

UNCLASSIFIED

Defense Technical Information Center Compilation Part Notice

ADP010444

TITLE: The Risk of Human Error: Data Collection,
Collation and Quantification

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: The Human Factor in System Reliability Is
Human Performance Predictable? [les Facteurs
humains et la fiabilite des systemes - Les
performances humaines, sont-elles previsibles?]

To order the complete compilation report, use: ADA388027

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, ect. 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:

ADP010439 thru ADP010446

UNCLASSIFIED

The Risk of Human Error: Data Collection, Collation, and Quantification*

J W Chappelow
Centre for Human Sciences
DERA Farnborough
Farnborough, Hants. GU14 0LX
United Kingdom

Summary: Human performance poses significant problems in system reliability assessment. Are realistic assessments of safety in systems involving humans possible? Can human performance be quantified? What aspects of human performance are predictable? Practical experience in the field of aviation safety suggests some answers to these questions.

Introduction: This is a historical account of a variety of projects concerned with human error in aviation. As a summary of personal experience it is necessarily partial, in both senses; that is to say it is an incomplete and biased view of human reliability. It may, nevertheless, cast some light on the themes of the workshop: Can the safety implications of human performance be addressed rigorously? What should be predicted? Is meaningful quantification possible?

Classification 1: Psychologists have assisted Royal Air Force Boards of Inquiry since 1972. By 1982, enough reports on aircraft accidents had been collected to allow a first attempt at organising the data and seeking patterns. The classification scheme devised then had no particular theoretical bias, was simply organised, and allowed the most prevalent contributory factors to be identified.^{1,2,3} They are shown in Table 1 grouped under arbitrary headings.

On the basis of this analysis, research projects addressing personality issues and cognitive failure were undertaken.⁴ Although some interesting findings resulted, neither project led to practical innovations to reduce risk beyond general guidance given to flying supervisors in flight safety courses. It is interesting to note, in retrospect, that both projects addressed individual susceptibility to particular types of error. This was probably a reflection of political rather than technical realities at the time. Although the role of design and organisational factors in human error was well recognised, there was still a remnant of “blame culture” to be overcome.

Table 1: The most common contributory factors

Aircrew		System	
Inexperience	23%	Training & briefing	25%
Personality	21%	Administration	23%
Life stress	14%	Ergonomics	22%
Social factors	11%	High workload	14%
Immediate causes			
Acute stress	26%	Inappropriate model	16%
Distraction	20%	Visual illusion	10%
Cognitive failure	17%		

* © Defence Evaluation and Research Agency

Two sorts of insight resulted from these initial efforts: Identification of the more important contributory factors; and the recognition that both the size of contribution to overall risk and the tractability of the problem were important in determining where to invest remedial effort. Tractability and quantifiability turned out, initially at least, to be associated.

Quantification 1: Few emergencies in aviation require an immediate response. Helicopters have more than their fair share of those that do. A prime example is total power failure. It requires an immediate reduction in collective pitch. How long the pilot has to achieve this depends on the inertia in the rotor disc, and this is an issue of relevance to the certification requirements for helicopters.

Reaction times are relatively easily and objectively measured. They have long been a mainstay of experimental psychology. Unfortunately, it is difficult to generalise with convincing precision from laboratory studies, however sophisticated, to real world situations. It was necessary to resort to flight simulator experiments. Figure 1 shows some of the results for three helicopter types: means and 90th percentiles for detection time (the interval between the emergency onset and the first indication of an appropriate response) and response time (the time taken to complete the action).⁵ It seems that reaction times even for well-practised responses to easily identified conditions can be surprisingly long, particularly when the normal variability of behaviour is taken into account.

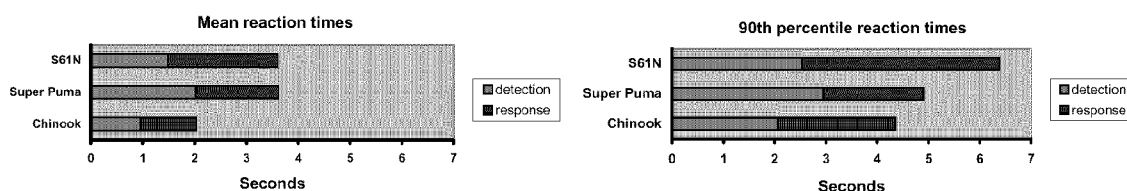


Figure 1: Reaction times to total power failure

These results have a direct bearing on the mechanical design of helicopters. When the probability of total power failure is known, they allow the risk of an unfavourable outcome to be estimated in a way that allows cost-benefit analysis to inform design decisions.

A variety of helicopter emergencies were addressed in this study.⁶ Although some instructive differences were found, similar results were obtained in several cases and in a dissimilar case – an untrained-for and (from the designer’s perspective) unpredictable control malfunction in a fixed-wing aircraft. The findings do provide general guidance on the reaction times to be expected in a range of situations within aviation, at least. It is also clear that there are limits to this generalisability, and it is not clear how wide a range of similar studies would be required to provide comprehensive guidance on reaction times in real situations. Such guidance would, however, be valuable to system designers and regulators, and could be relatively easily obtained. A sensible first step would be the classification of situations in terms of the types of task and responses involved.

Quantification 2: The UK Low Flying System (UKLFS) is uncontrolled airspace from ground level to 2000ft. It is used by a variety of civilian aircraft – hang-gliders, microlights, gliders, fixed- and rotary-wing light aircraft – as well as military helicopters, transports, and fast jets operating at speeds in excess of 400kt. All operate on the “see-and-avoid” principle. The risk of random mid-air collision is real. Collisions involving two fast jet aircraft not surprisingly provide the most numerous examples of this risk. They also represent an extreme and, therefore, relatively simple case, the most important features of which (the psychophysical aspects) can be modelled sufficiently precisely to allow useful predictions to be made.

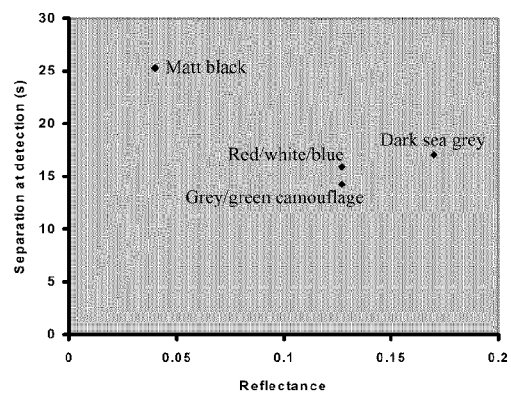


Figure 2: Flight trial results (paint schemes)

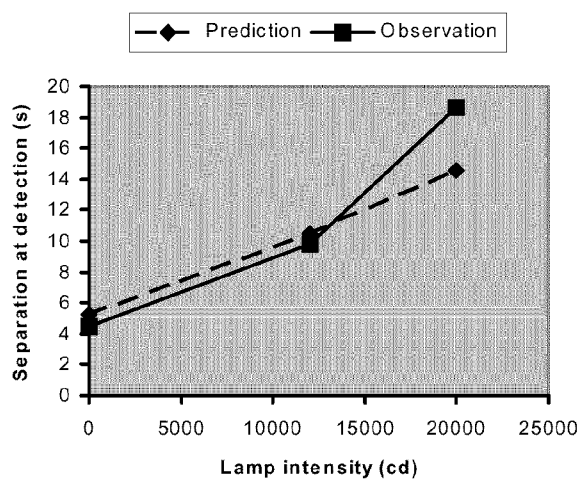


Figure 3: Flight trial results (lamps)

An initial, approximate attempt at such modelling suggested advantages for black paint schemes and for very bright, fixed, steady lights – as opposed to the high intensity strobe lights commonly fitted to aircraft.⁷ It also allowed the risk reduction achievable through electronic collision warning systems to be estimated. Flight trials confirmed the predictions (Figures 2 and 3 show sample results), and supported refinement of the model.^{8,9,10}

In a further project, the psychophysical model was combined with a computer simulation of activity in the UKLFS.¹¹ It was necessary to collect a large amount of data to support this modelling exercise (Figure 4). The resulting predictions were validated against reported conflict rates (from the Joint Airprox Working Group) and the historical record of collisions. The principal predictions (one fast jet–fast jet collision every two years and one military–civilian collision every six years) have continued to prove tragically accurate. However, the estimates of the effectiveness of remedies such as paint schemes and collision warning systems derived from the model have informed the continuing debate on safety in the UKLFS and influenced policy decisions.

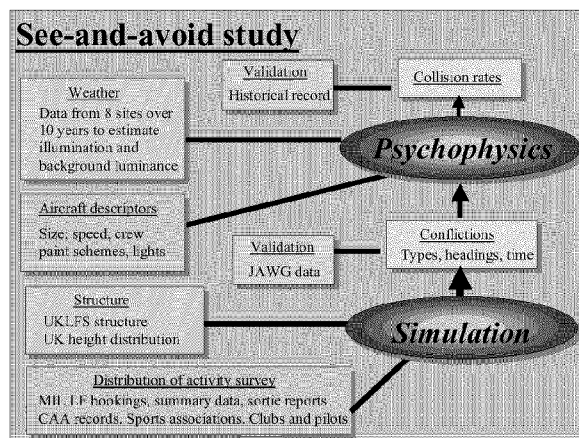


Figure 4: Construction of a predictive model

Classification 2: The need for a precise and useful classification scheme for human error and its underlying causal factors has become more pressing. Involvement with NATO RSG 25 allowed a less aviation-specific model to be drafted, and this formed the basis of a recent project aimed at developing a causal factors database for both the human factors and the engineering domain within military aviation.¹²

Although computer systems are changing the picture, the engineering domain has been characterised by a plethora of subsystems and components each of which has a limited range of functions (usually only one each) and only a few ways of failing. The human factors domain is characterised by one component (*Homo sapiens*) which serves a multitude of goals (rather than simple functions), and has many ways of failing.

Accident and incident databases in aviation have tended to follow a model appropriate to the engineering domain, and have been relatively uninformative as to the causes of human error. Indeed, there is a parallel between the traditional engineering approach (identify the defective component and replace it) and the old-fashioned approach to human error (find out who is to blame and punish them). The new database allows for simple classification of human errors and a flexible, hierarchical coding of causal mechanisms designed to identify all types of contributory factors (Figure 5 is an outline). By imposing a similar model on the engineering domain (Figure 6), a different perspective on the causes of mechanical failure has been obtained, which has resulted in at least one unexpected insight concerning the detection of problems between flights.

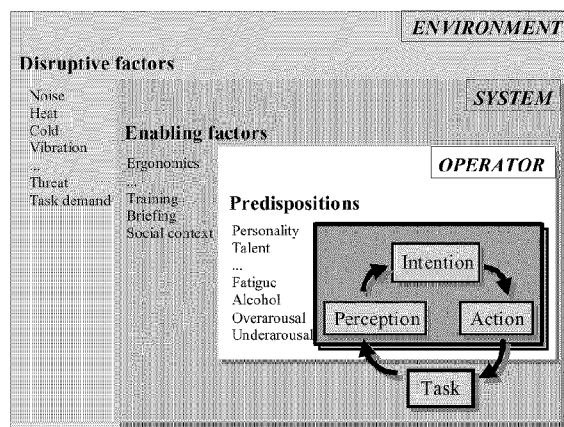


Figure 5: Outline human factors classification

The database has also been used to prototype a risk analysis system. By using historical data to estimate the quality of underlying causal factors and the strength of their influence on failure mechanisms, relatively objective sensitivity analysis has been made possible. Comparison of Figures 7 and 8 shows the broader

perspective and added complexity derived using this approach in comparison with a similar procedure based on experts’ opinions when both approaches were used to analyse the factors underlying one type of accident.

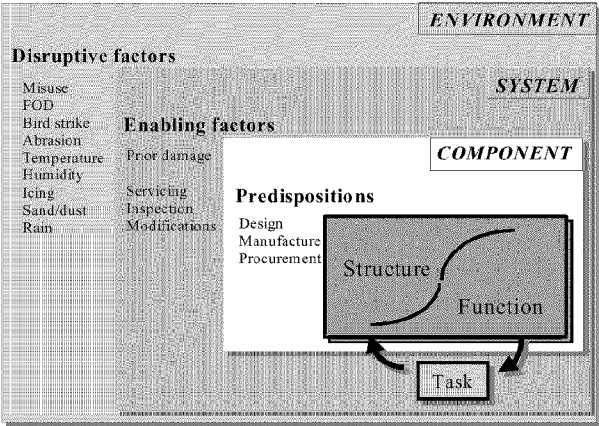


Figure 6: Outline engineering classification

Sensitivity analysis applied to the whole range of accidents has revealed the strongly influential character of social factors in military aircraft accidents – a fact not evident in simpler analyses. These factors can be addressed via training programmes – a relatively cheap and immediate option in comparison with other remedies for error such as hardware modification, for example. The fact that they are influential as well as relatively tractable makes them an important target in flight safety programmes.

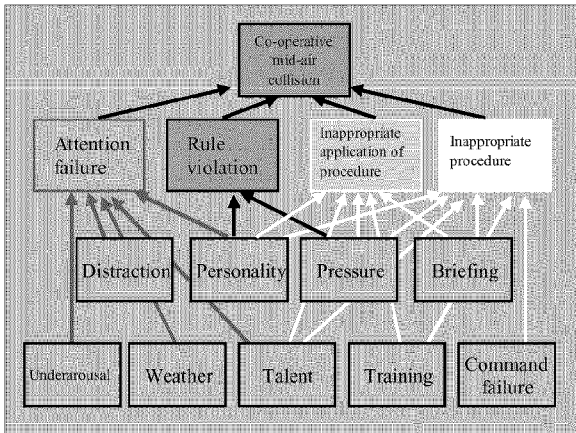


Figure 7: Influence diagram generated by experts

A recent review of social factors in accidents was intended to refine the RAF crew resource management training programme by identifying social factors influencing ground-based as well as airborne activity.¹³ The factors identified include not only communication problems and decision making biases already known to affect small teams, such as the “risky shift” phenomenon, but also organisationally-induced tendencies to more risky behaviour.¹⁴ There may be parallels here with the risk conservation behaviour reported in the road safety context.¹⁵ It is certainly clear that, whatever the intention behind the design of a system, individual operators, small groups or teams, and even whole organisations may use it for aims undreamed of by the designer. Individuals derive status, satisfaction, fun, even thrills from the use of systems, and teams and organisations may similarly add to or even subvert the formally defined purpose. The social contexts that promote these parallel or supplementary purposes deserve attention since they define a whole category of risk otherwise ignored.

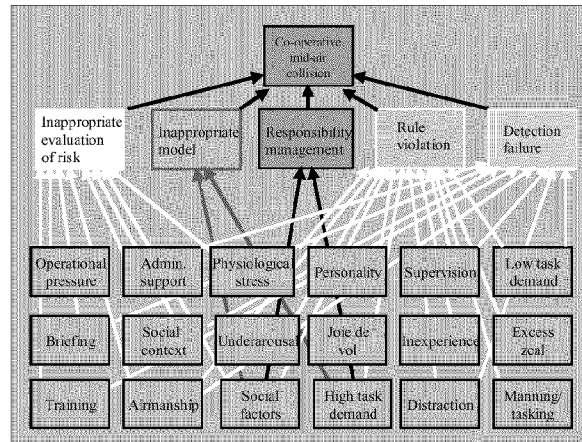


Figure 8: Influence diagram based on historical data

Incident data (and Quantification 3): Accident investigations provide a rich source of detailed information on risk and reliability, but accumulate only slowly. Incidents (near misses) are more numerous, and logically deserve equal attention since in principle, they could provide much more data. Confidential incident reporting schemes have been introduced in many industries besides aviation as a way of increasing the amount of data collected. Recent experience in the RAF suggests that open reporting may be even more effective in uncovering unsuspected problems. Such a system requires the prior establishment of an appropriate organisational culture so that a guarantee of immunity from punishment for honest mistakes will be accepted at face value. There always remains, however, a problem in assessing the magnitude of a reported risk, as the following example illustrates.

Ejection seats are intended to save life, but are potentially lethal. Most are made safe by inserting mechanical barriers into the firing mechanism - usually pins. In thirty years the RAF has recorded two fatal accidents involving ejection seat pins. In one case the seat was safe when it should have been live (a Type 1 error). In the other case it was live when it should have been safe (a Type 2 error). Only eleven *incidents* involving seat pins were formally reported in the same period.

Shortly after the introduction of open reporting, a change in the procedures used at one flying station resulted in several Type 2 errors, which were reported. As interest focussed on this particular location, a small number of Type 1 errors appeared as well. These could not have been caused by the change in procedure. They appeared to have come to light simply because of the locally heightened interest in seat pins procedures. On this basis it was suggested that other aircraft types and flying stations must also be experiencing seat pins errors. This encouragement produced a small crop of reports of both types of error. At this stage, it was clear that a problem of unknown magnitude had been uncovered. To estimate its prevalence, questionnaires were used to capture all seat pins errors occurring during one month.

The results of this survey suggest that about 100 Type 1 errors and 200 Type 2 errors are made every year in the RAF. These potentially lethal errors have presumably been occurring since the introduction of ejection seats, and have barely come to official notice except when accidents occurred. To obtain a realistic estimate of the error rate, it was necessary not only to advance beyond mandatory and confidential incident reporting programmes, but also to collect data on this specific topic for a defined period.

A simple count of the frequency of an error is not enough to gauge its importance. Combining the probability of a Type 1 error with the probability (obtained from accident data) that ejection will be required enables the risk of a fatal outcome to be calculated. This gives real meaning to conventional reliability standards such as 1 fatality in 10^6 or 10^7 sorties. On present estimates, the risk due to Type 1 errors warrants serious consideration of modifications to current operating practices and a re-evaluation of the general approach in future ejection seat designs. If a different standard were adopted, 1 in 10^9 for example, the implications would be far more severe, and immediate, drastic action would be required.

Conclusions: The practical experience described here suggests some conclusions that might possibly have general relevance. Meaningful quantification of human performance in a form that is useable in reliability assessment does seem to be possible. It is, however, probably significant that the two major examples given involve relatively simple aspects of behaviour – reaction times and visual psychophysics. In both examples, the stimulus conditions and the required responses were closely defined. The third (ejection seat) example also involves relatively simple behaviour. In tasks demanding more interpretation or complex decision making, the challenge of meaningful and rigorous quantification may be considerably more daunting.

Although laboratory studies can provide a rigorous understanding of specific error mechanisms, a realistic appreciation of the potential for human error can only be obtained by close scrutiny of real systems. This implies thorough investigation of the human factors aspects of accidents and the collection of data on “near-miss” incidents. Such data are, of course, useless unless organised and collated in a way that illuminates failure mechanisms and allows practical remedies to be devised. We have demonstrated that classification can be developed to the point of permitting relatively objective risk assessment. However, the ejection seat example demonstrates that considerable, focussed effort is required to obtain reliable estimates of error rates in the real world, and that reliance on accident statistics or conventional incident data alone is likely to result in a substantial underestimate.

Finally, although it is possible to quantify the probability of error in, say, dial reading or switch operation in a way that parallels reliability assessment of engineering components, this ignores important facts about human operators. They have goals rather than functions. Some of their goals are not those envisaged by system designers. Some are determined by characteristics of the teams they work in or of the organisation as a whole. These factors are also amenable to systematic analysis, possibly even to quantification. In addressing system reliability, we need to consider not just the artefact-system, or the man-machine system, but the whole system-complex.

References:

1. Chappelow, J. W. (1984). *Human error in aircraft accidents: A review of psychologists' reports on RAF accidents 1972-1982*. IAM Report 663. Farnborough, UK: RAF Institute of Aviation Medicine.
2. Chappelow, J. W. (1988). *Causes of aircrew error in the Royal Air Force*. In *Human Behaviour in High Stress Situations in Aerospace Operations*. AGARD Conference Proceedings 458. Neuilly sur Seine, France: NATO Advisory Group on Aerospace Research and Development.
3. Chappelow, J. W. (1989). *Remedies for aircrew error*. IAM Report 664. Farnborough, UK: RAF Institute of Aviation Medicine.
4. Smith, A., Chappelow, J. W., and Belyavin, A. J. (1995). *Cognitive failures, focused attention, and categoric search*. Applied Cognitive Psychology, Vol. 9, S115-S126.
5. Smith, P. R. and Chappelow, J. W. (1995). *Pilot Intervention Times*. In *Human Factors in Aviation Operations, Proceedings of the 21st Conference of the European Association for Aviation Psychology (EAAP)* Vol 3 ed. R. Fuller, N. Johnson and N. McDonald. Aldershot, UK: Avebury Aviation.
6. Chappelow, J. W. and Smith, P. R. (1997). *Pilot intervention times in helicopter emergencies: Final report*. PLSD/CHS/HS3/CR97020/1.0. Farnborough, UK: DRA Centre for Human Sciences.
7. Chappelow, J. W. and Belyavin, A. J. (1991). *Random mid-air collisions in the low flying system*. IAM Report 702. Farnborough, UK: RAF Institute of Aviation Medicine.
8. Chappelow, J. W. and Belyavin, A. J. (1992). *A trial to assess aids to conspicuity*. IAM Report 723. Farnborough, UK: RAF Institute of Aviation Medicine.
9. Chappelow, J. W., Belyavin, A. J., and O'Connor, E. M. (1992). *A further evaluation of fixed steady lights as aids to conspicuity*. IAM Report 732. Farnborough, UK: RAF Institute of Aviation Medicine.
10. Chappelow, J. W., Belyavin, A. J., and Smith, P. R. (1993). *Aircraft conspicuity and paint schemes: A further trial*. IAM Report 747. Farnborough, UK: RAF Institute of Aviation Medicine.

11. Chappelow, J. W., Belyavin, A. J., Ruskell, L., Procopides, M., Smith, P. R., and O'Connor, E. M. (1997). *See-and-avoid operational analysis study*. PLSD/CHS/HS3/CR97029/1.0 DRA Farnborough, UK: DRA Centre for Human Sciences.
12. Chappelow, J. W., O'Connor, E. M., Johnson, C., and Blair, R. C. (1999). *Causal factors in military aircraft accidents*. DRA/CHS/MID/CR990245/1.0. Farnborough, UK: DERA Centre for Human Sciences.
13. Chappelow, J. W. and O'Connor, E. M. (1998). *Crew resource management training and military flight safety*. DERA/CHS/MID/CR980268/1.0. Farnborough, UK: DERA Centre for Human Sciences.
14. Snizek, J. A. and Henry, R. A (1989). Cited in Sutherland, S. (1992). *Irrationality*. London: Constable and Company Ltd..
15. Janssen, W. (1994). *Seat-belt wearing and driving behaviour: an instrumented-vehicle study*. Accident Analysis and Prevention, Vol. 26, pp 249-261. Pergamon, London.