated ATC querying the position of the aircraft and giving radar vectors.

Similar errors in over-running the hold were made in two other runs. In one, immediately preceding the error, P1 and P2 were discussing in the same way the length of the landing runway and the fuel dump required.

On departing Riverhead, P1 asked P3 to read the "One Engine Out" approach check list. ATC then called to change frequency to Kennedy approach. This instruction was not understood the first time and ATC repeated it. P1 complied and Kennedy approach control requested a change of transponder code and identification. At 68.05 P1 took control of the aircraft back from P2, and they set the approach and landing speeds and critical heights by moving the "bugs" or "lubber marks" on the airspeed indicators, pressure, and radio altimeters.

They departed Riverhead, having decided that the hold at Bohemia was 10 mi distant.

P3 indicated that they had too high a gross weight for the landing runway, and P1 and P2 discussed the appropriate flap setting. During this time they flew for 4 min before P2 noticed that they were through the hold and set the reverse heading of 084. Thus, they went 24 instead of 10 mi before commencing the hold.

In another run, following an animated discussion between P1 and P2 concerning the conditions that had to be met to use the short, wet, out-of-wind runway, P1 was the "flying" pilot and continued to fly the aircraft without giving it to P2 for a minute.

P1 and P2 decided to start the hold 10 mi by DME past Riverhead. However, P1 continued for 7 min at about 5 mi per min before turning to reverse the hold.

During this time, P1 was having difficulty in hand-flying the aircraft accurately. While P1 and P2 discussed the approach speed and flap settings for the intended landing runway, both height and speed were affected. Indicated airspeed was 20 knots less than the recommended minimum without flaps at the current gross weight, almost at the stick-shake condition. Throttle handling by P1 was rough, with large power changes — straight from idle to 1.25 EPR. Later EPR's were up to 1.44 or above for 50 sec. Vertical speed varied from -759 to +829 ft/sec.

This was a typical demonstration of the effect of a complex mental task on manual skill. P1 could have resolved his difficulties by requiring P2 to fly the aircraft during the hold, thus allowing P1 to marshal his thoughts.

Examples of Errors in Engine Handling

The important errors in engine handling were of two kinds: overboosting of the remaining engine after No. 2 had been shut down and the misinterpretation of fluctuations of the EPR caused by icing of the pressure probes in the engine nacelles (see page 14).

Examples of errors involving overboosting— There were five examples of this type of error:

1. After shutting down the engine, P1, who was the flying pilot, overboosted the remaining engines to 1.89 EPR, presumably in an attempt to maintain height and speed. P2 was engaged in communication with ATC and P3 was running the "dump" check list. Thus effective monitoring was absent. It seems that this error might have been prevented had P1 handed over the flying of the aircraft to P2 with the start of the engine problem so that he could have given his attention to decisionmaking rather than to continuing to fly.

2. A similar incident occurred when P1 was the flying pilot. The remaining engines were overboosted to 1.68 EPR about 10 min after No. 2 had been shut down. During this time P1 had been discussing with P2 the implications of the diversion to the short runway at JFK, arguing about the power required to "drift down," obtaining further weather reports, and discussing landing gross weight and flap settings.

   P3 had also been busy during this period calculating the amount of fuel to dump and running the dump check list. All three crew members had their attention diverted, and the monitoring of the aircraft parameters lapsed. It seems likely that this error would have been prevented had P1 given the flying of the aircraft to P2 as soon as the engine problem became evident.

3. A comparable incident occurred in yet another run when P1 was the flying pilot. Some 4 min after No. 2 engine had been shut down, P1 wrongly decided it was possible to maintain 31,000 ft at the current gross weight and proceeded to overboost the remaining engines. At this time P2 was distracted by communications with ATC. P3, however, noticed the high EPR values and remonstrated with P1.

4. On this occasion, P2 was the flying pilot and immediately the No. 2 engine was shut down, P1 took the aircraft from P2 and without consulting documents, and after obtaining clearance from ATC, decided to return to JFK. Next, he turned on the autopilot and advanced the throttles to 1.83 EPR, which activated the red lights on the exhaust gas temperature (EGT) gauges. After some 10 sec, P2 noticed the EGT's and throttled back. P1 and P3 did nothing during this time.

5. In this example P1 was the flying pilot and maintained control of the aircraft during all the discussions of weather, alternate airfields, and landing gross weight. Then, during descent, about 15 min after the engine problem was noticed, the autopilot warning horn came on as the hydraulic system was depleted. This produced a "startle" reaction by P1, who advanced the throttles enough to overboost the remaining engines to 1.88 EPR and activate the EGT warning lights.

It is difficult to think of a logical reason for this action. However, as he was descending, P1 may have had a "set" to regard any warning as related to ground proximity unless it was proved otherwise.

Examples of errors involving inopportune reduction of power— There were two examples of this type or error:

1. Icing protection for engine nacelles is not activated prior to takeoff although reported conditions warranted this. P2 was the flying pilot. Three minutes after the power was applied for takeoff, engine icing was simulated. The EPR gauges then fluctuated while all other engine instruments showed normal readings. P3 pulled back the throttle levers during a climbing turn; the speed dropped to 140 knots and the stick shaker began to operate. P3 pushed the throttles forward then brought them back. P1 then reapplied takeoff power.

The original error of failing to activate engine nacelle ice protection may be because it only occurs once in the pre-taxi check list and in no other prior to take off. The error made by P3 in pulling back the throttles when the EPR gauges fluctuated may be due to his failure to scan more than one indicator of power during the takeoff and initial climb rather than to lack of vigilance.

This incident could indicate that most P3's are overconscious of problems of engine life rather than being fully integrated with the total aircraft situation.

2. With P1 as the flying pilot, icing protection for engine nacelle was not actuated prior to takeoff as in example No. 1. Three minutes after takeoff power was applied, engine icing was simulated, leading to fluctuating EPR gauges. P3 pulled back the throttle levers, P2 almost simultaneously activated engine nacelle ice protection, then P3 reapplied power.

In the second example the lapse was not as serious as in the first because the engine icing was recognized more quickly and power immediately restored. The initial error, however, was probably due to failure to comply with the pre-taxi check list. Although it did not result in action to reduce power, there was another incident when engine nacelle ice protection was omitted during the pre-taxi check, leading to similar malfunction of the EPR gauges 2 min after takeoff, but the problem was quickly diagnosed and ice protection turned on. In this event, only a small change in airspeed resulted.

That the same initial error was made by 3 crews out of 18 may be due to ice protection not being repeated as part of the pre-takeoff check list.

APPENDIX B

STATISTICAL STUDY

James Stevenson

INTRODUCTION

During the course of an experiment undertaken to assess the effects of workload on the performance of civil transport aircrews, the opportunity was taken to measure a variety of parameters. These included errors, vigilance and decision times. The heart rates of the three crew members were also recorded continuously.

An attempt was made to see if any of the recorded parameters were in any way related to each other or to other factors known about the crew, such as their age, experience, and category, and whether they were the flying pilot for any particular sector.

The calculations made in these respects are as follows.

Calculations

1. Relationship of heart rate (HR), expressed as percentage rise over lowest recorded "preflight" to factors relating to crew members and phases of flight.

Percentage rise in heart rate regressed on age of P1 and P2 flying and not flying.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
First Sectors					
During takeoff					
Mean age of P1 + P2	49.6	7.0	0.122	-0.744	0.462
Flying or not	---	---	.344	2.090	.044
Rise in HR, %	36.5	22.3			
During approach					
Mean age of P1 + P2	49.6	7.0	.071	.525	.603
Flying or not	---	---	.660	4.858	.000
Rise in HR, %	31.6	19.4			

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
On landing					
Mean age of P1 + P2	49.5	7.1	0.071	-1.277	0.211
Flying	---	---	.660	3.8	.001
Rise in HR, %	45.0	30.2			
Second Sectors					
During takeoff					
Mean age of P1 + P2	49.3	7.0	.059	.366	.717
Flying	---	---	.489	3.036	.005
Rise in HR, %	33.4	17.6			
During approach					
Mean age of P1 + P2	49.8	6.9	.044	.333	.741
Flying	---	---	.690	5.191	.000
Rise in HR, %	35.0	22.3			
On landing					
Mean age of P1 + P2	49.7	6.9	-.089	-.748	.461
Flying	---	---	.814	6.980	.000
Rise in HR, %	53.2	30.7			

2. Relationship of lowest recorded heart-rate preflight to age (P1, P2, and P3).

Lowest preflight raw heart rate regressed on age.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
First Sectors					
Mean age of P1 + P2 + P3	49.6	7.00		0.052	0.959
HR	76.7	11.7			

3. Relationship of heart rate expressed as percentage rise over lowest recorded preflight to age of P1 and P2 while flying.

Percentage rise of heart rate while flying regressed on age (P1 and P2).

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
<u>First Sectors</u>					
<u>During approach</u>					
Mean age of P1 + P2	48.1	7.3	0.133	0.537	0.599
Rise in HR, %	43.9	19.8			
<u>On landing</u>					
Mean age of P1 + P2	47.7	7.3	.071	.277	.79
Rise in HR, %	66.8	29.4			

Percentage rise of heart rate while flying regressed on age (P1 and P2).

<u>Second Sectors</u>					
<u>During approach</u>					
Mean age of P1 + P2	51.6	6.7	0.133	0.26	0.789
Rise in HR, %	50.9	20.1			
<u>On landing</u>					
Mean age of P1 + P2	51.6	6.7	-.290	-1.174	.259
Rise in HR, %	76.3	23.4			

Note: Because of the lack of effect of age on the rise in heart rate, either for the preflight, approach, or landing condition, this variable (age) has not been used in subsequent calculations.

4. Relationship between the time taken to notice the clogged oil filter in the second sectors and the rise in heart rate of the flying and not flying pilot.

Time to notice oil clog regressed on percentage rise in heart rate.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
Rise in HR, %	9.7	8.5	-0.410	-1.31	0.22
Flying or not	---	---	.122	.391	.705
Time	22.5	28.3			

There is no correlation of the time taken to notice the engine oil filter clog with either the rise in heart rate over the lowest preflight value of the pilots or whether P1 or P2 was the flying pilot in any particular second sector.

5. Relationship between the time taken to notice the depletion of No. 3 hydraulic system and percentage rise in heart rate of those 10 P3's who were first to notice the onset of the problem.

Time to notice hydraulic depletion regressed on percentage rise in HR.

P3's	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
Rise in HR, %	9.5	6.9	0.279	0.82	0.44
Time	28.9	28.7			

Thus, the heart rate of these pilots was not associated with their vigilance for this problem.

6. Relationship between the heart rate of those P3's who saw the hydraulic problem first and those who did not.

P3's	Mean	Standard deviation	T
Saw	9.5	7.0	1.20
Did not see	13.6	7.7	

7. Relationship between the number of errors made in first and second sectors and the percentage rise in heart rate over resting levels of P1, P2, and P3, collectively and individually.

Total error for sector regressed on percentage rise of heart rate for period prior to clogging of oil filer.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
Sector 1					
Mean rise in HR of P1 + P2 + P3, %	17.7	5.6	-0.142	-0.574	0.574
Errors	7.0	4.4			
Sector 2					
Before oil clog					
Mean rise in HR of P1 + P2 + P3, %	11.5	4.6	.026	.106	.917
Errors	18.2	6.2			
After oil clog					
Mean rise in HR of P1 + P2 + P3, %	15.4	5.6	.199	.812	.429
Errors	18.2	6.2			

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
P1					
Sector 1					
Rise in HR, %	19.7	11.8	-0.136	-0.549	0.591
Error	7.0	4.4			
Sector 2					
Before oil clog					
Rise in HR, %	11.3	9.4	-.360	-1.541	.143
Error	18.2	6.2			
After oil clog					
Rise in HR, %	14.6	10.4	-.192	.781	.445
Error	18.2	6.2			
P2					
Sector 1					
Rise in HR, %	22.9	9.6	-.275	-1.146	.269
Error	7.0	4.4			
Sector 2					
Before oil clog					
Rise in HR, %	15.3	11.6	.092	.367	.717
Error	18.2	6.2			
After oil clog					
Rise in HR, %	20.3	13.7	.313	1.316	.206
Error	18.2	6.2			
P3					
Sector 1					
Rise in HR, %	10.5	6.1	.302	1.267	.223
Error	7.0	4.4			
Sector 2					
Before oil clog					
Rise in HR, %	7.7	5.3	.508	2.36	.031
Error	18.2	6.2			

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
P3 (Concluded)					
Sector 2 (Concluded)					
After oil clog					
Rise in HR, %	11.3	7.4	0.148	0.589	0.566
Error	18.2	6.2			

The only significant relationship between percentage rise in HR and errors seems to be for the heart rate of P3's in the first part of the second sectors and the total errors in them. This may be caused by those P3's who showed the higher heart rates as being the ones who were aroused by a situation that was already stressing their capacity before the onset of the increased workload due to subsequent problems.

8. Relationship between the number of errors in each sector of the runs and the number of days since the crew member's previous flight, using the sum of the days for P1, P2, and P3, as well as individually.

All errors for first and second sectors regressed on sum of days since last flight for P1, P2, and P3.

Errors regressed on days since last flight	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
Sector 1					
Days	12.1	11.7	0.442	1.97	0.066
Errors	7.0	4.4			
Sector 2					
Days	12.1	11.7	.428	1.9	.076
Errors	18.2	6.2			

There is a weak relationship in both the first and the second sectors.

All errors for first and second sectors regressed on day's last flight for P1, P2, and P3, individually.

Errors regressed on days since last flight	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
P1					
Sector 1					
Days	8.8	9.7	0.495	2.279	0.037
Errors	7.0	4.4			
Sector 2					
Days	8.8	9.7	.249	1.030	.318
Errors	18.2	6.2			
P2					
Sector 1					
Days	13.9	19.7	.140	.568	.578
Errors	7.0	4.4			
Sector 2					
Days	13.9	18.7	.257	1.064	.303
Errors	18.2	6.2			
P3					
Sector 1					
Days	13.5	18.9	.429	1.899	.076
Errors	7.0	4.4			
Sector 2					
Days	13.5	18.9	.415	1.822	.087
Errors	18.2	6.2			

The results of these calculations show that there is a strong association between the number of days since the last flight by P1's and the number of errors that were recorded for the first sector they operated.

There is also a less strong association between the number of days since the last flight of the respective P3 and the number of errors seen in both the first and second sectors.

9. Relationship between months flying by P1, P2, and P3 in this model aircraft and the number of errors recorded for the first and second segments.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
Sector 1					
Mean months of P1 + P2 + P3	30.8	14.7	-0.118	-0.426	0.641
Errors	7.0	4.4			
Sector 2					
Mean months of P1 + P2 + P3	30.8	14.7	.043	.173	.864
Errors	18.2	6.2			
P1					
Sector 1					
Mean months	33.9	23.6	-.107	-.429	.674
Errors	7.0	4.4			
Sector 2					
Mean months	33.9	23.6	.267	1.110	.283
Errors	18.2	6.2			
P2					
Sector 1					
Mean months	27.7	18.9	-.050	-.202	.843
Errors	7.0	4.4			
Sector 2					
Mean months	27.7	18.9	-.266	-1.105	.285
Errors	18.2	6.2			
P3					
Sector 1					
Mean months	52.3	18.7	.203	.831	.418
Errors	7.0	4.4			
Sector 2					
Mean months	52.3	18.7	-.109	-.440	.665
Errors	18.2	6.2			

10. Relationship between the number of errors recorded for the first and second sectors and the mean of the number of hours out of bed prior to starting the simulator runs for P1, P2, and P3.

Total errors for Sector 1 and Sector 2 regressed on hours out of bed for P1, P2, and P3.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
Sector 1					
Mean hours out of bed for P1 + P2 + P3	7.7	4.1	-0.075	-0.302	0.766
Errors	7.0	4.4			
Sector 2					
Mean hours out of bed for P1 + P2 + P3	7.7	4.1	-.320	-1.352	.195
Errors	18.2	6.2			

11. Relationship between the number of errors recorded for the first and second sectors and the time of day the crew reported for the simulation run. These times were 06.00, 10.15, 15.00, 19.15 hr and are numbered one to four.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
Sector 1					
Simulator session	2.8	1.2	-0.161	-0.653	0.523
Errors	7.0	4.4			
Sector 2					
Simulation session	2.8	1.2	-.251	-1.037	.315
Errors	18.2	6.2			

12. Relationship between the time taken to decide to shut down the engine and the heart rate taken 20 sec before the decision expressed as the percentage rise over the lowest preflight and whether this time was influenced by P1 being the "flying" or "nonflying" pilot.

	Mean	Standard deviation	Standard regression coefficient	T	P (2 tail)
P1					
Flying or not	---	---	0.596	2.566	0.023
Percentage rise	14.8	14.2	-.185	-.799	.439
Decision time in sec	187.5	66.9			

This shows a strong effect. It takes longer for the P1 to make this decision if he is the "flying pilot." However, it seems that heart rate is not correlated with decision time.

13. Relationship between vigilance and decision times and the number of errors.

Errors in the second sectors regressed on (1) time to notice the oil clog, (2) time to shut down the engine, (3) time to decide to return to New York, and (4) time to decide to dump fuel.

	Mean	Standard regression coefficient	T	P (2 tail)
1. Time to notice oil clog	30.12	0.43	1.517	0.157
2. Time to shut down engine	196.3	.29	1.076	.305
3. Time to decide to return to New York City	250.19	.10	.374	.715
4. Time to decide to dump	403.06	.06	.233	.820
	Multiple "R"	0.5987		

Because of the small sample size, multiple regression was performed on the two most important variables.

Errors in second sectors regressed on (1) time to notice the oil clog and (2) time to shut down the engine.

	Mean	Standard regression coefficient	T	P (2 tail)
Time to notice oil clog	30.12	0.42	1.812	0.093
Time to shut down engine	196.31	.33	1.434	.175
	Multiple "R"	0.59		

The time taken to notice the oil clog is a weakly significant prediction of total errors in Sector 2. The two variables, time to notice and time to shut down, explain about one-third of the variance.

Errors in second sectors regressed on the number of errors in the first sectors that preceded them.

First Sector	Mean	Standard coefficient	T	P (2 tail)
Errors	7.00	0.208	0.849	0.408

This is not significant. Thus, errors in the first sector of a run do not predict the errors in the second sector.

RESULTS

There was always a strong relationship between increased heart rate (HR) and being the "flying" pilot at all stages of flight. There was no relationship between age and heart rate in either the less stressed preflight condition or during landings which showed the largest increases in heart rate.

No effect was seen between the time taken to notice the engine oil pressure problem and the heart rate of the crew in the immediately preceding period.

The 10 P3's who noticed the depletion of the hydraulic fluid before the P1's or P2's did not have a significantly different increase in HR compared with the 8 who did not.

The rise in HR of the P3's during the early part of the second sector was associated with increased error rate for those sectors, but no other association between HR and errors was found.

The number of errors made in the first sectors seemed associated with the number of days since P1's had flown but not for second sectors. There was also a less strong association between the errors in the first and second segments and the number of days since the P3's last flight.

There was no relationship between the number of months the crew members had been flying this type of aircraft, nor the number of hours out of bed prior to commencing the simulator runs, nor with the time of day the run began.

A relationship was shown between whether the P1 was the "flying" pilot or the "nonflying" pilot and the time taken to make the decision to shut down the engine.

The number of errors in the low workload sector did not predict the number of errors in the following high workload sector; they are, however, weakly predicted by the vigilance measure of time to notice the clogged oil filter.

This finding contrasts with the lack of effect of arousal as indicated by the rise in heart rate of P1 and P2 on the number of errors in different runs. Although the heart rate of the P3's is associated with the number of errors in the second sectors, it is interesting that the behavioral measure of vigilance seems to be a better overall predictor of errors than is heart rate.

REFERENCES

1. Howitt, J. S.: Flight-Deck Workload Studies in a Civil Transport Aircraft. AGARD Conference Proceedings, 56, Dec. 1969.

2. Ruffell Smith, H. P.: Heart Rate of Pilots Flying Aircraft on Scheduled Airline Routes. Aerosp. Med., vol. 38, 1967, pp. 1117-1119.

3. Billings, Charles E.; Eggspuechler, Jack J.; Gerke, Ralph J.; and Chase, Robert C.: Studies of Pilot Performance: II. Evaluation of Performance During Low Altitude Flight in Helicopters. Aerosp. Med., vol. 39, 1968, pp. 19-31.

4. Billings, C. E.; Wick, R. L.; Gerke, R. J.; and Chase, R. C.: Effects of Ethyl Alcohol on Pilot Performance. Aerosp. Med., vol. 44, 1973, pp. 379-382.

5. Ruffell Smith, H. P.: Pilots' Activities Immediately Preceding a Fatal Accident. Annual Science Meeting, Aerosp. Med. Assn. Conference Proceedings, Washington, D. C., May 6-9, 1974, pp. 83-84.