Aerospace Science and Technology, Volume 9, Issue 6, September 2005, Pages 525-532.

Using SHERPA to Predict Design-Induced Error on the Flight Deck

Don Harris[1*], Neville A.Stanton[2], Andrew Marshall[3], Mark S. Young[2], Jason Demagalski[1] and, Paul Salmon[2]

[1]Cranfield University, Human Factors Group, Cranfield, Bedford, MK43 OAL, UK

Tel: +44 (0)1234 750111 x5196

Fax: +44 (0)1234 750192

E-mail: d.harris@cranfield.ac.uk

[2]Brunel University, Department of Design, UK

[3]Marshall Associates, Havant, UK

RUNNING HEAD: Predicting Design-Induced Error

Abstract

Human factors certification criteria are being developed for large civil aircraft. The objective is to reduce the incidence of design induced error on the flight deck. Many formal error identification techniques currently exist, however none of these have been validated for their use in an aviation context. This paper evaluates SHERPA (Systematic Human Error Reduction and Prediction Approach) as a means for predicting design induced pilot error. Since SHERPA was developed for predicting human error in the petrochemical and nuclear industries, a series of validation studies have suggested that it is amongst the best human error prediction tools available. This study provides some evidence for the reliability and validity of SHERPA in a flight deck context and concludes that it may form the basis for a successful human error identification tool.

Keywords

Design induced error/ Error Prediction/ Flight deck design/ Reliability/ Validity

1.0 Introduction

For the past 50 years there has been a general decline in the commercial aircraft accident rate, however during the last decade the serious accident rate has remained relatively constant at approximately one per million departures [3]. With the projected increase in the demand for air travel, if this rate remains unchanged, by 2015 there will be one major hull loss per week. As the reliability and structural integrity of aircraft has improved the number of accidents directly resulting from such failures has reduced dramatically, hence so has the overall number of accidents and the accident rate. However, human reliability has not improved to the same degree during this period. Figures vary slightly but it can be estimated that up to 75% of all recent aircraft accidents now have a major human factors component. Human error is now the primary risk to flight safety [4].

The roots of human error are manifold and often have complex interrelationships with many aspects of the system of operating a modern airliner. However, during the last 10 years 'design induced' error has become of particular concern to the major airworthiness authorities. The Captain of a modern commercial aircraft is now a manager of flight crew and of complex, highly automated aircraft systems. These systems began to be introduced into airliners during the 'glass cockpit' revolution of the 1980s. The high levels of automation offered a considerable advance in safety over their 'clockwork cockpit' forbearers, however new types of error began to emerge on these flight decks.

This was exemplified by accidents such as the Nagoya Airbus A300-600, the Cali Boeing 757 accident and the Strasbourg A320 accident.

As a direct result of such accidents, the US Federal Aviation Administration (FAA) commissioned an exhaustive study of the pilot-aircraft interface on modern flight decks [8]. The report identified many major flight deck design deficiencies and shortcomings in the design process. There were criticisms of the flight deck interfaces, identifying problems in many systems, such as pilots' autoflight mode awareness/indication; energy awareness; position/terrain awareness; confusing and unclear display symbology and nomenclature; a lack of consistency in FMS interfaces and conventions, and poor compatibility between flight deck systems. The report also heavily criticised the flight deck design process, identifying a lack of human factors expertise on design teams and placing too much emphasis on the physical ergonomics of the flight deck, and not enough on the cognitive ergonomics. Fifty-one specific recommendations came out of the report, including:

'The FAA should require the evaluation of flight deck designs for susceptibility to design-induced flightcrew errors and the consequences of those errors as part of the type certification process'.

In July 1999 the US Department of Transportation assigned a task to the Aviation Rulemaking Advisory Committee to provide advice and recommendations to the FAA administrator to 'review the existing material in FAR/JAR 25 and make recommendations about what regulatory standards and/or advisory material should be updated or developed to consistently address design-

related flight crew performance vulnerabilities and prevention (detection, tolerance and recovery) of flight crew error' [21]. The European Joint Airworthiness Authorities (JAA), as a part of the airworthiness regulatory harmonisation efforts, also subsequently adopted this task. The rules and advisory material being developed as part of this process will be applied to both the Type Certification and Supplemental Type Certification processes for large transport aircraft [19, 20]. In the meantime, in 2001 the JAA issued an interim policy document [9] that will remain in force until the new harmonised human factors regulations encompassed in Part 25 come into force.

Compliance with an airworthiness requirement must be established either through inspection, demonstration, evaluation, analysis and/or test. Some form of formal error analysis will probably be the most feasible way to evaluate formally the pilot interface and demonstrate that the likelihood of 'design-induced error' is as low as is reasonably practicable. Formal error analysis is not new, however it is a novel approach as a means of demonstrating compliance with a certification requirement, and as such faces unique challenges. Any technique used for a formal approval process must be reliable and valid and, for the purposes of certification, the method should also be capable of being used by non-human factors experts within the certification authorities (*e.g.* the certification test pilots). As a corollary, any formal error prediction technique should also be capable of being used by the flight deck design teams to verify that at the early stages of design, their flight deck interface is likely to comply with the human factors certification requirement. Furthermore, any error prediction methodology for the

flight deck must be designed to encompass the specific demands of the environment

The objective of the work reported here was to assess the sensitivity, predictive validity and reliability of one such formal technique, SHERPA (Systematic Human Error Reduction and Prediction Approach) [5] to evaluate its suitability for application in an aerospace context. This technique was selected on the basis of the results of a comparative study of six human error identification techniques [10] in which SHERPA achieved the highest overall rankings on a number of assessment criteria for its performance (comprehensiveness, accuracy, consistency, theoretical validity, usefulness and acceptability), and on the basis of a follow-up study which showed that the method also performed well in predicting subsequent actual errors [2]. Further empirical studies have shown that SHERPA also has acceptable test/re-test reliability [16].

There is great caution and scepticism in the aerospace industry concerning formal methods that purport to produce a probability of error associated with any aspect of crew performance. Furthermore, Advisory Circular AC25.1309-1A [7] also suggests that the reliable quantitative estimation of the probability of crew error is not possible. In this study emphasis is placed upon the identification of potential errors using formal methods, *not* their quantification. Unless all significant errors are identified the probability of error generated by any formal technique will be underestimated [9], so the first requirement must be to assess a technique's potential to identify all possible errors resulting from poor human factors design.

To achieve the objectives described, this study progressed in three stages. Firstly, the SHERPA technique was applied to the task of conducting an approach and landing in a modern, highly automated, glass cockpit commercial airliner (Aircraft X) to produce predictions of the errors likely to occur. This was done on two occasions in order to assess the test/re-test reliability of analysts.

Secondly (and independently) using a questionnaire, low-level error data were collected from flight crew currently flying Aircraft X concerning the errors that they had made during the approach and landing flight phase. A fundamental problem when validating formal error identification techniques is obtaining ecologically-valid and reliable criterion data [13]. Accidents are very infrequent events and investigation reports do not contain sufficient detail to establish the design-induced errors that may have contributed to the sequence of events. Incident data are much more plentiful, however, these reports contain even fewer details about the pilots' interactions with their equipment. As a result, a self-completion questionnaire had to be used for this task.

Once the above two data collection stages were completed, the final stage was to compare the error data with the predictions made by SHERPA, using a signal detection paradigm (*q.v.* references 2, 16) to assess the predictive validity of the method.

2.0 Systematic Human Error Reduction and Prediction Approach (SHERPA)

SHERPA [5] (also known as PHEA – Predictive Human Error Analysis [6]) uses a Hierarchical Task Analysis (HTA) [1] in conjunction with an error taxonomy to identify credible errors associated with a sequence of human activity. SHERPA is from a family of human error identification tools that takes a psychologically-based approach [11] hence is ideally suited for use on the flight deck, especially when it is considered that flight deck design has been criticised for its lack of due consideration of cognitive ergonomics [8]. The technique has been used successfully in a variety of non-aviation, safety critical situations, for example in the handling and transportation of hazardous chemicals [12] and offshore oil and gas exploration and exploitation [17]. Non-safety critical applications of SHERPA include the analysis of the usability of ticket vending machines [2] and of car radio/cassette players [18].

The technique operates by trained analysts making judgements about which error modes are credible for each task step based upon this analysis of the work activity. A SHERPA proceeds through six basic steps:

- **Hierarchical Task Analysis**: A hierarchical task analysis is conducted for the task or scenario under analysis.
- **Task Classification**: Each bottom level task in the HTA is taken in turn and classified into one of the five behaviours from the SHERPA behaviour taxonomy.
    - **Action** (e.g. pressing a button or pulling a switch): Errors in this

category are classified into: operation too long/short; operation mistimed; operation in wrong direction; operation too little/much; misaligned; right direction on wrong object; wrong operation on right object; operation omitted; operation incomplete; wrong operation on wrong object.

- **Retrieval** (e.g. getting information from a screen or manual): Errors in this category are classified into: information not obtained; wrong information obtained; information retrieval incomplete.
- **Checking** (e.g. conducting a procedural check): Errors in this category are classified into: check omitted; check incomplete; right check on wrong object; wrong check on right object; check mistimed; wrong check on wrong object
- **Selection** (e.g. choosing one alternative over another): Errors in this category are classified into: selection omitted; wrong selection made.
- **Information communication** (e.g. talking to another party): Errors in this category are classified into: information not communicated; wrong information communicated; information communication incomplete.

- **Human Error Identification**: After each task is classified into a behaviour, the analyst considers the error modes associated with that behaviour. A credible error is one the analyst judges to be possible. For

example, if the task step is classified as an 'Action' the analyst takes the associated 'Action' error modes and considers whether any are credible for that task. For any credible error types, the analyst describes the nature of the error, notes down the associated consequences, highlights whether it is recoverable or not and suggests possible remedial measures.

- **Consequence Analysis**: The analyst considers the consequences of each identified error, which has implications for the criticality of the error.
- **Recovery Analysis:** If there is a task step at which the error can be recovered this is entered next. If no recovery step is possible then 'None' is entered.
- **Tabulation**: The information gained from the SHERPA analysis method is converted into a tabular output. The ordinal probability (P) of an error is then categorised as low (hardly ever occurs); medium (has occurred once or twice) or high (occurs frequently). The same process is applied to the criticality (C) of the error.

The main strengths of the SHERPA method are that it provides a structured and comprehensive approach to error prediction, gives an exhaustive and detailed analysis of potential errors and the error taxonomy prompts the analyst for potential errors, however SHERPA is somewhat repetitive and time consuming to perform.

3.0 SHERPA Analysis

An HTA of a fully-coupled autoland approach to New Orleans airport undertaken in Aircraft X was performed. This consisted of some 22 subtasks under the main headings of setting up for approach, lining up for the runway, and preparing the aircraft for landing. The approach and landing considered was completely normal with no non-routine procedures included. This task analysis formed the basis of the following formal error prediction analysis. An extract of this HTA is included in figure 1.

INSERT FIGURE 1 ABOUT HERE

Eight graduate engineering participants aged between 22 and 55 years took part in this study. All participants were trained in the SHERPA methodology. The training comprised an introduction to the key stages in the method and a demonstration of the approach using a non-aviation example, using an in-car task (see reference [18]). Participants were then required to apply the method to another non-aviation task (described in reference [16]) under the guidance of the instructors. The purpose of this was to ensure that they had understood the workings of the SHERPA method. A debriefing followed, where participants could share their understanding with each other. When the instructors were satisfied that the training was completed, the main task was introduced. This required participants to make predictions of the errors that pilots could make in the autoland task.

To make their error predictions, participants were given an HTA of the autoland task developed by the authors; a demonstration of performing an autoland using Microsoft™ flight simulator; the SHERPA error taxonomy; and colour photographs of Aircraft X's Flight Control Unit, flap levers, landing gear lever, speed brake, primary flight displays, and an overview of the flight deck. An example of the nature of the errors predicted by SHERPA for one task step in the HTA is provided in table 1.

INSERT TABLE 1 ABOUT HERE

Participants were required to make predictions of the pilot errors on two separate occasions, separated by a period of four weeks. This enabled intra-analyst reliability statistics to be computed. The predictions made were compared with error data reported by pilots using autoland (as described in the following section). This enabled SHERPA's validity coefficient to be computed.

4.0 Collection of Error Data

From the approach and landing HTA for Aircraft X a list was compiled of all the possible errors that could be made during the landing phase of flight using the Flight Control Unit (FCU) as the main controlling interface. A few additional systems were also included such as the speed brake and flaps. An additional list of possible errors was developed from observations made during a series of orientation flights on Aircraft X and comments from further interviews with type

rated pilots. These data were used to develop a design induced error questionnaire specific to Aircraft X.

The objective of the questionnaire was to survey pilots to obtain a comprehensive picture of the low-level errors they knew had ever been made on the flight deck while flying a fully-coupled autoland approach and landing in the particular type of aircraft. To achieve this, respondents were not only asked if they had ever made the error themselves but also if they knew of a fellow pilot who had made the same error. A simple 'yes/no' response format was used. As it was highly probable that the list of questions was not exhaustive, space was provided to report additional errors or for further comments to be given.

Following an initial pilot administration of the instrument to a small sample of senior pilots to check for major errors and to refine the wording of the survey items further, the final questionnaire was sent to pilots flying Aircraft X in three UK Airlines.

The final instrument contained 70 questions specifically concerned with the flight deck interface. The survey was divided into 13 subsections, with items regarding speed brake setting (7 questions); flap selection (10 questions); lowering landing gear (1 question); airspeed (11 questions); checking ALT is engaged (1 question); altitude (8 questions); changing headings (4 questions); checking HDG is engaged (1 question); engaging the approach system (4 questions); checking APPR is engaged (1 question); tracking the localiser (7 questions); tracking the glideslope (2 questions); and other miscellaneous items (13 questions).

After the return of the questionnaires, a number of additional interviews were conducted to clarify the additional comments received on many of the survey instruments.

5. 0 Error Data Results

5.1 Sample

Forty-six completed questionnaires were returned in time for analysis. Of the 46 respondents the majority were Captains (45.7%). First Officers provided a further 37% of the sample with the remainder being either Training Captains (13.3%) or they failed to state their position (two respondents). The number of total flying hours ranged from less than 2,000 hours to over 16,000 with a mean of 6,832 hours and a standard deviation of 4,524 hours. With regard to type specific experience this ranged from less than 1,000 hours to over 5,000 hours. The mean time on type was 1,185 hours with a standard deviation of 1,360 hours).

5.2 Survey Results

A total of 57 different types of error were reported, either as responses to the structured survey items or in the additional comments section of the questionnaire. For brevity, in table 1 only the data from the survey items where more than two respondents reported a particular error are presented. In table 2, the column 'ME' contains the frequency data where the respondent has made the error in question themselves; the column labelled 'OTHER' contains the data which indicates that

they have seen someone else make the error or that they are aware of someone who has.

INSERT TABLE 2 ABOUT HERE

5.3 Predictive Validity and Reliability of SHERPA for Predicting Pilot Error

Overall, the analysts using SHERPA predicted 56 errors. These predictions were compared with error data collected in the survey of pilots. This enabled validity statistics to be computed using a signal detection paradigm [14]. This approach provides a useful framework for testing the power of formal human error identification methods [2, 16]. In addition to comparing correct predictions of error with actual errors (hits) it identifies type I analytical errors (a miss: when the error analyst predicts the error will not occur and it does) and type II analytical errors (a false alarm: when the error analyst predicts that there will be an error and there is not). The results of this analysis are presented in table 3.

INSERT TABLE 3 ABOUT HERE

The overall hit rate was very high. Of the 57 errors reported by flight crew, 52 were predicted by one or more of the analysts. Four errors were predicted which were not yet reported by any of the pilots surveyed. The validity coefficient, expressed as a mean of the 'hit' and 'false alarm' rates was *circa* 0.6. This value is toward the lower end of being acceptable. However, if the error predictions across the participants are pooled, the validity statistic rises to *circa* 0.9, which is exceptionally high.

The test/re-test reliability of analysts between the first and second applications of the SHERPA methodology (assessed using Pearson's r) was *circa* 0.7. This value is moderate, but it should be noted that this was the first time the participants had applied the SHERPA and they were not experts in either human factors or from the aerospace industry.

The complete results for all analysts (Hits, False Alarms, Misses, Correct Rejections and Sensitivity calculations) are included in table 4. The formula for calculation of the Sensitivity Index is included at equation 1.

INSERT TABLE 4 ABOUT HERE

INSERT EQUATION 1 ABOUT HERE

6.0 Discussion

The identification of errors made on the flight deck committed as a result of poor design is difficult. In many cases, even in accident and incident reports, there is a paucity of data on the low-level errors made that may have contributed to the accident sequence. As a result of the scarcity of these data the validation of formal error prediction methods in this context is problematic.

The questionnaire survey conducted in this study does, however, suggest that design-induced errors are an everyday issue for the pilots of modern airliners. Pilots who responded to the survey reported that between them, 57 errors related to the design of the pilot interface occurred in just the approach and landing phase of flight. Simply because many of these errors do not appear in accident and

incident reports does not make them unimportant issues. At best, these aspects of flight deck design that encourage error represent a daily source of frustration for pilots, as they have to remain vigilant to their own mistakes which they then need to correct. At worst, it is possible that the types of errors identified may be significant factors in accidents and incidents. It is simply too difficult, however, to link them to the accident sequence with any degree of certainty.

Many of the potential errors that were identified in the SHERPA analysis were the types of errors that most pilots were aware of and have simply had to accept on the flight deck during everyday operations. It is hoped that the development of human factors certification standards will help to ensure that many of the design induced errors identified are eradicated in future aircraft. The initial results using SHERPA for this type of task are promising. Whilst more studies are needed to investigate more flight tasks in different phases of flight, the current study shows that novice analysts were able to acquire the approach with relative ease and reach an acceptable level of performance within a relatively short period of time, which supports the results of previous studies [16]. This is essential given the nature of how, and by whom, the technique may be applied in both the flight deck design and certification context.

The validity coefficient for the non-pooled analyst data at first seems a little disappointing however this is unlikely to be a major problem within a certification programme. Firstly, Analyst skills improve with familiarity with the technique [16]. A previous study has shown that there is very little change over time in the frequency of ‘hits’ and ‘misses’ however, the frequency of ‘false alarms’ falls

over time and consequently, the frequency of correct rejections increases. In the present study a slightly different pattern of results was noted. There was a trend for 'hits' to increase at the second attempt but 'misses' to decrease. In contrast to the previously reported study, though, one of the consequences of an increased 'hit' rate was also a moderate increase in 'false alarms' (see table 4). However, it is argued that in a safety critical industry, an increase in the 'false alarm' rate is perfectly acceptable if it is accompanied by a concomitant increase in the 'hit' rate, as the latter is what really matters. Secondly, it is unlikely that any aspect of certification will be accepted on the basis of a single analyst. It has also been shown that the error identification rate goes up when multiple analysts are used as in this case.

The results indicate that existing human error identification techniques developed for use in other domains (*e.g.* Nuclear Power) can be applied with some success in an aviation context. This study suggests that SHERPA seems particularly well suited in this respect. Furthermore, empirical evidence from comparisons of SHERPA with other formal human error prediction techniques [15] strongly suggests that SHERPA is the best of the currently available human error prediction techniques. SHERPA's error taxonomy is well suited to the tasks carried out by a civil aircraft pilot, with 'actions' and 'checks' being the most prominent tasks involved. It is suggested that with suitable development and validation, SHERPA (or a SHERPA-like approach) may be used in the future to predict design-induced error for certification purposes.

Acknowledgements

The authors wish to acknowledge that this research was made possible through funding from the UK Department of Trade and Industry as part of the European EUREKA! Programme. Our thanks go to Gillian Richards and John Brumwell, at the DTI and Richard Harrison (now at QinetiQ) who have provided invaluable support to this research effort. We are most grateful to our European colleagues in the programme, Sidney Dekker from Linköping University, Sweden and Thomas Waldmann of the University of Limerick, Ireland for their help and advice. Our thanks also go to Air2000, British Midland and JMC airlines in allowing us access to their pilots and to the pilots for taking the time to complete the questionnaire.

References

[1] Annett, J., Duncan, K.D, Stammers, R.B., Gray, M.J. Task Analysis. Training Information No.6. HMSO, London, 1971.

[2] Baber, C., Stanton, N.A. Human error identification techniques applied to public technology: predictions compared with observed use, Applied Ergonomics 27 (1996) 119-131.

[3] Boeing Commercial Airplanes Group. Statistical Summary of Commercial Jet Airplane Accidents: Worldwide Operations 1959-1999. Boeing, Seattle WA, 2000.

[4] Civil Aviation Authority. Global Fatal Accident Review 1980-96 (CAP 681). Civil Aviation Authority, London, 1998.

[5] Embrey, D.E. SHERPA: A systematic human error reduction and prediction approach. Paper presented at the International Meeting on Advances in Nuclear Power Systems, Knoxville, Tennessee, 1986.

[6] Embrey, D.E. Quantitative and Qualitative Prediction of Human Error in Safety Assessments. Institute of Chemical Engineers Symposium Series, 130 (1993) 329-50

[7] Federal Aviation Administration. Advisory Circular: System Design and Analysis (AC 25.1309-1A). Federal Aviation Administration, Washington DC, 1988.

[8] Federal Aviation Administration. Report on the Interfaces between Flightcrews and Modern Flight Deck Systems. Federal Aviation Administration, Washington DC, 1996.

[9] Joint Airworthiness Authorities. Human Factors Aspects of Flight Deck Design: Interim Policy Paper INT/POL/25/14. Joint Airworthiness Authorities, Hoofdorp, 2001.

[10] Kirwan, B. Human Reliability Assessment, in, J.R. Wilson and E.N. Corlett (Eds), Evaluation of Human Work. London, Taylor and Francis, 1990, pp. 706-754.

[11] Kirwan, B. Human error identification in human reliability assessment. Part 2: detailed comparison of techniques, Applied Ergonomics, 23 (1992) 371-381.

[12] Kirwan, B. A Guide to Practical Human Reliability Assessment. London, Taylor and Francis, 1994.

[13] Kirwan, B. Validation of three Human Reliability Quantification Techniques – THERP, HEART and JHEDI: Part I- Technique Descriptions and Validation Issues, Applied Ergonomics 27 (1996) 359-374.

[14] Macmillan, N.A., Creelman, C.D. Signal Detection Theory: a user's guide. Cambridge: Cambridge University Press, 1991.

[15] Salmon, P.M., Stanton, N.A., Young, M.S., Harris, D., Demagalski, J.M., Marshall, A., Waldmann, T., Dekker, S. Predicting Design Induced Pilot Error: A comparison of SHERPA, Human Error HAZOP, HEIST and HET, a newly developed aviation specific HEI method, in, D. Harris, V. Duffy, M. Smith and C. Stephanidis (Eds.), Human Centred Computing. Mahwah, NJ. Lawrence Erlbaum Associates, 2003, pp. 567-571.

[16] Stanton, N.A., Stevenage, S.V. Learning to predict human error: issues of reliability, validity and acceptability, Ergonomics 41 (1998) 1737-1756.

[17] Stanton, N.A., Wilson, J.A. Human Factors: Step Change Improvements in Effectiveness and Safety, Drilling Contractor, Jan/Feb (2000) 36-41.

[18] Stanton, N.A., Young, M.S. A Guide to Methodology in Ergonomics: Designing for Human Use, London, Taylor and Francis, 1999.

[19] UK Civil Aviation Authority. Joint Airworthiness Requirements, (JAR 25 – Large Aeroplanes). Civil Aviation Authority, London, 1978.

[20] US Department Of Transportation. Federal Aviation Regulations, (Part 25 – Airworthiness Standards). Revised January 1, 2003. US Department Of Transportation, Washington, DC, 1974.

[21] US Department of Transportation. Aviation Rulemaking Advisory Committee; Transport Airplane and Engine: Notice of new task assignment for the Aviation Rulemaking Advisory Committee (ARAC). Federal Register, Vol. 64, No. 140, July 22 1999.

Table 1

Extract of SHERPA analysis. In the following extract, error mode A3 is an operation in the wrong direction; error mode A6 is an operation in the correct direction but on the wrong object. Estimates (low, medium or high) of the probability of occurrence and criticality of the error are given in columns labelled P and C, respectively.

Task Step	Error mode	Description	Consequence	Recovery	P	C	Remedial measures
3.2.2	A3	Pilot turns the Speed/MACH selector knob the wrong way	The wrong airspeed is entered and the plane speeds up instead of slowing down	3.2.1	M	M	- Clearer control labeling - Auditory signal informing increase/decrease
3.2.2	A6	The pilot dials in the desired airspeed using the wrong control knob i.e. the heading knob	Before capture, the auto-pilot will attempt to switch course to the speed value entered causing the plane to leave the glideslope	Immediate	M	H	- Improved control labeling - Improved separation of controls

Table 2

Percentage of pilots reporting having made (or knowing about) common design induced errors during the approach and landing phase while performing an autoland in Aircraft X

ITEM	ME	OTHER
Speed Brake		
Moved the flap lever instead of the speed brake lever when intended to apply the speed brake	0%	6.5%
Flaps		
Checked the flap position and misread it	4.3%	4.3%
Moved the flap lever further or not as far as intended	17.4%	6.5%
Landing Gear		
Omitted to put the landing gear down until reminded	19.6%	37.0%
Airspeed		
Initially, dialled in an incorrect airspeed on the Flight Control Unit by turning the knob in the wrong direction	39.1%	37.0%
Having entered the desired airspeed, pushed or pulled the switch in the opposite way to the one that you wanted	26.1%	26.1%
Adjusted the heading knob instead of the speed knob	78.3%	65.2%
Altitude		
Entered the wrong altitude on the Flight Control Unit and activated it	15.2%	17.4%
Entered an incorrect altitude because the 100/1000 feet knob wasn't clicked over	26.1%	28.3%
Believed you were descending in flight path angle and found that you were in fact in Vertical speed mode or vice versa.	8.7%	13.0%
Failed to check ALT (Altitude) mode was active	8.7%	6.5%
Heading		
Entered a heading on the Flight Control Unit and failed to activate it at the inappropriate time	34.8%	34.8%
Failed to check HDG (Heading) mode was active	23.9%	19.6%
Approach System		
Tried to engage APPR (Approach) mode too late so that it failed to capture	28.3%	30.4%
Pressed the wrong button when intending to engage APPR such as EXPED (Expedite)	6.5%	8.7%
Failed to check APPR was active	28.3%	30.4%
Localiser		
Incorrectly adjusted heading knob to regain localiser and activated the change	4.3%	4.3%

Table 2 (continued)

Percentage of pilots reporting having made (or knowing about) common design induced errors during the approach and landing phase while performing an autoland in Aircraft X

ITEM	ME	OTHER
Glideslope		
Failed to monitor the glide slope and found that the aircraft had not intercepted it	39.1%	52.2%
Other		
Had an incorrect barometric air pressure set	45.7%	45.7%
Set an altitude ‘out of the way’ and then out of habit pulled the altitude knob	15.2%	32.6%

Table 3

Performance of SHERPA in predicting pilot error during approach and landing using pooled error data

		ERROR OBSERVED?	
		Yes	**No**
ERROR PREDICTED?	**Yes**	*Hits* 52	*False Alarms* 4
	No	*Misses* 5	*Correct Rejections* 179

Table 4

Hits, False Alarms, Misses and Correct Rejections and

SIs for time 1 and time 2

	Time 1					**Time 2**				
Sub	**Hits**	**FA**	**Misses**	**CR**	**SI**	**Hits**	**FA**	**Misses**	**CR**	**SI**
1	71	66	32	203	0.72	51	43	53	226	0.67
2	54	38	50	231	0.69	62	45	42	224	0.71
3	52	38	52	231	0.68	44	27	60	242	0.66
4	62	76	42	193	0.66	98	116	6	153	0.76
5	48	44	116	285	0.58	60	76	44	193	0.65
6	68	42	36	227	0.75	74	51	30	218	0.76
7	83	86	21	183	0.74	80	83	24	186	0.73
8	38	29	66	240	0.63	71	42	33	227	0.76
				Mean SI	**0.68**				**Mean SI**	**0.71**

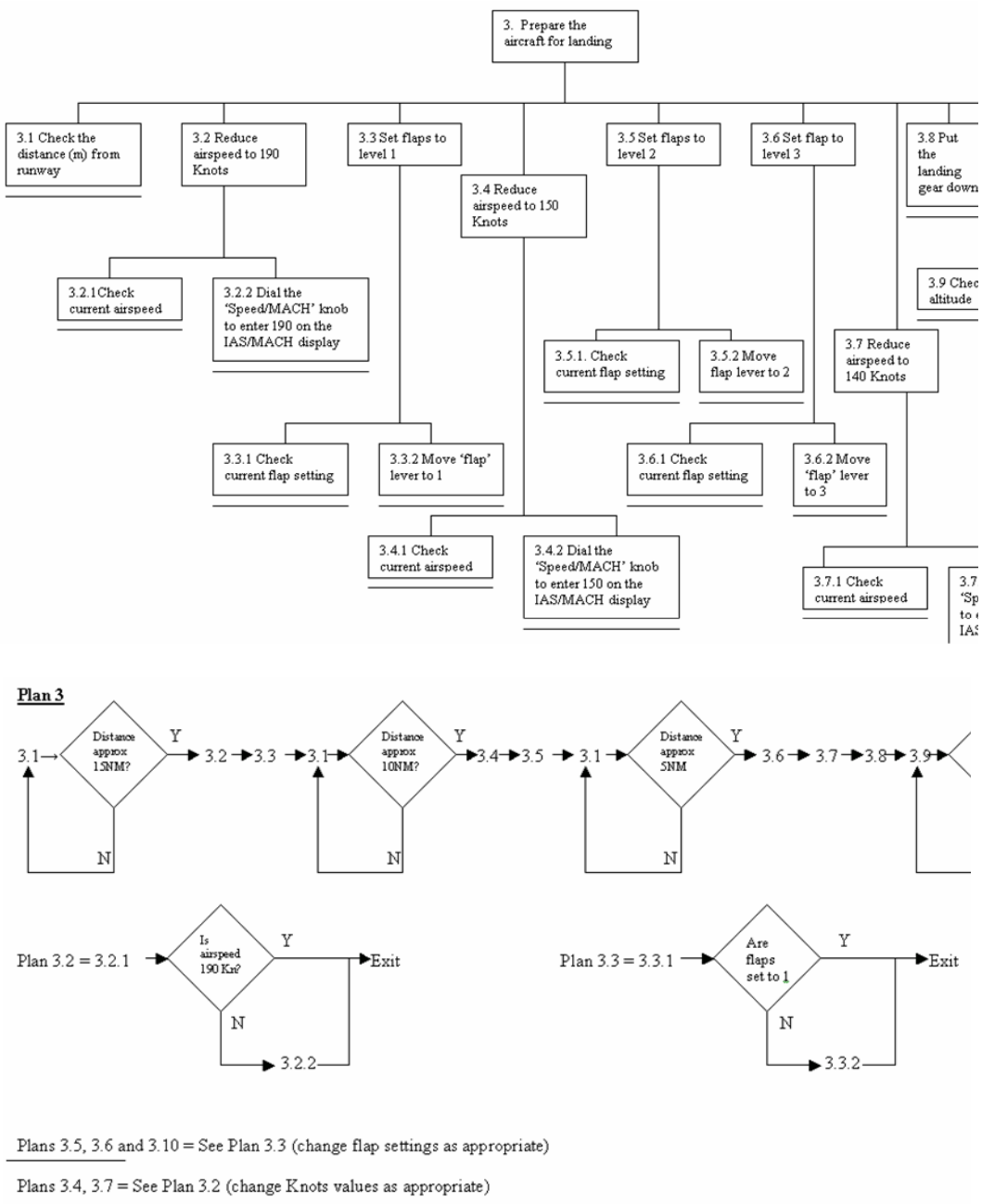
Formula 1

Sensitivity Index formula

$$Si = \left(\frac{\left(\frac{Hit}{Hit + Miss} \right) + 1 - \left(\frac{False\ Alarm}{FA + Correct\ Rejection} \right)}{2} \right)$$

Figure 1

Extract of HTA for landing aircraft (including an extract of plans) upon which error predictions were made





https://dspace.org/

Using SHERPA to predict design-induced error on the flight deck.

Harris, Don; Stanton, Neville A.; Marshall, Andrew; Young, Mark S.; Demagalski, Jason M. ...

2005-09
Elsevier
http://hdl.handle.net/1826/980
Don Harris, Neville A. Stanton, Andrew Marshall, Mark S. Young, Jason Demagalski and Paul Salmon, Using SHERPA to predict design-induced error on the flight deck, Aerospace Science and Technology, Volume 9, Issue 6, September 2005, Pages 525-532.

Downloaded from a DSpace repository. https://dspace.org/