

NOV 29 1984

COMPONENT PART NOTICE

THIS PAPER IS A COMPONENT PART OF THE FOLLOWING COMPILATION REPORT:

A

(TITLE): Flight Test Techniques. Proceedings of the Flight Mechanics Panel  
Symposium Held in Lisbon, Portugal on 2-5 April 1984.

(SOURCE): Advisory Group for Aerospace Research and Development, Neuilly-sur-Seine  
(France).

TO ORDER THE COMPLETE COMPILATION REPORT USE AD-A147 625.

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDINGS, ANNALS, SYMPOSIA, ETC. HOWEVER, THE COMPONENT SHOULD BE CONSIDERED WITHIN THE CONTEXT OF THE OVERALL COMPILATION REPORT AND NOT AS A STAND-ALONE TECHNICAL REPORT.

THE FOLLOWING COMPONENT PART NUMBERS COMPRISE THE COMPILATION REPORT:

AD#:	TITLE:
AD-P004 098	Determination of External Store Drag.
AD-P004 099	The Flight Test of an Automatic Spin Prevention System.
AD-P004 100	High Angle of Attack Test and Evaluation Techniques for the 1980s.
AD-P004 101	Determination of Performance and Stability Characteristics from Dynamic Manoeuvres with a Transport Aircraft Using Parameter Identification Techniques.
AD-P004 102	Application of Advanced Parameter Identification Methods for Flight Flutter Data Analysis with Comparisons to Current Techniques.
AD-P004 103	A Technique to Determine Lift and Drag Polars in Flight.
AD-P004 104	Flight Testing a Digital Flight Control System: Issues and Results.
AD-P004 105	Aircraft Tests in the Clearance Program for the Use of a Combat Aircraft from a Runway After Damage Repair.
AD-P004 106	The Handling and Performance Trials Needed to Clear an Aircraft to Act as a Receiver During Air-To-Air Refuelling.
AD-P004 107	Reference Systems for the Evaluation of Dead-Reckoning Navigation Equipment.
AD-P004 108	Autopilot Performance Evaluation: Tornado Experience and Future Applications.
AD-P004 109	Assessing Pilot Workloads in Flight.
AD-P004 110	Simulation Applied to the Avionics System Testing in the F/A-18.
AD-P004 111	Flight Test Techniques Employed in the Nimrod MR MK 2 Weapon System Performance Trials.
AD-P004 112	F-16 and A-10 Diffraction Optics Head-Up Display (HUD) Flight Test Evaluation.
AD-P004 113	In-Flight Accuracy and Coverage Tests of ESM and ECM Systems.
AD-P004 114	Flight Test Techniques Used for Proof of Structural Integrity of Tornado When Carrying External Stores.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P004 115	Real Time Testing - The Next Generation.
AD-P004 116	Using an Airborne CO <sub>2</sub> CW Laser for Free Stream Airspeed and Windshear Measurements.
AD-P004 117	NAVSTAR GPS Applications to Test and Training.
AD-P004 118	An Extended Real-Time Microwave Airplane Position System.
AD-P004 119	General Integrated Multipurpose Inflight Calibration System.
AD-P004 120	The Development of an Airborne Instrumentation Computer System for Flight Test.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

This document has been approved for public release and sale; its distribution is unlimited.

## ASSESSING PILOT WORKLOAD IN FLIGHT

by

Alan H. Roscoe  
 Royal Aircraft Establishment  
 Clapham  
 Bedford MK41 6AE

AD-P004 109

## SUMMARY

The importance of being able to assess pilot workload in real flight is generally acknowledged by people concerned with cockpit design and operational efficiency and safety. Currently, the most used and probably most reliable methods of estimating levels of workload in flight are those based on some form of subjective reporting by experienced test pilots. But subjective opinions are susceptible to bias and to pre-conceived ideas and so the use of a second and more objective measure to augment these opinions would seem to offer distinct advantages.

This paper describes the way in which a pilot's heart rate can be recorded to support, or occasionally question, his subjective rating of workload. A small number of examples from RAE Bedford trials are presented to illustrate the technique, and a short description is given of the BAe 146 Crew Complement Certification exercise. Finally, a current flight experiment to compare heart rate levels and workload ratings in a more scientific manner is described. The rationale for using heart rate in this way is discussed briefly.

## INTRODUCTION

Whether one is attempting to optimise workload levels on the flight deck of a civil airliner to improve safety, or to reduce workload in the cockpit of a combat aircraft to improve mission effectiveness, it is important to be able to assess workload in flight.

First of all in any paper devoted to assessing pilot workload it is clearly desirable to define what is meant by the term, for even a brief survey of the literature highlights the confusion that exists. Although there are several definitions they tend to be vague and to vary according to the disciplines and interests of their authors. It is possible, however, to place most of these definitions into one of two broad conceptual groups, those that are related to the demands of the flight tasks - input load, and those that are associated with the response to those demands - operator effort. Of course, there is a fundamental difference between these two groups each of which has its devotees. A confliction like this is difficult, if not impossible, to resolve and so it is not surprising that there is no general agreement on what is meant by the term pilot workload.

For the sake of completeness it is worth pointing out that a small number of people would argue for a third conceptual group containing interpretations of workload based on work results - or performance. Although it is an important attribute of workload it is suggested that performance *per se* should not be considered as the basis for working definitions. Nevertheless, whatever interpretation is used when assessing workload it is still essential to define and monitor performance.

Designers of aircraft cockpits and flight decks find it convenient to think of workload in terms of the task. In this way they can predict levels of workload for different flight tasks and operational situations by using data derived from various types of task and time analysis in mock-ups and simulators. (1), (2) (3). Levels or indices of theoretical workload (perhaps more appropriately termed 'task load'), usually expressed as a function of time required and time available, are extremely valuable in the design stage but, eventually, levels of actual workload will need to be assessed in real flight.

In this paper workload is considered to be related to effort - an interpretation which appears to be consistent with the views of more than 80% of professional pilots (4). It is also an interpretation that agrees well with the influence on the piloting task of such individual factors as natural ability, training and experience, physical fitness, age, and the individual response to stress. Cooper and Harper (5), in the introduction to their handling qualities rating scale, suggested a definition of pilot workload which, in a slightly modified form, seems to be most appropriate: Pilot workload is the integrated mental and physical effort required to satisfy the perceived demands of a specified flight task.

At present the most used and probably the most reliable methods for assessing workload in flight are based on some form of subjective reporting by experienced test pilots. Unfortunately, subjective opinions are susceptible to bias and pre-conceived ideas and so there are clear theoretical advantages in using more objective techniques to assess pilot workload. Among the many techniques studied are those based on measuring physiological variables; but the method of choice must be non-intrusive as well as being compatible with flight safety. Heart rate, which satisfies these criteria, is relatively simple to record and the discrete nature of the signal allows various forms of analysis to be carried out with ease. Moreover, the use of a physiological variable to assess workload fits in well with a concept that allows for the individual nature of piloting skill.

Heart rate has been recorded in flight extensively at the Royal Aircraft Establishment at Bedford over a period of some 15 years. After a three year exploratory study, during which test pilots were monitored while flying a wide variety of aircraft types (6), it was decided to record heart rate routinely to support pilot opinion during those flight trials where assessing workload is important.

This paper describes examples taken from a number of flight trials where this method of assessing workload has proved to be of some value

#### METHODOLOGY

The technique used at Bedford for recording heart rate in flight is based on the electrocardiograph (ECG). Amplified ECG signals, detected by means of two disposable electrodes applied to the pilot's chest, are recorded in analogue form on magnetic tape along with speech and various aircraft parameters. In the first instance heart rate is plotted out in beat-to-beat form together with the ECG 'R' wave - the basic signal, (Fig 6 is an example); subsequently mean rates for consecutive 30 sec epochs are plotted against time, these are found to be most useful (Fig 2 is an example). Mean heart rate values for particular flight manoeuvres, tasks, or sub-tasks may be calculated as required.

Initially pilot opinions of workload were given in a relatively unstructured descriptive manner but the need for some form of rating scale soon became obvious. After much trial and error, and with the help of numerous comments and criticisms from test pilots, a ten-point scale using the concept of spare capacity was developed (Fig 1). The overall design is based on the Cooper-Harper Handling Qualities Rating Scale (5), familiar to Bedford test pilots and sometimes used previously, though mistakenly, to rate workload (7).

The scale is not linear and probably lacks sensitivity at the lower end but is readily accepted by most pilots who have found it easy to use without the need to always refer to the decision tree. Half ratings are allowed within each decision branch and tend to be used frequently; originally it was decided not to permit the use of half ratings between the decision branches but the occasional difficulty of deciding between the last two branches, (in effect between ratings 3 and 4) was resolved by accepting a rating of  $3\frac{1}{2}$ . Pilots seem to find it much easier to rate a flight task if it is short and well defined.

Most of the flight trials of interest from the point of view of workload have involved the take-off or the approach and landing, two tasks that are well defined and where performance can usually be monitored by on-board instrumentation and by airfield-sited kinetheodolites or radar.

Heart rates and opinion ratings indicate only relative differences in workload so that it is helpful to have some form of datum for purposes of comparison. Although it is not always possible to compare heart rate responses for different experimental variables during the same sortie, or even under similar flight conditions, the advantages of doing so are obvious.

The individual nature of heart rate responses makes it necessary, especially when dealing with small numbers of pilots, for each pilot to be considered as his own control. This restriction, together with the high cost of operating research aircraft, usually makes it impossible to obtain enough data for worthwhile statistical analysis. Nevertheless, obvious trends in heart rate changes together with pilot ratings can provide valuable and reliable indications of differences in workload levels.

#### EXAMPLES

The brief examples presented in this section have been selected from different flight trials to demonstrate the practical use of recording pilots' heart rates as a means of augmenting their subjective assessments of workload.

Heart rate was first recorded routinely during a series of flight trials to evaluate different types of reduced-noise landing approaches (8), (9). The first two trials used a HS 748 Andover twin turbo-prop transport to compare simple-segment steep approaches with gradients of  $6^\circ$ ,  $7\frac{1}{2}^\circ$  and  $9^\circ$ , and, later two-segment approaches of  $7\frac{1}{2}^\circ$  changing to  $3^\circ$  at 200 ft, with conventional  $3^\circ$  approaches. There was generally good agreement between the project pilots' heart rates and their subjective estimates of workload; and in the case of the single-segment steep approach also with expected levels of difficulty - the workload being expected to increase with steeper approach paths and higher rates of descent. Figs 2 and 3 are examples of overall mean heart rate plots (different pilots). Interestingly, despite their relatively low heart rate responses the project pilots initially rated the workload for the two-segment approaches as high. It later transpired that these two pilots had instinctively disliked the idea of changing from a steep gradient - with the higher rate of descent - to a normal gradient at 200 ft. After their first sortie they modified their views and then consistently rated the  $7\frac{1}{2}^\circ/3^\circ$  approaches as being as easy as, if not easier than, the normal  $3^\circ$  approaches. This example highlights the fact that subjective assessments may be biased by allowing instincts and misconceptions to influence judgement. It also illustrates the advantage of using heart rate to augment - or sometimes to question - subjective assessments of workload.

Similar heart rate responses during a later trial using a BAC VC-10 four-jet transport to evaluate  $5^\circ/3^\circ$  two-segment approaches (with a transition height of 500 ft)

confirmed pilots' subjective assessments that the experimental approach profile generated similar levels of workload to the conventional 3° approach (Fig 4).

Pilots' heart rates were recorded during a flight test programme to assess the benefits of direct lift control (DLC) fitted to a BAC 1-11(10). Fig 5 shows overall mean rates for one of the three project pilots recorded during 3° and 6° approaches flown with and without DLC. It can be seen that the only appreciable reduction in heart rate with DLC occurred during the glide slope acquisition and early part of the 6° approaches. But glide slope performance was significantly better during all DLC approaches and all three pilots reported lower workload with DLC. Although heart rate did not appear to discriminate between the two experimental conditions, nor did it agree entirely with pilot opinion, it is worth noting that performance improved which suggests that pilots increased their effort without being aware of it.

These examples are typical of flight trials in which workload levels for different degrees of task difficulty can be compared with some forms of datum or with each other. Perhaps a well designed rating scale would have proved useful for assessing workload in these instances but where direct comparison is not possible such a scale becomes almost essential.

The final version of the Bedford rating scale was introduced during the HS Harrier ski-jump take-off trials (10). Fig 6 shows a typical beat-to-beat heart rate plot recorded during a ski-jump take-off - rated at 4. This can be compared with Fig 7 which was recorded from the same pilot during his first ski-jump take-off - rated at 6. There was good overall agreement between heart rate responses and workload ratings in demonstrating that workload levels for these ramp take-offs are no higher than those for conventional short take-offs from a runway.

During a more recent series of flights to Evaluate Economic Category 3 landings - consisting of autopilot approaches to a 50 ft or 60 ft decision height (depending on aircraft type) ending in a manual landing - pilots' heart rates and workload ratings were recorded (11). Fig 8 is of a typical heart rate plot indicating the increase in workload as decision height is neared and manual control assumed for the landing. The scatter diagram (Fig 9) illustrates graphically the relationship between 32 heart rate responses and workload ratings in real fog for the senior project pilot.

These brief accounts of in-flight workload assessment at Bedford are offered solely as examples of the way in which the technique of recording pilot's heart rates to supplement their subjective opinions has been developed and used in practice. More detailed reports of individual trials have been published elsewhere (see references).

#### Bae 146 WORKLOAD CERTIFICATION

A short description of the use of this technique in the certification of a new passenger transport aircraft - the British Aerospace 146 four-jet feederliner - is relevant.

Following a long, and sometimes acrimonious, debate between pilot unions on the one hand and operators and manufacturers on the other, on whether jet transport aircraft can be flown safely by two pilots President Reagan attempted to resolve the controversy by establishing a Task Force to examine the question impartially(12). The President's Task Force on Aircraft Crew Complement, which reported in July 1981, identified flight deck workload analysis and measurement as a major issue and pointed out that the only generally accepted method for evaluating workload at present is task/time-line analysis based on a comparison with previous aircraft and flight deck designs. It was suggested that this method, supplemented by improved subjective methods by suitably qualified pilots, should offer the best means for demonstrating compliance with FAA complement criteria for new aircraft (FAR 25 1523 Appendix D).

In the United Kingdom the CAA has adopted the Joint Airworthiness Requirements (JARs) which include a direct reproduction of FAA 25 1523 (as JAR 25 1523). British Aerospace elected to use a combination of techniques to assess workload in the Bae 146 - the first aircraft to be certified under JARs - during only one evaluation exercise. The basis for evaluation was a mini-airline exercise of the type already performed by Boeing - for the 757 and 767, McDonnell-Douglas - with the DC9-80, and Airbus - with the A300 FF. In late 1982 three teams of two pilots each flew consecutive three-day intensive flight schedules around a circuit of three major high intensity airfields, with crew duty hours considerably in excess of those allowed by the CAA for passenger carrying operations. Crew workload was assessed by means of subjective estimates from the pilots and flight observer (using a rating scale and post-flight questionnaire) and by recording heart rate; flight deck activity and performance - including error counts - were monitored by video cameras situated on the flight deck.

Heart rate was recorded continuously from before pre-start checks to after shut-down checks. Isolation pre-amplifiers situated on the cockpit floor fed the ECG signal, as pulses, to a Hewlett Packard 9826A computer. Heart rate in beat-to-beat format was displayed in real time on the computer's in-built CRT and then plotted out by a HP 2763 graphic printer. Each plot - or frame - was for 300 sec of heart rate plus a 60 sec plot in a negative direction, ie from the previous frame. Fig 10 shows a typical frame (recorded during a take-off from Amsterdam).

Workload ratings - using the ten-point scale developed at Bedford - were obtained from both pilots and the flight observer on verbal request and light signal from the exercise controller by means of small keyboards fitted to the control column and the observer's clip board. Ratings were plotted at the time of request - according to a pre-determined plan - on the heart rate plot (fig 10). Pilots were instructed in the use of the rating scale before the exercise started and, in particular, were asked to consider the load during the previous 30 sec. Ratings were requested frequently during high workload phases of flight such as the take-off and departure, the approach and landing, and when simulated in-flight failure occurred; during the cruise ratings were requested less frequently.

As the main reason for recording heart rate was to augment subjective assessments of workload, mean rates for the 30 sec proceeding each set of ratings were calculated and then collated with those ratings. Results for the high workload phases were of particular interest; fig 11 shows plots of heart rate (30 sec) and of ratings from a handling pilot and from a flight observer recorded during a take-off and departure from Hatfield (A) and during an approach and landing at Amsterdam (B).

Overall there was reasonably good agreement between heart rates and workload ratings, but there were a number of exceptions which were probably due partly to inexperience with the rating scale and partly to rating the instantaneous level of workload rather than the level experienced during the previous 30 seconds. For example, whereas fig 10 shows quite good agreement fig 11 shows some disagreement between the (two different) pilot's heart rates and their ratings recorded during a take-off and departure at Amsterdam (A) and an approach and landing at London (B); it is interesting that the flight observers' ratings showed better agreement with the pilot's heart rates.

In addition to relating heart rate with subjective ratings each beat-to-beat plot was examined for unduly high heart rates that may have suggested inappropriately high levels of workload; for rapid changes in rate indicating sudden changes in workload; and, in the absence of changes in overall rate, for suppression of heart rate variability consistent with increased mental load.

#### RELATIONSHIP BETWEEN HEART RATES AND WORKLOAD RATINGS

The use of heart rate to augment pilot's subjective opinions of workload prompts the question: how good is the relationship between the two measures? It is a question that should really have been examined in more detail some years ago, but now, after previous plans failed to materialise, a more scientific study is underway at Bedford. The experiment involves four short but well defined flight tasks, (generating theoretically markedly different levels of difficulty) being flown in a BAe 125 twin-jet 'business' aircraft by several pilots. The sequence of tasks - to be flown within laid down performance limits - consists of :

- 1 A 360° turn in 2 min at constant altitude, IAS, and rate of turn.
- 2 A 360° turn in 2 min with a simultaneous loss of 2000 ft in altitude at a constant IAS and rate of turn.
- 3 A 360° turn in 2 min with a simultaneous 2000 ft altitude loss followed by a reverse 360° turn in 2 min with a simultaneous gain of 2000 ft at a constant IAS and rate of turn.
- 4 A 360° turn in 2 min with a simultaneous altitude loss of 2000 ft and speed reduction of 100 kts.

The subject pilot's heart rate is monitored throughout and he is also asked to rate each task, or if he desires each sub-task, using the ten-point workload rating scale referred to earlier (Fig 1).

It was originally intended that the experiment would follow a design based on the Latin square with each pilot flying the sequence four times in different order. In practice, partly because the aeroplane is available only on an 'opportunity' basis and partly because more pilots have offered to participate, the sequence is now flown in the most convenient order at the time; but with the first task to be flown being repeated at the end of the sequence.

Performance was measured during the first few flights but as it became clear that pilots were striving to achieve their best performance anyway, and as it was proving difficult to apply error scores to heart rates and ratings, subsequent sequences have not been measured.

All subject pilots are highly experienced but those not current on the aeroplane are given at least 30 mins familiarisation before being asked to rate the tasks; similarly, pilots unfamiliar with the rating scale are given a full briefing beforehand.

To date 11 pilots have flown a total of 14 sequences and results show an extremely good relationship between their mean heart rates and their workload ratings for 5 pilots, reasonably good agreement for another, and no agreement at all for one pilot. This latter finding is in accord with previous observations on individual responses that occasionally a pilot's heart rate, whilst responding qualitatively with expected changes in workload,

fails to agree with his subjective ratings. Figs 13 and 14 are of scatter diagrams showing the relationship between heart rate means and workload ratings for 7 pilots.

## DISCUSSION

There is now substantial evidence to suggest that during normal but demanding flight heart rate responses in experienced pilots are determined almost entirely by their workload, so it is interesting to speculate on the neuro-physiological mechanisms that are likely to be involved. Rarely is the heart rate change due to the influence of physical activity or to environmental stressors which in normal flight are quite low, although higher rates recorded from pilots in the manual control loop suggest that increased neuro-muscular activity of some form may play a part.

Piloting an aeroplane, especially during the more difficult manoeuvres, requires the brain to collect, filter and process information quickly, to exercise judgement and make decisions, and to initiate rapid and appropriate actions. This neurological activity - which must have been essential for the survival of primitive man - is associated with a state of preparedness sometimes known as arousal. Furthermore, there is experimental evidence that increased arousal, up to a moderate level, enhances a person's capacity for complex skills and thus improves performance (13). It has been suggested by several people that the relationship between performance and arousal can be described by an inverted 'U'-shaped curve - though there is only meagre evidence in support (14). Nevertheless, a theoretical relationship of this type has a particular attraction in the context of flying as there is evidence that both under- and over-arousal have preceded landing accidents where performance was clearly below an acceptable level.

There is also some experimental evidence that a similar inverted 'U'-shaped function describes the relationship between performance and task demands (15); and it has been suggested that levels of arousal are determined by task characteristics or demands, by how the individual perceives the situation, and by how he responds to his environment (16). So one can speculate that a pilot is more likely to produce an adequate - if not optimum - level of performance by matching his level of arousal to the perceived difficulty of the flight task. The result will depend largely on his training and experience although if the task is a novel one, as happens often in test flying, a significant element of empiricism must be involved. Of course, the level of arousal should be high enough for the task itself and also high enough to allow for the unexpected; for example, an engine failure on take-off may require extremely rapid and appropriate actions.

On occasions at Bedford, it has been obvious from the sudden increase in heart rate after the start of a manoeuvre that a pilot had failed to anticipate the difficulties of the task and 'set' his arousal accordingly. Conversely, high heart rates have been recorded when there was an element of uncertainty about the task; this was particularly noticeable for the novel 'ski-jump' take-offs and for a pilot's first approach and manual landing in fog. At such times there has been a tendency for heart rate levels to disagree occasionally with workload ratings.

Support for these speculations is provided by experimental evidence showing that appropriate pathways in the brain and central nervous system do exist. The concept of arousal is an oversimplification of complex neuro-physiological mechanisms but it is functional and, providing it is not confused with emotion, it explains the relationship between a pilot's workload and his heart rate in a convenient manner.

The use of physiological variables to indicate levels of workload has been viewed with suspicion by many people and the use of only one variable - such as heart rate - has been criticised in particular. However, many of these criticisms have been based on the results of laboratory and flight simulator experiments where quite often the task and the levels of workload were unrealistic.

It can be argued that in using heart rate to augment pilots' subjective ratings two variables are being used; but it is questionable whether they are really separate measures - the relationship, already discussed, between perceived difficulty of the task, arousal, and heart rate precluding true independence; although, as the actual neurological mechanisms are uncertain the possibility of conscious unawareness must exist.

Heart rate and subjective ratings are relatively coarse measures - a fact which is often criticised by scientists accustomed to using more precise measuring techniques. But when one considers the individual variations in the different aspects of skill between pilots, and even within the same pilot from time to time, the search for minor differences in workload may be unrealistic in real-world conditions. It is also worth noting that expected differences in workload may be more theoretical than real and so before deciding whether heart rate or subjective ratings can differentiate between workload levels it is important to be sure that there is in fact a real difference (10).

Experience at Bedford has shown that when a pilot is in the control loop, or is expecting to enter the loop, and when the flight task is reasonably demanding, heart rate will usually identify meaningful changes or differences in workload. When the task is relatively undemanding or when the pilot is in a purely monitoring role, as happens frequently in the new generation of civil transport aircraft, heart rate alone may not discriminate between small differences in workload - although subjective ratings may well do so. Assessment of cognitive activity of this type is more difficult but often, in

these instances, visual inspection of beat-to-beat plots will reveal changes in the degree of physiological heart rate variability (sinus arrhythmia) which may signify significant changes in mental workload (17).

Monitoring heart rate during flight seems to be readily accepted by pilots; Bedford test pilots have co-operated to the extent of applying their own electrodes and preparing recording equipment for flight on many occasions. The heart rate data are often studied with interest by the pilots who find their results helpful in recalling various aspects of the sortie.

In its final form the workload rating scale has been generally welcomed by pilots who find it relatively easy to use in practice - even, surprisingly, after the briefest of introductions. Recently the rating scale has been used quite effectively to assess workload on the Boeing 737 by airline pilots unfamiliar with the technique. These favourable observations are probably due to the use of a definition of workload accepted by pilots and to basing the scale on the idea of spare capacity.

The experiment to examine the relationship between heart rates and workload ratings has so far produced some encouraging results. The reasons for the poor agreement for one pilot in this series, and a small number of pilots during flight trials, are not known. Are the heart rate responses at fault? or are the subjective assessments unreliable? At times some pilots tend to give a rating of the instantaneous workload when asked rather than to consider the workload level over the previous period or task. A pilot may misinterpret the demands of the task - this is a recognised cause of aircraft accidents; or perhaps the physiological mechanisms involved in tuning the arousal level are at fault.

In the BAe 146 certification exercise there was a reasonable level of agreement between heart rates and ratings but inconsistencies and anomalies did occur, although neither ratings nor heart rate levels suggested unusually high workload. In this trial and in others anomalies and inconsistencies in heart rate responses and workload ratings have been resolved by discussion with the pilots and re-analysis of the data in most instances. For example, several inconsistencies could be traced to rating instantaneous workload rather than that for the previous 30 sec. It was interesting to note that on several occasions a flight observer's ratings agreed more closely with the pilot's heart rate than did his own ratings (fig 11).

Such observations underline the need to use at least three pilots when assessing workload.

#### CONCLUSIONS

So far evidence supports strongly the use of the methodology described above in the practical assessment of workload in flight; but experience has underlined the need for the following points to be born in mind:

it is important to use a well designed rating scale which is easy to use and fully understood by pilots;

such ratings are increased in value by recording the pilot's heart rate;

heart rate responses are idiosyncratic and so each pilot should be used as his own control;

to a large extent the same applies to subjective ratings;

neither heart rate responses nor ratings are absolute measures of workload.

In the long term a more sophisticated, reliable and sensitive measure of workload may be developed, but it is suggested that for the time being the use of a technique based on the one described in this paper is worth contemplating.



## REFERENCES

- 1 Gulick Ramona. Validation of Pilot Workload Estimates Utilizing In-flight Data. Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics. AFFTC-TR-82-5 254-274 1982.
- 2 E L Brown, G Stone, W E Pearce. Improved Cockpits Through Flight Crew Workload Measurement. Paper presented to Second Advanced Aircrew Display Symposium. Naval Air Test Centre Patuxent River 1975.
- 3 R T White. Task Analysis Methods. A Review of Development Techniques for Analysing Mental Workload in Multiple-task Situations. McDonnell Douglas Corporation Report No MDC J 5291 1971.
- 4 G A Ellis, A H Roscoe. The Airline Pilot's View of Flight Deck Workload: A Preliminary Study Using a Questionnaire. Royal Aircraft Establishment Technical Memorandum FS(B) 465 1982.
- 5 G E Cooper, R P Harper. The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities. NASA Technical Note 5153 Washington DC 1969.
- 6 A H Roscoe. Use of Pilot Heart Rate Measurement in Flight Evaluation. Aviat Space Environ. Med 47 86-90 1976.
- 7 G A Ellis. Subjective Assessment Pilot Opinion Measurements. In: Assessing Pilot Workload Ed. Roscoe A H. AGARDograph AG-233 AGARD Paris 1978.
- 8 A H Roscoe. Heart Rate Monitoring of Pilots During Steep Gradient Approaches. Aviat. Space Environ. Med 46 1410-1415 1975.
- 9 A H Roscoe. Pilot Workload During Steep Gradient Approaches In: Conference Proceedings No 212 Aircraft Operational Experience and its Impact on Safety and Survivability. AGARD Paris 1976.
- 10 A H Roscoe. Handling Qualities, Workload, and Heart Rate In: Survey of Methods to Assess Workload Eds. B O Hartman and R E McKenzie. AGARDograph No 246 AGARD Paris 1979.
- 11 A H Roscoe. Pilot Workload and Economic Category 3 Landings In: Conference Pre-prints. Aerospace Medical Association Annual Scientific Meeting 1980.
- 12 J L McLucas, F J Drinkwater, H W Leak. Report of the President's Task Force on Aircraft Crew Complement. Washington DC 1981.
- 13 E Duffy. Activation and Behaviour. J Wiley and Sons New York 1962.
- 14 A T Welford. Arousal, Channel Capacity and Decision. Nature 194 365-366 1962.
- 15 A T Welford. Stress and Performance. Ergonomics 16 567-580 1973.
- 16 D Kahneman. Attention and Effort. Prentice-Hall Inc Englewood Cliffe 1973.
- 17 H J G Zwaga. Psychophysiological Reaction to Mental Tasks - Effort or Stress. Ergonomics 16 61-67 1973.

**PILOT WORKLOAD RATING SCALE**  
(for a specified piloting task)

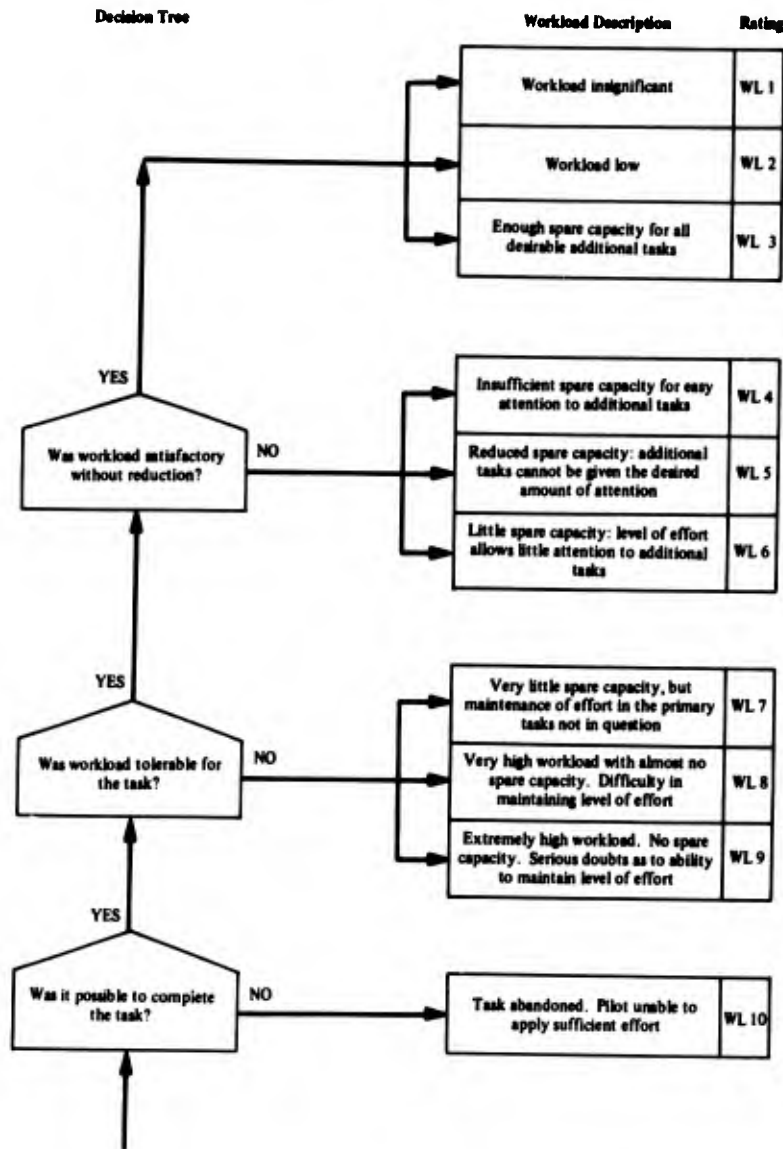


Fig 1 Pilot Workload Rating Scale

The pilot starts his decision-making process at the bottom left corner of the 'decision tree'.

The workload being assessed is that involved in the execution of the primary task. The pilot will almost certainly be performing additional tasks, but the effort expended on them must be included as part of his spare capacity.

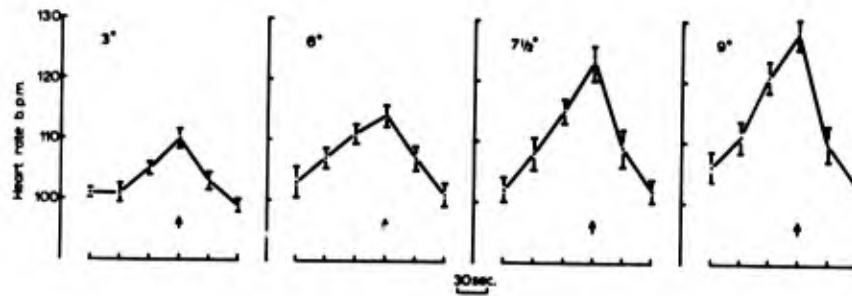


Fig 2 Overall mean 30 sec heart rates ( $\pm$ SEM) for four statistically designed sorties of single-segment experimental noise abatement approaches in an HS 748 Andover - one pilot. The arrows indicate the epochs centred on the touchdown.

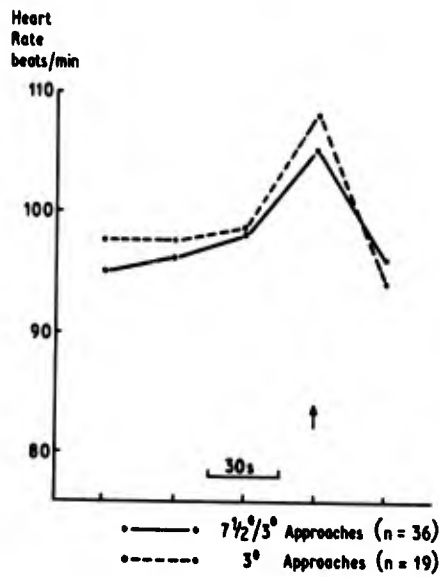


Fig 3 Overall mean 30 sec heart rates for 7½°/3° two-segment approaches and for conventional 3° approaches in an HS 748 Andover - one pilot. The arrow indicates the touchdown epoch.

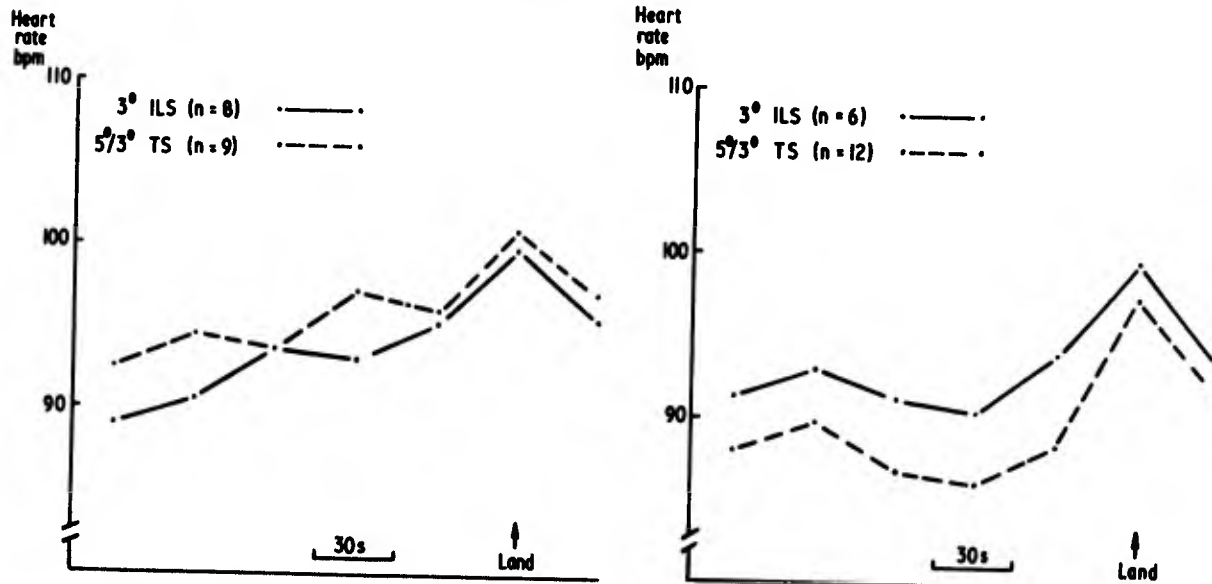


Fig 4 Overall mean 30 sec heart rates for 5°/3° two-segment approaches and for conventional 3° approaches in a VC-10 - two pilots. The arrows indicate the touchdown epoch.

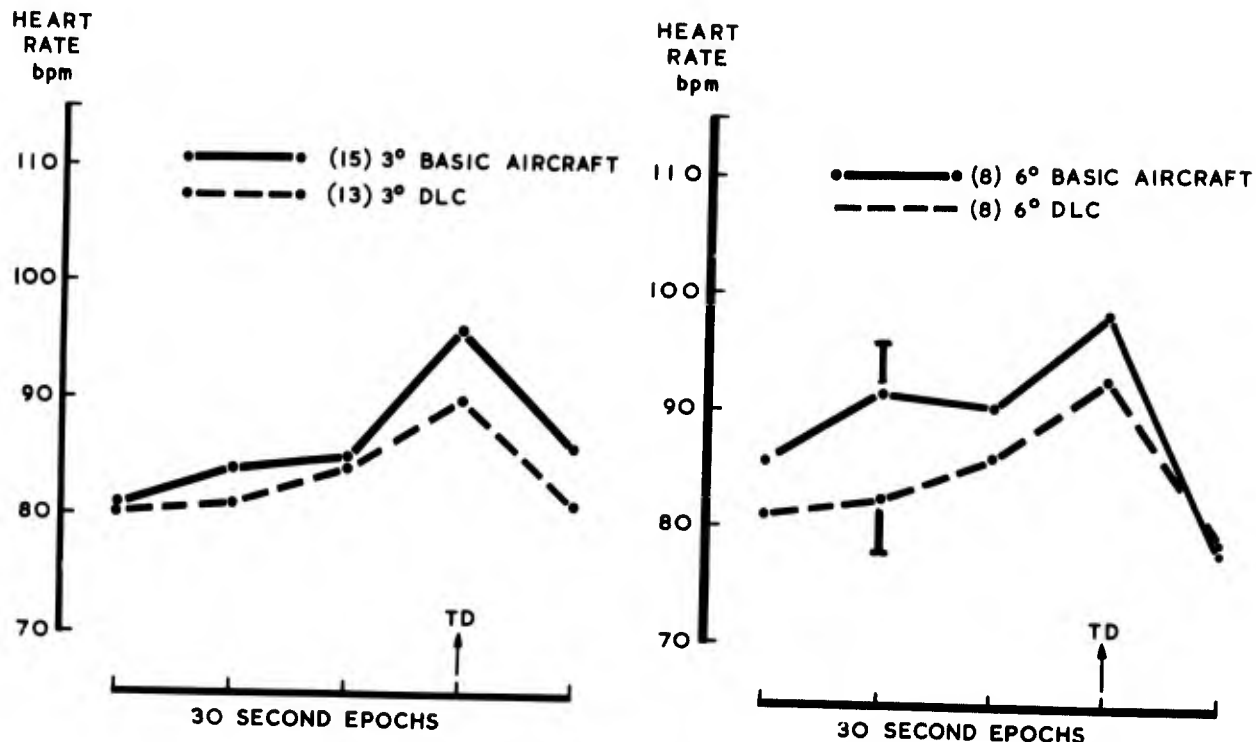


Fig 5 Overall mean 30 sec heart rates for 3° and 6° approaches and landings flown with and without DLC in a BAC 1-11 - one pilot.

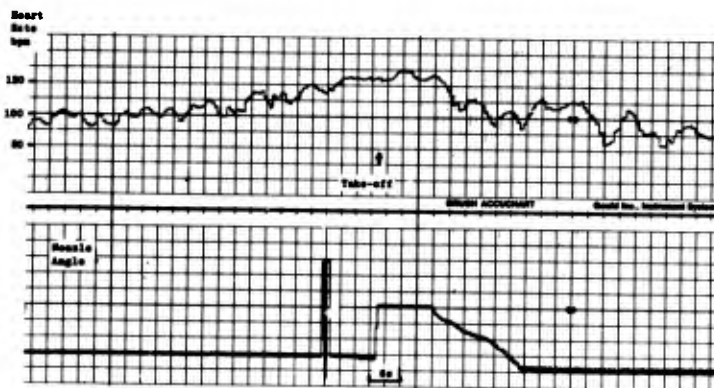


Fig 6 Typical beat-to-beat heart rate and nozzle angle traces recorded during a ski-jump take-off in an HS Harrier. (Note the rapid downwards rotation of engine thrust (upwards in the trace) as the aircraft left the ramp (arrow) followed by the gradual rearwards rotation (downwards in the trace) as speed was increased.)

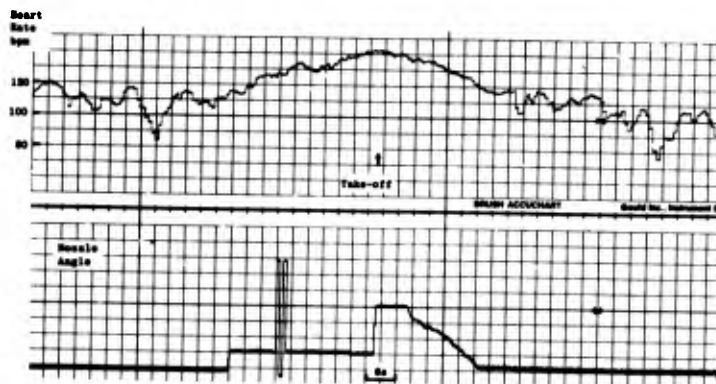


Fig 7 Beat-to-beat heart rate and nozzle angle recorded during the first ski-jump take-off by this pilot - the same pilot as in Fig 6.

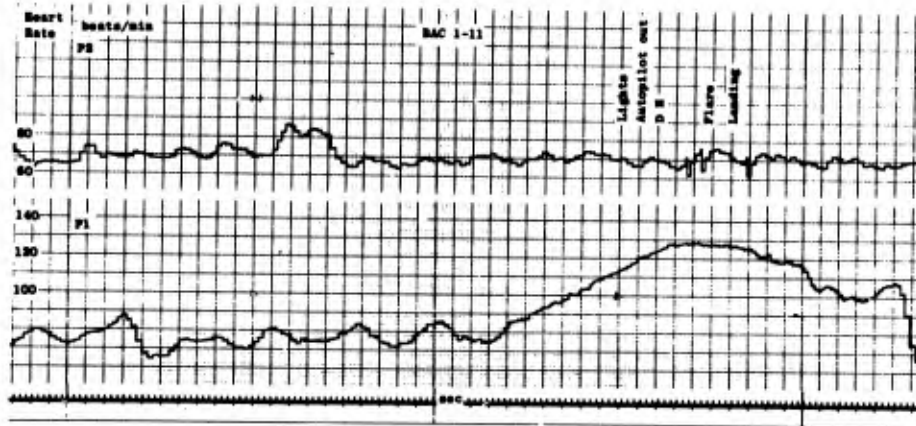


Fig 8 Typical beat-to-beat heart rates (P1 and P2) recorded during an autopilot approach and manual landing in fog (RVR 200 m) BAC 1-11.

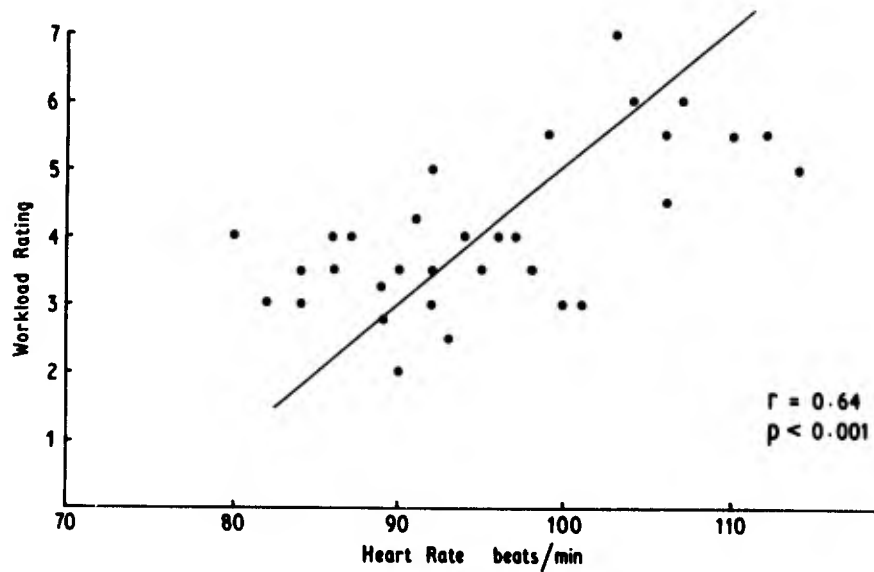


Fig 9 One pilot's heart rate responses and workload ratings for manual landings in fog.

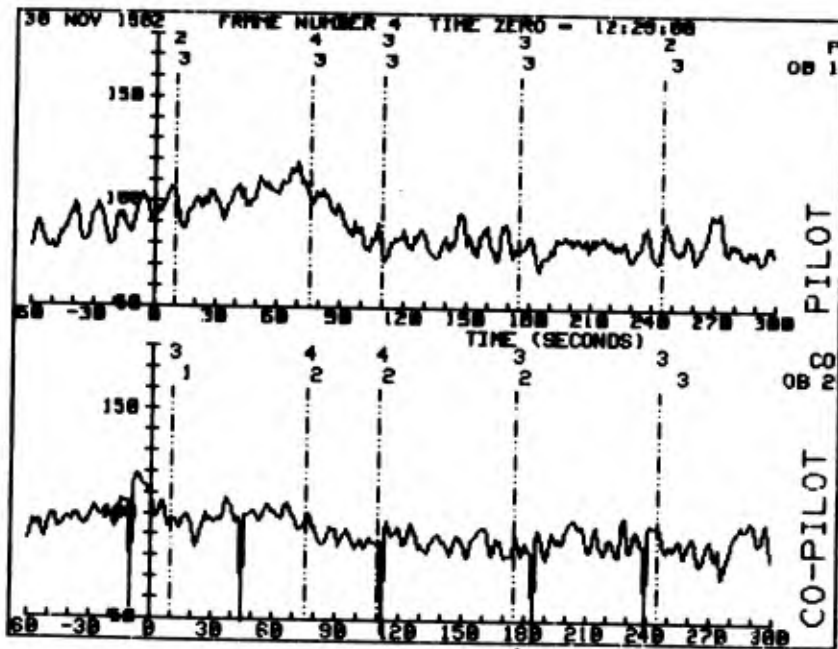


Fig 10 Example of beat-to-beat heart rate plots and workload ratings from the BAe 146 certification programme. (Recorded during a take-off from Amsterdam.)

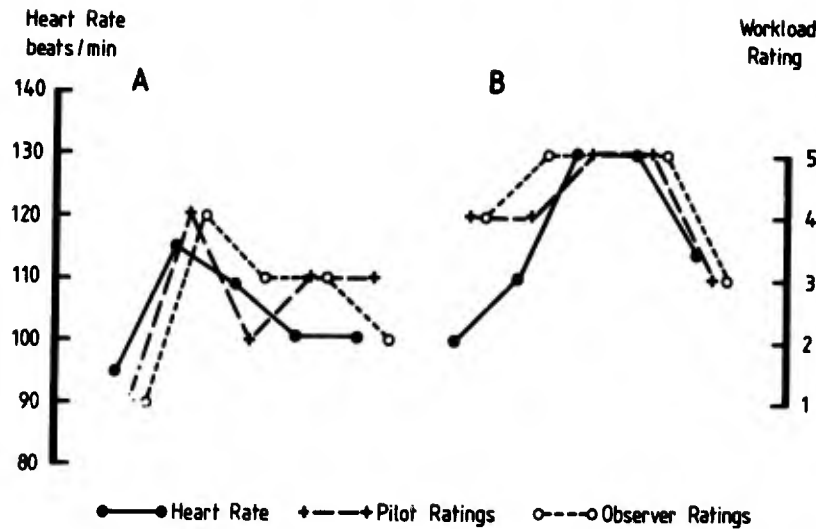


Fig 11 Plots of man heart rate responses (30 sec) and of workload ratings from the handling pilot and flight observer - BAe 146.  
A - Take-off and departure from Hatfield.  
B - Approach and landing at Amsterdam.

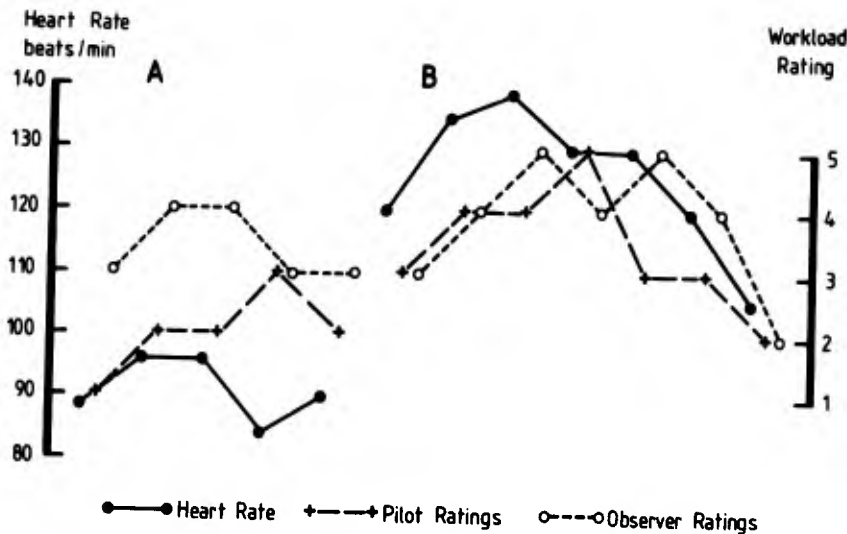


Fig 12 Plots of mean heart rate responses (30 sec) and of workload ratings from the handling pilot and flight observer - BAe 146.  
A - Take-off and departure from Amsterdam.  
B - Approach and landing at London.

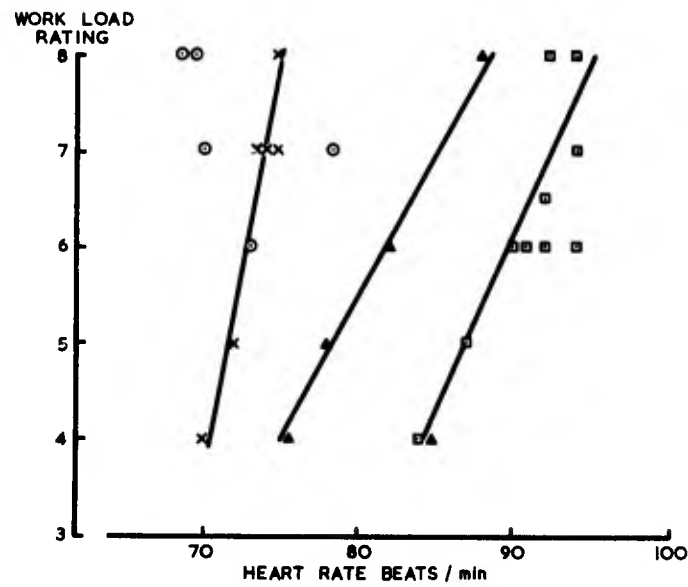


Fig 13 Heart rate responses and workload ratings for four pilots.  
HS 125.

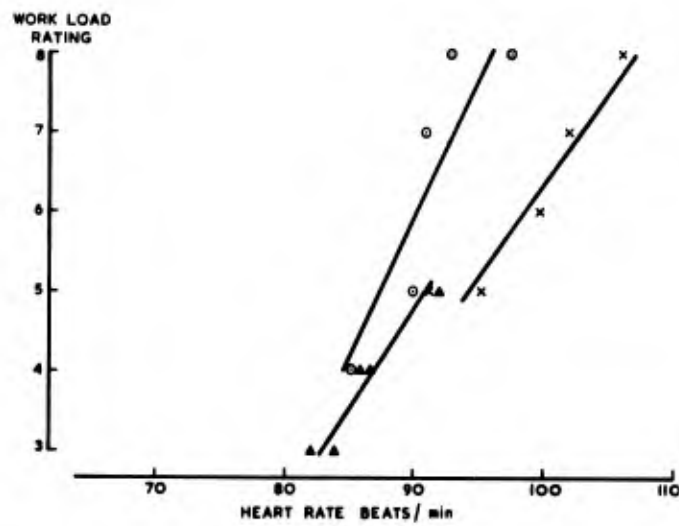


Fig 14 Heart rate responses and workload ratings for three pilots.  
HS 125.