COMPONENT PART NOTICE

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

TITLE: The Practical Assessment of Pilot Workload: Flight Mechanics Panel of AGARD.

TO ORDER THE COMPLETE COMPILATION REPORT, USE AD-A184 834.

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDING, 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#: P005 629 thru P005 643	AD#:
AD#:	AD#:
AD#:	AD#:

Accession For		
NTIS GRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A-1		

DTIC FORM 463 MAR 85

OPI: DTIC-TID

AD-P005 639

CHAPTER 12

IN-FLIGHT ASSESSMENT OF WORKLOAD USING PILOT RATINGS AND HEART RATE

by

Alan H Roscoe
Britannia Airways
Luton, England

INTRODUCTION

At present the most used and probably the most reliable methods for assessing pilot workload in flight are based on some form of subjective reporting by experienced test pilots. Unfortunately, subjective opinions are susceptible to bias and preconceived ideas and so may occasionally result in false estimates of workload. For more than fifteen years subjective reporting by pilots at RAE Bedford has been augmented by recording their heart rates. At first pilots described workload in a relatively unstructured manner but the need for some form of rating scale was soon apparent. After much trial and error and with the valuable assistance of practising test pilots a ten-point rating scale using the concept of spare capacity was developed (fig 1). The overall design is based on the Handling Qualities Rating Scale of Cooper and Harper (1) already familiar to Bedford test pilots and sometimes used previously, though mistakenly, to rate workload (2).

During the last eight years a number of flight trials at Bedford, including the Harrier 'ski-jump' take-off trial and the Economical Category 3 landing trials, have used pilot ratings and heart rate responses to assess workload (3) (4).

The rationale for using heart rate in assessing pilot workload is based on the concept of neurological arousal. Flying an aeroplane, especially during the more difficult manoeuvres, requires the pilot's 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. There is evidence that increased arousal up to a moderate level enhances a person's capacity for complex skills; and it has been suggested that the relationship between performance and arousal can be described by an inverted 'U'-shaped curve (5)(6). There is also some experimental evidence that a similar shaped function describes the relationship between performance and task demands. In addition it has also been suggested that levels of arousal are determined by task characteristics or demands, by how an individual perceives the situation, and by how he responds to his environment (7)(8). It is hypothesised that a pilot is more likely to produce an adequate — if not optimum — level of performance by matching his arousal to the perceived demands or difficulty of the flight task. A coarse setting of his arousal may be followed by fine tuning as the task develops. Heart rate tends to reflect neurological arousal via activity in the autonomic nervous system. An appropriate definition of pilot workload, modified slightly from that proposed by Cooper and Harper in the introduction to their Handling Qualities Rating Scale, is: pilot workload is the integrated mental and physical effort required to satisfy the *perceived* demands of a specified flight task. The interpretation of workload as effort is one that appears to agree with the views of more than 80% of military pilots and civil airline pilots (9), as well as being consistent with the effect on piloting ability of a number of individual variables.

Description of the Technique

Workload ratings — It is almost essential when using a workload rating scale to specify the flight task in reasonably precise terms. The workload being assessed should be 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.

Ratings, which should be given in flight wherever possible, may be for a complete flight task, for example, an instrument approach and landing, or for a sub-task, such as becoming established on the glide slope. On the other hand an experimental protocol may require regular ratings at specified time intervals which might very according to the stage of flight; perhaps being more frequent during expected high workload phases of flight. Regular ratings of this kind tend to be less reliable unless related to a particular flight task.

The rating scale is not linear and probably lacks sensitivity at the lower end; 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 3½

It is important that pilots are fully briefed on the scale to be used. In its final form this particular scale has been generally welcomed by pilots who find it relatively simple to use in practice, especially so if the task to be rated is short and well defined. Somewhat surprisingly, airline pilots unfamiliar with rating techniques have recently used the scale with good effect in assessing workload on Boeing 737 and 767 aircraft. 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.

Recording Heart Rate Heart rate recording is non-intrusive and it is compatible with flight safety; pilots seem readily to accept being 'wired up'; and the discrete nature of the basic data encourages various forms of analysis. The technique used to record heart rates from pilots during flight is based on the electrocardiogram (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 (which might include workload ratings) and, where possible, other aircraft parameters. In the first instance the basic signal — the 'R' wave of the ECG — is plotted out along with heart rate in instantaneous or 'beat-to-beat' form (derived from the 'R' waves by cardiotachometer). Subsequently mean rates for a particular task, sub-task, or time interval may be calculated according to the requirements for workload ratings. Plots of mean rates for 30 sec epochs are often useful in demonstrating

PILOT WORKLOAD RATING SCALE
(for a specified piloting task)

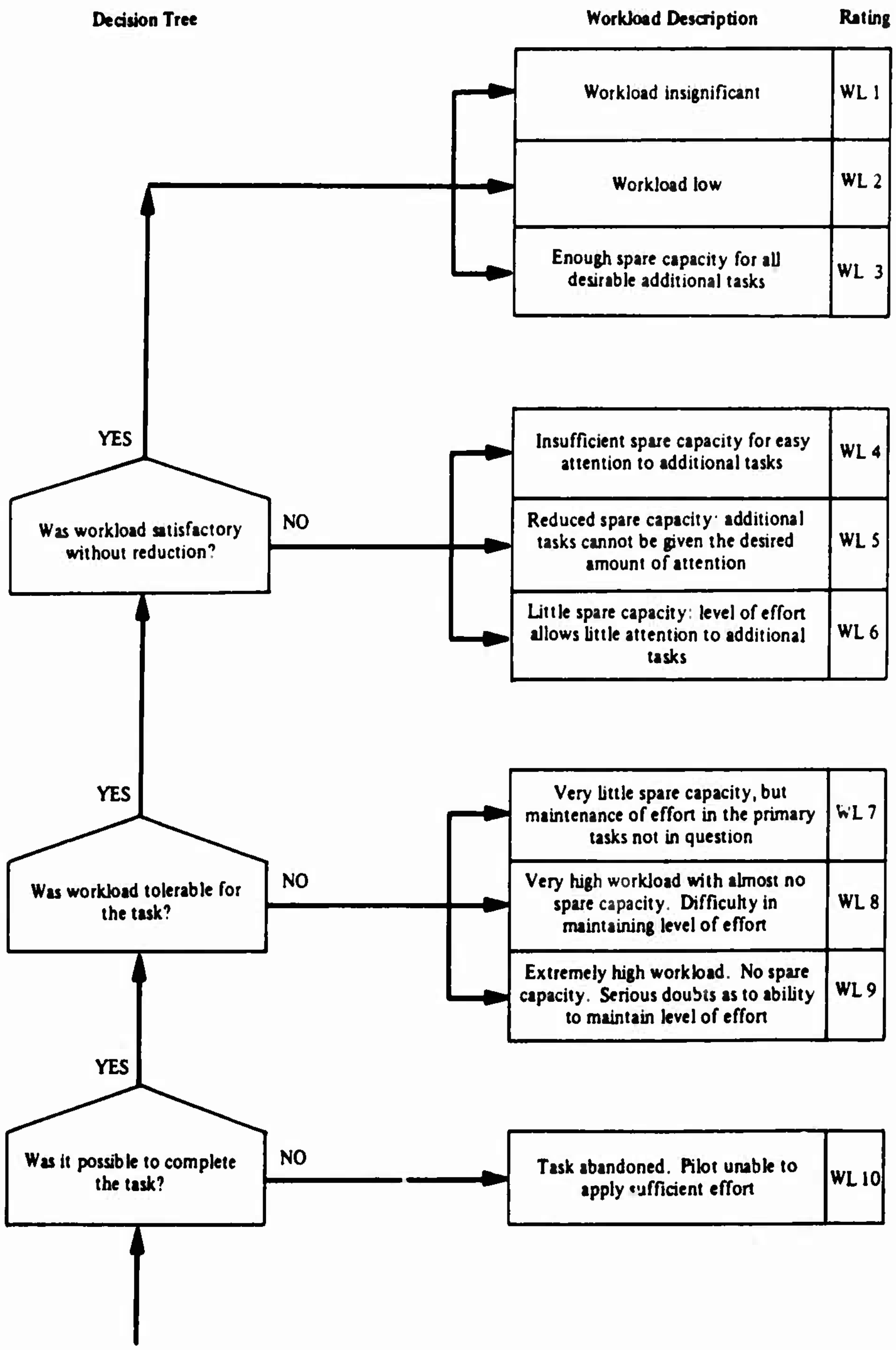

Fig.1 Pilot workload rating scale
The decision-making process is started at the bottom left corner of the 'decision tree'*

*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.

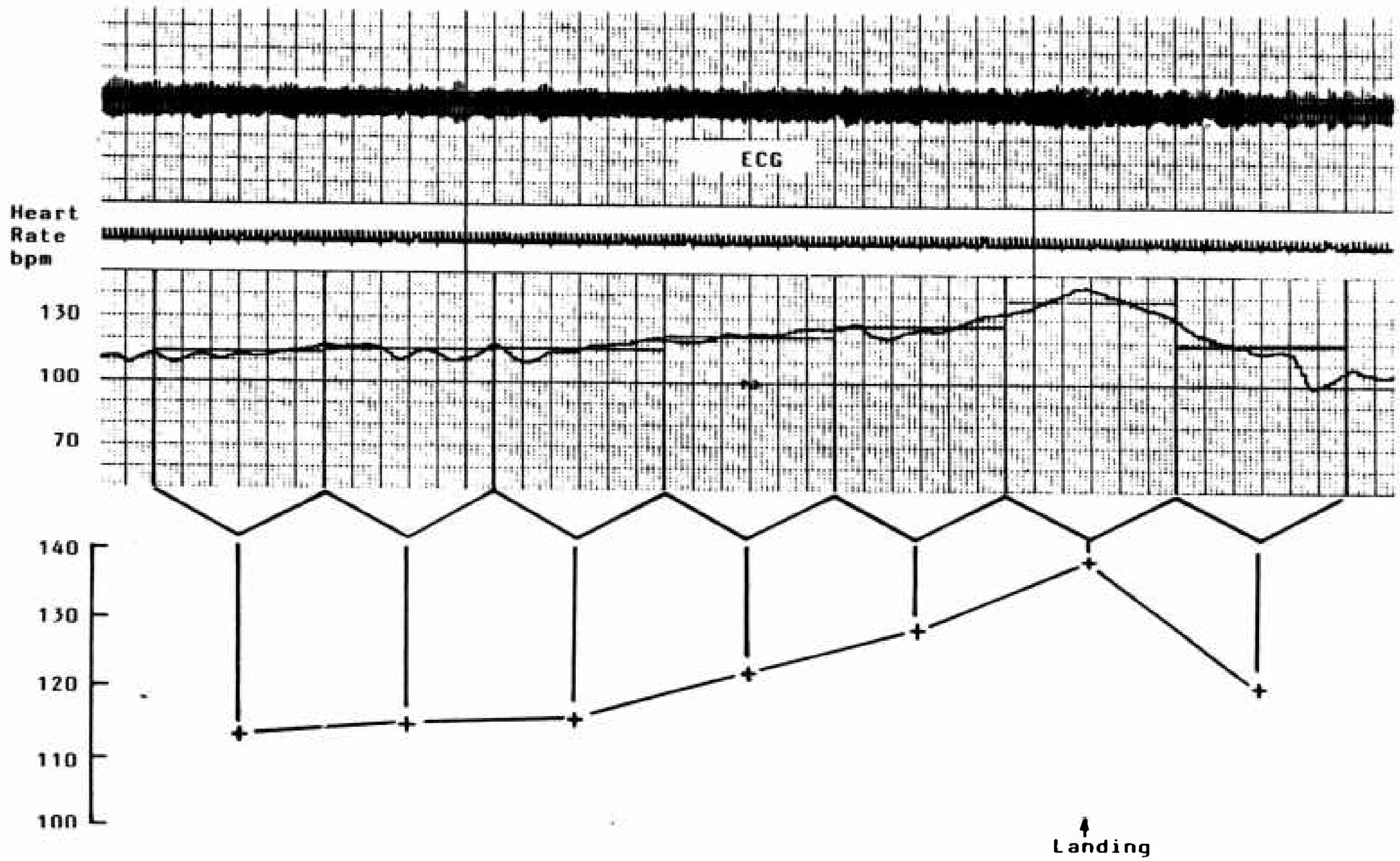

Fig.2 Schematic representation of a 30 sec plot derived from beat-to-beat heart rate; approach and landing B767

significant heart rates changes by smoothing (fig 2 is an example). On the other hand, beat-to-beat plots have the advantage of showing rapid and sometimes short term changes of interest.

In the absence of any significant change in overall heart rate the degree of sinus arrhythmia (physiological heart rate variability) may be of value in assessing changes in mental workload. Changes in sinus arrhythmia are usually evident on visual inspection of beat-to-beat plots; a number of techniques are available for scoring sinus arrhythmia although none seem to be reliable and so results must be interpreted with caution.

Performance — As workload and performance are to a large extent interdependent it is important when assessing the former to monitor the latter. In some flights it is a relatively simple matter to record actual performance in the air by means of aircraft recorders or on the ground by kinetheodolites. Where this is not practicable realistic performance limits should be defined and monitored by a flight observer, by video recording, or by the pilot himself.

Example of using the Technique

Assessing pilot workload during a manually flown instrument approach and landing using a flight director system in a twin-jet transport flown by a crew of two pilots. (See Appendix I for details). The defined task lasts five minutes.

Heart rates are recorded from both pilots continuously throughout. Workload ratings are requested from both pilots and for each pilot from an experienced flight observer seated on the flight deck as follows:

1. At 3,000ft — starting the final descent onto the glide slope.
2. At 1,000ft QFE — for glide slope acquisition.
3. At 100ft QFE — for final approach.
4. On deceleration to 60K — for flare and touchdown.

Untoward events are rated on an *ad hoc* basis. Performance is monitored by the flight observer. Mean heart rates for the appropriate periods before the ratings are calculated and bracketed with the rating scores. The beat-to-beat heart rate plot is examined for evidence of inappropriate or sudden changes and also for suppression of the sinus arrhythmia. (Inspection of heart rate plots by the pilots will often act as an *aide memoire*). Ambiguities and inconsistencies are of particular interest and are studied in more detail.

These data provide some idea of workload levels but become more valuable when compared with data from the same pilots recorded on other occasions when using different techniques or systems, or when flown in different weather conditions. For example, this flight director approach may be compared with an approach using a different type of flight director, with a raw ILS approach, or with an autoland.

Pitfalls and Limitations

The technique described above does not result in the more precise measurements associated with experiments carried out in the controlled conditions of laboratories. Furthermore, there are a number of important limitations and pitfalls to be aware of when assessing levels of workload in real flight.

1. Ratings depend largely on the personal experience of the pilot and do not result in absolute values of workload, comparisons between pilots are, therefore, not valid; minor inconsistencies between different pilots flying the same aeroplane should be expected.
2. In-flight ratings may not be possible when assessing workload in single-seat aircraft.
3. As the rating scale is non-linear statistical treatment of rating numbers must be treated with caution.
4. The idiosyncratic nature of the heart rate response precludes comparison of results derived from different pilots — each pilot must be used as his own control — unless large numbers of pilots are involved.
5. Heart rate responses recorded during flight tasks involving increased physical effort or physical stressors such as high 'g' manoeuvres must be interpreted with care.
6. Ambiguities and inconsistencies between a pilot's ratings and his heart rate responses are sometimes due to a pilot rating a particular aspect of part of a task or epoch rather than the entire task or period of time.
7. The technique is most valuable when the handling pilot is manually flying the aeroplane during a relatively demanding task or when he is anticipating taking manual control at short notice. Both ratings and heart rate responses for non-flying pilots in a purely monitoring role are less valuable, although changes in beat-to-beat heart rate variability can be most useful in detecting changes in mental load.
8. Finally, experience so far suggests that results from one pilot in five show poor agreement between subjective ratings and heart rate responses. The reason for this disagreement is not known for certain but may be due to the failure of heart rate to reflect accurately levels of central arousal in these individuals.

REFERENCES

1 COOPER G E, HARPER RP — The use of pilot rating in the evaluation of aircraft handling qualities. NASA Technical Note 5153, Washington DC, 1969

2 ELLIS G A — Subjective assessment pilot opinion measurements in: Roscoe A H (Ed.) Assessing pilot workload, AGARDograph AG-233, AGARD Paris 1978

3 ROSCOE A H — Assessing pilot workload in flight in: Conference Proceedings No 373, Flight test techniques. AGARD Paris 1984

4 ROSCOE A H — Pilot workload an Economic Category 3 landings in: Conference pre-prints. Aerospace Medical Association Annual Scientific Meeting. 1980

5 DUFFY E — Activation and behaviour. J Wiley and Sons New York 1962

6 HOCKEY R — Stress and the cognitive components of skilled performance in: Hamilton V and Warburton D M (E.Js) Human stress and cognition. John Wiley and Sons Chichester 1979

7 WELFORD A T — Stress and Performance. Ergonomics 16 567-580 1973

8 KAHNEMAN D — Attention and Effort. Prentice — Hall Inc Englewood Cliffe 1973

9 ELLIS G A, ROSCOE A H — The airline pilot's view of flight deck workload: preliminary study using a questionnaire. Royal Aircraft Establishment Technical Memorandum. FS(B) 465 1982