DOT/FAA/AR-10/27

Air Traffic Organization
NextGen & Operations Planning
Office of Research and
Technology Development
Washington, DC 20591

Flight Crew Intervention Credit in System Safety Assessments

Evaluation by Manufacturers and Definition of Application Areas

October 2010

Final Report

This document is available to the U.S. public through the National Technical Information Services (NTIS), Springfield, Virginia 22161.

This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof. The United States Government does not endorse products or manufacturers. Trade or manufacturer's names appear herein solely because they are considered essential to the objective of this report. This document does not constitute FAA certification policy. Consult your local FAA aircraft certification office as to its use.

This report is available at the Federal Aviation Administration William J. Hughes Technical Center's Full-Text Technical Reports page: actlibrary.tc.faa.gov in Adobe Acrobat portable document format (PDF).

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AR-10/27		

4. Title and Subtitle	5. Report Date
FLIGHT CREW INTERVENTION CREDIT IN SYSTEM SAFETY ASSESSMENTS EVALUATION BY MANUFACTURERS AND DEFINITION OF APPLICATION AREAS	October 2010
	6. Performing Organization Code
7. Author(s) R. Wever and A.L.C. Roelen	8. Performing Organization Report No. NLR-CR-1009-618
9. Performing Organization Name and Address NLR Air Transport Safety Institute Anthony Fokkerweg 2 P.O. Box 90502 1006 BM Amsterdam The Netherlands	10. Work Unit No. (TRAIS)
	11. Contract or Grant No.
12. Sponsoring Agency Name and Address U. S. Department of Transportation Federal Aviation Administration Air Traffic Organization NextGen & Operations Planning Office of Research and Technology Development Washington, DC 20591	13. Type of Report and Period Covered Final Report
	14. Sponsoring Agency Code ANM-111

15. Supplementary Notes

The Federal Aviation Administration Airport and Aircraft Safety R&D Division technical monitor was Cristina Tan.

16. Abstract

According to current regulations for type certification of large commercial aircraft, certification credit may be taken for correct and appropriate action for both quantitative and qualitative assessments provided that some general criteria are fulfilled. According to the same regulations, quantitative assessments of the probabilities of flight crew errors are not considered feasible. As a consequence, the system designer is allowed to take 100% credit for correct flight crew action in response to a failure. Previous research indicates that this leads to an overestimation of flight crew performance.

The goal of this research effort was to develop a method that would allow certification credit for good human factors design practice in certification regulations. This method consists of a scoring algorithm that combines key flight deck design characteristics into an overall level of certification credit for flight crew intervention in response to a system failure.

Aircraft manufacturers applied the method to a selection of system failure cases, which provided feedback on the method, characteristics, and application areas.

Seven options for implementation of this method in the design and certification process have been identified. Each option sets particular requirements for further research and development activities.

17. Key Words	18. Distribution Statement
Human factors, Design, Certification, Safety assessment	This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	38	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGEMENTS

This research project was sponsored by the Federal Aviation Administration (FAA) and Civil Aviation Authority (CAA) the Netherlands. The authors are indebted to John Lapointe and Cristina Tan of the FAA William J. Hughes Technical Center, Atlantic City International Airport, NJ; Loran Haworth of the FAA Transport Airplane Directorate in Seattle, WA; and Ron van de Leijgraaf of the CAA in the Netherlands, for overall coordination and support. The authors are very grateful for the voluntary participation and effort by Gulfstream, Embraer, and Fokker to support the evaluation of the method. The European Aviation Safety Agency is acknowledged for facilitating a meeting with the European Human Factors Advisory Group. The authors also are grateful for the opportunity to present this project to avionics and engine system designers from Honeywell, USA.

TABLE OF CONTENTS

			Page
EXECUTIVE SUMMARY			ix
1.	INTRODUCTION		1
	1.1	Background	1
	1.2	The Method	2
	1.3	Objective	3
	1.4	Structure of the Report	3
2.	EVALUATION OF THE METHOD BY INDUSTRY		4
	2.1	Objective	4
	2.2	General Setup of the Evaluation	4
	2.3	Organization of the Evaluation by Gulfstream	4
	2.4	Organization of the Evaluation by Embraer	5
	2.5	Organization of the Evaluation by Fokker	5
	2.6	Results From the Application and Evaluation	6
		2.6.1 Recommended Improvements for the Method	6
		2.6.2 Time Needed to Apply the Method	6
		2.6.3 Application Areas	7
	2.7	Feedback From Authorities	8
3.	OPTIONS FOR IMPLEMENTATION		9
	3.1	Introduction	9
	3.2	Current Acceptable Means of Compliance	9
		3.2.1 Means of Compliance According to AMC 25.1302	10
		3.2.2 Means of Compliance According to AMC 25.1309	11
	3.3	Option for Applying the Method in the Aircraft Design Process	12
		3.3.1 Option 1: The Method as a Design Evaluation Tool	12
	3.4	Options for Applying the Method in 25.1309	13
		3.4.1 Option 2: Evaluation Method to Determine if Full or Zero Credit can be Taken for Flight Crew Intervention	13

3.4.2 Option 3: Evaluation Method in Support of Discussion Between Applicant and Certification Authorities 14

3.4.3 Option 4: Evaluation Method to Determine Quantitative Probability of Flight Crew Intervention 15

3.5 Option for Applying the Method in 25.1302 19

3.5.1 Option 5: Assessment Method for Acceptability of the Design 19

3.6 Options for Applying the Method in Accident Investigations and Flight Test Pilot School Curriculum 19

3.6.1 Option 6: Diagnostic Tool for Accident Investigators 19

3.6.2 Option 7: Implementing the Method Into the FAA Flight Test Pilot School Curriculum 20

4. CONCLUSIONS AND RECOMMENDATIONS 20

4.1 Conclusions 20
4.2 Recommendations 21

5. BIBLIOGRAPHY 22

APPENDICES

A—Recommended Improvements and Comments on the List of Design Characteristics
B—Credit Level per Characteristic

LIST OF TABLES

Table		Page
1	Relationship Between Failure Severity and Minimum Score	14
2	Relationship Between Probability and Severity of the Failure Condition	16
3	Linking the Score With a PSF	17
4	Linking the Score With Flight Crew Intervention Probability	18

LIST OF ACRONYMS

AC	Advisory Circular
AFM	Aircraft Flight Manual
AMC	Acceptable Means of Compliance
CAA	Civil Aviation Authority
CFR	Code of Federal Regulations
CS	Certification Specifications
EASA	European Aviation Safety Agency
FAA	Federal Aviation Administration
FMEA	Failure Modes and Effects Analysis
GIA	Garmin Integrated Avionics Unit
NLR-ATSI	NLR Air Transport Safety Institute
PSF	Performance-shaping factor

EXECUTIVE SUMMARY

According to current regulations for type certification of large commercial aircraft, certification credit may be taken for correct and appropriate flight crew action for both quantitative and qualitative safety assessments provided that some general criteria are fulfilled. According to the same regulations, quantitative assessments of the probabilities of flight crew errors are not considered feasible. As a consequence, the system designer is allowed to take 100% credit for correct flight crew action in response to a system failure. Previous research indicates that this leads to an overestimation of flight crew performance.

The NLR Air Transport Safety Institute (NLR-ATSI) was commissioned by the Federal Aviation Administration (FAA) and Civil Aviation Authorities of the Netherlands to develop a method that would allow certification credit for good human factors design practice in certification regulations. NLR-ATSI developed a method that offers a scheme for evaluating designs based on a set of design characteristics, resulting in a level of certification credit for flight crew intervention in response to a system failure.

Aircraft manufacturers (Gulfstream, Embraer, and Fokker) applied the method to a selection of system failure cases, which provided feedback on the method, characteristics, and application areas.

Based on the input from the manufacturers and further discussion with the FAA and the European Aviation Safety Agency (EASA), application areas of the method were evaluated to show that the method could support the design and development within a manufacturer to encourage good human factors design practices.

The following conclusions were drawn: (1) the evaluation of the method by aircraft manufacturers provided recommendations for improving the method, (2) the manufacturers considered the application of the method in the design and definition phase as beneficial since it provides a structured guidance tool to support the analysis and discussion on human factors aspects of the flight deck design, (3) the method is easy to use and the effort to apply the method is reasonable, and (4) seven options for implementation of this method in the design and certification process have been identified. Each option sets particular requirements for further research and development activities.

With additional development, the method could be applied to the design and development of aircraft and equipment and aircraft certification (e.g., Certification Specifications (CS) and Acceptable Means of Compliance (AMC) 25.1302 or Title 14 Code of Federal Regulations, CS, and AMC 25.1309). In addition, the method could be used in the FAA flight test pilot school curriculum and in support of accident investigations. Further research and development is advised: (1) it is recommended to improve the list of design characteristics with feedback from the manufacturers during the evaluation, (2) the method could also be adapted in view of the Notice of Proposed Rulemaking change to 25.1322 about flight crew alerting, (3) the FAA is recommended to formulate an opinion on the method and the way to proceed with the development of this method, and (4) it is recommended to disseminate the findings of the study to interested parties, such as the FAA small airplane directorate, EASA, and the European Human Factors Advisory Group.

1. INTRODUCTION.

1.1 BACKGROUND.

Current certification requirements for aircraft system safety analysis were initially developed, in approximately 1970, when long-range, wide-body aircraft, such as the Boeing 747, the Douglas DC-10, and the Lockheed L-1011, appeared. These requirements are based on the principle that an inverse relationship should exist between the probability of malfunctions and the degree of hazard to the aircraft and its occupants. A detailed Failure Modes and Effects Analysis (FMEA) and Fault Tree Analysis are often necessary to demonstrate compliance with these requirements.

One way to mitigate the effect of a system failure is by appropriate corrective action by the flight crew. The current Advisory Circular (AC) 25.1309-1A states:

> "When assessing the ability of the flight crew to cope with a failure condition, the warning information and the complexity of the required action should be considered. If the evaluation indicates that a potential failure condition can be alleviated or overcome during the time available without jeopardizing other safety-related flight crew tasks and without requiring exceptional pilot skill or strength, credit may be taken for correct and appropriate corrective action, for both qualitative and quantitative assessments."

The AC also states that quantitative assessments of the probabilities of flight crew errors are not considered feasible. As a consequence, probabilities of either zero or one are entered into the fault trees. Accordingly, the designer is allowed to take 100% credit for correct flight crew action in response to a failure. Since the development of certification regulations in approximately 1970, much research was conducted in the field of human factors, resulting in a better understanding of the effectiveness of various design features for human-machine interfaces. This is reflected in the design of current-generation flight decks and aircraft certification regulation, such as European Aviation Safety Agency (EASA) Notice of Proposed Amendment 15/2004 that proposes to add a new paragraph, Certification Specifications (CS) 25.1302, Human Factors, to the existing airworthiness code. However, good design practice, which would be expected to enhance the reliability, timeliness, or effectiveness of flight crew intervention, is not rewarded in the quantitative analyses for demonstrating compliance with Title 14 Code of Federal Regulations (CFR) 25.1309. Current regulations and associated guidance material do not provide criteria that encourage or require manufacturers to develop and follow a flight deck design process that comprehensively addresses human performance considerations. To provide such an incentive, a methodology is needed that would provide certification credit for desirable design features.

The Air Transport Safety Institute of the Dutch National Aerospace Laboratory (NLR-ATSI) was commissioned by the Federal Aviation Administration (FAA) and the Civil Aviation Authority (CAA) of the Netherlands to conduct research on the development of a method that would allow appropriate levels of certification credit for good human factors design practice in flight decks, for those situations in which flight crews are expected to take action to mitigate system failure conditions. This methodology should establish a rational scheme for evaluating

designs based on easy-to-understand characteristics, eventually resulting in a number that is allowed to be used in the system safety assessment. The methodology should provide several levels of credit to differentiate significantly different designs and should give full credit to what is considered to be best design practice. The method was intended to promote and reward good human factors design practice.

1.2 THE METHOD.

NLR-ATSI developed a scheme for evaluating designs based on a set of design characteristics, resulting in a numerical outcome. The method is described in Roelen and Wever, 2004, and Roelen, et al., 2005. A list of 14 key design characteristics that influence flight crew intervention was developed and is presented in appendix A. The overall numerical outcome is obtained by summation of the scores obtained for each of the 14 key characteristics. Each characteristic can get full, partial, or no score, and the total score is normalized to obtain a value between 0 and 1 (see equation 1). The scores for the different levels of each design characteristic are defined in appendix B.

$$S = \frac{1}{14}\sum_{i=1}^{14} C_i \tag{1}$$

where

S = overall score
C_i = score for an individual design characteristic i; C_i can be 1 (full score), 0.5 (partial score), or 0 (no score)

The numerical outcome of the method, the total score, represents the quality of the flight deck and aircraft systems design in supporting the flight crew in the detection, diagnosis, and mitigating actions to cope with or alleviate a system failure condition. The maximum score that can be obtained is 1 and the minimum is 0. The better the design of the flight deck and the aircraft systems supporting the flight crew in failure detection, analysis, and recovery, the higher the total score. For example, if a system failure condition obtained a total score of 0.8, this is an indication that, from a human factors perspective, the flight deck design is better than the failure case that obtained a score of 0.3. The numbers represent a rank order, they are not probabilities.

Some design characteristics may not be applicable in a particular failure case. If some of the characteristics are not applicable, the sum of the individual scores is still divided by the total number of key characteristics (14) to maintain an equal weight of each characteristic. A nonapplicable characteristic receives a full score to ensure that the applicant obtains a full overall score if each of the remaining (applicable) characteristics receives a full score. Characteristic 13 covers the way in which deferred actions[1] are prompted to the flight crew. If deferred actions are not prompted, there is no score for this characteristic. If deferred actions are

[1] Deferred actions are specific actions, envelope restrictions, or unavailable systems as a direct result of the failure, which applies after the immediate failure procedure has been completed.

prompted, there is a full score for this characteristic. However, there may be cases in which there are no deferred actions. Characteristic 13 is then not applicable but still receives a full score to avoid being unnecessarily penalized.

A situation could also arise where more than one level of score is applicable to a certain design characteristic, for example, if the situation requires manipulation of several controls, but only one of these controls provides feedback (characteristic 14, feedback on control input). In that case, the research team suggests to take a conservative approach and to select the lowest score for that particular characteristic.

Wever and Roelen, 2009, showed that the method could also be used to evaluate a case of a system condition other than a system failure that requires a particular flight crew action, e.g., landing an aircraft that was dispatched with an inoperative thrust reverser. Another use of the method includes assessing the safety effect of design modifications that relate to flight crew intervention.

This method is not intended to provide a final judgement on a particular design, but rather to serve as a starting point for the discussion between the applicant and the certification authorities. Nevertheless, it is considered beneficial if the discussion between the applicant and certification authorities is as efficient as possible and focuses on the important issues. Therefore, every effort was made to make the method as complete, objective, and unambiguous as possible.

1.3 OBJECTIVE.

The objective was to conduct research and development, focusing on activities that further develop the method to maturity for practical implementation in the aircraft certification environment, specifically addressing the method's integration into the system safety assessment and certification process.

1.4 STRUCTURE OF THE REPORT.

Section 2 describes the evaluation of the method with three aircraft manufacturers. In section 3, various options for the practical application and implementation of the method within regulations and an industry environment are described. Section 4 provides the conclusions and recommendations.

Throughout this report, the following regulations will be referred as:

- 25.1302—Certification Specifications (CS) 25.1302 and Acceptable Means of Compliance (AMC) 25.1302

- 25.1309—14 CFR 25.1309, CS 25.1309, and AMC 25.1309

- 25.1322—14 CFR 25.1322 as defined in FAA Notice of Proposed Rulemaking published July 9, 2009, Docket No. FAA 2008-1292; Notice No. 09-05.

2. EVALUATION OF THE METHOD BY INDUSTRY.

2.1 OBJECTIVE.

One of the objectives of this phase of the research was to apply the method in an industry environment to gain insight in the use of the method and its added value in a system safety assessment and to get feedback from manufacturers about the method. NLR-ATSI approached Gulfstream, Embraer, and Fokker with a request to support the development of the method, to which the manufacturers responded positively.

2.2 GENERAL SETUP OF THE EVALUATION.

The method was presented in the form of a questionnaire on key design characteristics, one question for each characteristic (see appendix A). The employees of the manufacturers that participated in the evaluation applied the questionnaire to selected system failure cases. The selection of system failure cases was done in consultation with the manufacturers for one of their aircraft models: the G550 at Gulfstream, the Phenom 100 at Embraer, and the Fokker 100 at Fokker. The participants provided answers to the questionnaire, determined the related individual scores by means of table B-1 of appendix B, and finally derived an overall score. The following information was used by the participants applying the method and then collected to improve the design:

- The system description
- The failure condition: description of the failure, nature, and location of the failure
- A complete description of the alerting and design characteristics on the flight deck that aid the crew in failure recovery, including inherent characteristics (e.g., vibration, smoke, etc.)
- The failure recovery procedure for the particular system failure (e.g., Aircraft Flight Manual (AFM) procedures)

It was not the intention of the evaluation to rate the design or system failure cases. Therefore, this report does not contain the numerical results of the failure case evaluations. Detailed results of the evaluation are considered proprietary and are not published in this report.

2.3 ORGANIZATION OF THE EVALUATION BY GULFSTREAM.

The NLR research team briefed Gulfstream (Savannah, GA) on the project, the method, and the application of the method. Four system failure cases on the Gulfstream 550 aircraft were selected for the evaluation: engine fire, low cabin pressure, main entry door open, and incorrect aircraft configuration. Within Gulfstream, three safety engineers and three flight test pilots

applied the method to the selected system failure cases. The engineers completed one case each, while the flight test pilots applied the method to all four system failure cases. Next, the NLR research team met with the Gulfstream engineers and a flight test pilot to review the results and the comments and suggestions. Potential application areas were discussed as well. Gulfstream prepared an internal report describing the evaluation process, results of the assessment of the system failure cases, and the recommendations for improvement (Kratchounova, 2009).

2.4 ORGANIZATION OF THE EVALUATION BY EMBRAER.

The NLR research team conducted a 3-day workshop at Embraer (Sao Jose dos Caomps, Brazil). After presenting the project and the method, a demonstration of how to apply the method to a system failure case was given; separate sessions were then organized to evaluate the method. Five systems were selected on the Phenom 100 aircraft: avionics, hydraulics, propulsion system, electrical, and flight control systems. Sessions dedicated to each of these aircraft systems were organized with engineers, flight test pilots, and flight test engineers to apply the method to failure cases of that particular system. The method was applied to nine system failure cases in total: GIA (Garmin Integrated Avionics Unit) 1 failure, GIA 2 failure, blocked pitot 1, battery 1 bus off, generator 1 and 2 failure, flap failure, hydraulic leak, dual-engine failure, and engine 1 failure.

The NLR research team was present during the Embraer personnel's discussion of the failure case and the appropriate answers to the questions related to the design characteristics. However, the method was applied with little to no guidance from the NLR research team.

The Embraer participants used the AFM, system diagrams, and the aircraft's Quick Reference Handbook to collect the necessary information for a failure case. After each system failure case was evaluated, the Embraer teams provided comments and feedback on the characteristics and the method in general.

Note: The Phenom 100 is a 14 CFR Part 23-certified aircraft. Although the method was originally developed for 14 CFR Part 25 aircraft, this difference did not play a role during the evaluation.

2.5 ORGANIZATION OF THE EVALUATION BY FOKKER.

At Fokker, the NLR research team briefed seven system engineers. The method was applied by the system engineers to one system failure case on the Fokker 100. The case was evaluated twice: once for a pre- and once for a postmodification design. The Fokker engineers discussed the failure case and applied the method with little to no guidance from the NLR research team. The AFM was used to support the evaluation. It took the participants approximately 2 hours to apply the method for both cases.

2.6 RESULTS FROM THE APPLICATION AND EVALUATION.

2.6.1 Recommended Improvements for the Method.

The NLR research team prepared a list of questions specifically addressing the advantages and areas for improvements. The manufacturers provided review comments on the list of design characteristics, the scoring algorithm, and the possible application areas. A summary of the most relevant observations is provided below.

In general, the support and response from the manufacturers was positive.

The following recommendations were given by the evaluation participants.

- Some characteristics need refinement to better capture what is really considered important from a human factors or design perspective. Appendix A shows the recommended improvements for each characteristic.

- For some characteristics, the definition should be improved to make the characteristics easier to understand and remove ambiguity.

- The explanation and examples following each characteristic need improvement and expansion so that the applicant gets more guidance.

- It was suggested to add a characteristic that describes how well the failure annunciation and accompanying failure-handling procedure fits in the overall flight deck design philosophy.

- It was suggested that there should be a relationship between a certain required score and the difference in criticality of the situation that results after correct or incorrect completion of the procedure. The bigger the difference, the higher the required score. In general, there was an intuitive feeling that the score should be related to failure probability and severity level, i.e., “higher score for more critical failures.”

2.6.2 Time Needed to Apply the Method.

In general, the amount of time needed for completion was as expected, although Gulfstream found the effort longer than claimed in the briefing package. From the sessions with Embraer, it appeared that once the participants understood the structure of the method and the meaning of the characteristics, they applied the method without any difficulty. In general, the completion of one failure case, including discussion, took about 1 hour.

2.6.3 Application Areas.

According to the manufacturers, the method could be useful in the following areas:

- Early design and definition phase. The method provides a structured guidance tool to support the discussions on the design and human factors aspects of the flight deck. All items that are addressed in this method are already considered in current design practice, but not in such a structured format. With input from safety assessments (e.g., FMEA, functional hazard analysis, preliminary system safety assessment), the method could help to design the flight deck and initial development of the corrective procedure for system failures. This method can also assist in the documentation of the considerations for certain design trade-offs. Additionally, the method is useful for assessing the severity level of a failure by taking into account the score. A failure that obtains a lower score is considered to be more severe than a failure with a higher score, because of the increased likelihood that the flight crew will not successfully complete the corrective procedure.

- An internal "quality control" method. Applying the method in the design phase could raise flags for designs or characteristics that receive a low score. The method would indicate which characteristics need attention and indicate a possible area for improvement. For example, characteristics that receive a 0 score could be evaluated to check if it is feasible to improve the characteristic to a higher score.

- Demonstrating compliance to 25.1309. Certification involves engineering judgement, and this method provides an extra tool to judge the flight crew's ability to cope with a failure condition. The method is seen as a structured way to consider relevant flight deck design aspects and human factors issues. It can be used in certification to demonstrate and document the particular subjects that have been reviewed.

- Demonstrating compliance to 25.1302. If adapted, the method could be used as a tool to conduct human error analysis in the context of 25.1302 and to show compliance with (parts) of 25.1302. This will require expansion of the list of characteristics to enable analyses of all flight crew actions, i.e., not restricted to response to system failures.

The following remarks were made by the manufacturers:

- The application of the method in the context of 25.1309 raised questions about the way in which the current numerical outcome could be used. Will the method set a minimum score that has to be obtained for full credit for flight crew intervention in 25.1309? Is the method providing a numerical outcome or probability that can directly be used in fault tree analysis?

- Validation of the method before being used in certification is important.

2.7 FEEDBACK FROM AUTHORITIES.

The project was presented to members of EASA, the European Human Factors Advisory Group, and the FAA. The following recommendations for improvement were made by these organizations:

- Some characteristics may need more resolution. For example, characteristic #1 on alerting should be reconsidered in view of new regulations on flight crew alerting, 25.1322. The manner in which the alerting is implemented could be addressed.

- It is recommended to match the list of 14 key design characteristics with the manufacturers' human factors design philosophy to check whether and how the characteristics capture the current philosophy and to check for omissions.

The following remarks were made regarding the application of the method:

- As for the application in 25.1309, the method could provide a structure to support decision making with respect to the assignment of full credit for successful flight crew intervention.

- The premise of this study was that there was a need for a method that could yield a credit value between 0 and 1 for flight crew intervention credit. However, the discussion with the FAA showed that the overall numerical outcome from the method was not considered valuable because it did not represent a probability of successful crew intervention. Still, the method itself is considered valuable, especially because of the systematic process in which the flight deck design characteristics are evaluated. The information about which characteristics receive a poor score, and not the overall score, is relevant because it indicates areas for improvement and focuses attention during the certification process.

- This method provides a structure for subjective judgement and can be a useful tool in the system safety assessment process.

- As for the application in 25.1302, the method could be a partial means to show compliance. It could support the applicant in showing how the requirements in the regulations and corresponding advisory material have been met. In particular, it is recommended to link the list of characteristics with the (sub) paragraphs of 25.1302. Because the list of characteristics for 25.1309 is not necessarily the same as the application in 25.1302, it is recommended to review the list of characteristics in view of the latter regulations.

- The use of flight crew performance probabilities is a topic of much debate. According to AC 1309-1A, quantitative assessments of the probability of flight crew error are not considered feasible. On the other hand, with the general shift from compliance-based to performance-based regulations, it is a logical and interesting step to study the use of quantitative flight crew performance probability in the context of 25.1309.

3. OPTIONS FOR IMPLEMENTATION.

3.1 INTRODUCTION.

This section defines the use of the method in the context of flight deck design, system safety assessments, and aircraft system certification regulations. The link between the method and the certification process is also discussed in this section. In section 3.2, the current acceptable means of compliance for certification regulations are explained. Section 3.3 explains the approach to implementing the method options. In sections 3.3 to 3.6, seven options are discussed for applying the method in support of the design and development process (section 3.3), for certification process (sections 3.4 and 3.5), and for flight test pilot school and accident investigation (section 3.6).

Implementation of the method within the current certification regulations should consider possible future changes in those regulations. The research study was initially set up as part of the FAA's intention to explore the possibilities for updating regulations and was seen as an input to the Aviation Rulemaking Advisory Committee.

The method could possibly be applied in the context of two parts of the current regulations concerning the certification of aircraft systems, i.e., 14 CFR Part 25 and CS 25:

- CS 25.1302/AMC 25.1302—Installed systems and equipment for use by the flight crew.

  This section pertains to the design and approval of installed equipment intended for the use of flight crew members from their normally seated positions on the flight deck, and recommendations for the design and evaluation of controls, displays, system behavior, and system integration, as well as design guidance for error management.

- 14 CFR 25.1309 and CS 25.1309/AMC 25.1309—Equipment, systems, and installations.

  This section pertains to the acceptable safety level for airplane equipment and systems in case of system failure conditions.

Using the method outside the direct domain of aircraft certification requirements may be investigated; in particular, the possible added value of using the method during the (initial) design of an aircraft. The following sections discuss in further detail the options for application. Each option sets specific requirements for further research and development.

3.2 CURRENT ACCEPTABLE MEANS OF COMPLIANCE.

This section contains a concise description of the current acceptable means of compliance with certification regulations. EASA's AMC provides guidance material for demonstrating compliance with CS 25.1302 and CS 25.1309.

3.2.1 Means of Compliance According to AMC 25.1302.

The following citation from AMC 25.1302 lists "calculation and engineering analysis" as possible means of compliance:

> Calculation and Engineering Analysis (paragraph 6.3.3): As one possible means of showing compliance with CS 25.1302(d), an applicant may document means of error management through analysis of controls, indications, system behaviour, and related flight crew tasks. This would need to be done in conjunction with an understanding of potential error opportunities and the means available for the flight crew to manage those errors. In most cases it is not considered feasible to predict the probability of flight crew errors with sufficient validity or precision to support a means of compliance. If an applicant chooses to use a quantitative approach, the validity of the approach should be established.
>
> Calculation and Engineering analysis is defined by the following characteristics [Paragraph 6.3.3, page 2-F-26]:

Description
Calculations or engineering analyses ("paper and pencil" assessments) that do not require direct participant interaction with a physical representation of the equipment.
Deliverable
Report detailing the analysis, its components, evaluation assumptions, and basis for decision making. The report details results and conclusions.
Participants
Conducted by the applicant.
Conformity
Not applicable.
Uses
Provides a systematic evaluation of specific or overall aspects of the human interface part of the product/system/flight deck. May be specified by guidance material.
Limitations
Carefully consider the validity of the assessment technique for analyses not based on advisory material or accepted industry standard methods. Applicants may be asked to validate any computational tools used in such analyses. If analysis involves comparing measured characteristics to recommendations derived from pre-existing research (internal or public domain), the applicant may be asked to justify the applicability of data to the project.
Example
An applicant may conduct a vision analysis to demonstrate that the flight crew has a clear and undistorted view out the windows. Similarly, an analysis may also demonstrate that flight, navigation and powerplant instruments are plainly visible from the flight-crew member station. The applicant may need to validate results of the analysis in ground or flight test.

3.2.2 Means of Compliance According to AMC 25.1309.

According to AMC 25.1309, compliance with CS 25.1309(b) should be shown by analysis and, where necessary, by appropriate ground, flight, or simulator tests. Analysis is further described in AMC 25.1309 section 9.b.(5) (page 2-F-46) and in section 12a (page 2-F-53).

AMC 25.1309 states in section 9.b.(5):

> "(i) Where an analysis identifies some indication to, and/or action by, the flight crew, cabin crew, or maintenance personnel, the following activities should be accomplished:\
>
> Verify that any identified indications are actually provided by the system.
> Verify that any identified indications will, in fact, be recognised.
> Verify that any actions required have a reasonable expectation of being accomplished successfully and in a timely manner.
>
> (ii) These verification activities should be accomplished by consulting with engineers, pilots, flight attendants, maintenance personnel and human factors specialists as appropriate, taking due consideration of the consequences if the assumed action is not performed or mis-performed.
>
> (iii) In complex situations, the results of the review by specialists may need to be confirmed by simulator or flight tests. However, quantitative assessments of the probabilities of crew or maintenance errors are not currently considered feasible. If the failure indications are considered to be recognisable and the required actions do not cause an excessive workload, then for the purposes of the analysis, the probability that the corrective action will be accomplished, can be considered to be one. If the necessary actions cannot be satisfactorily accomplished, the tasks and/or the systems need to be modified."

In addition, AMC 25.1309 section 12 "Operational and Maintenance Considerations" mentions:

> "This AMC addresses only those operational and maintenance considerations that are directly related to compliance with CS 25.1309; [...] However, quantitative assessments of crew errors are not considered feasible. Therefore, reasonable tasks are those for which full credit can be taken because they can realistically be anticipated to be performed correctly when they are required or scheduled. [...].
>
> a. *Flight Crew Action.* When assessing the ability of the flight crew to cope with a Failure Condition, the information provided to the crew and the complexity of the required action should be considered. If the evaluation indicates that a potential Failure Condition can be alleviated or overcome without jeopardising other safety related flight crew tasks and without requiring exceptional pilot skill or strength, credit may be taken for both qualitative and quantitative assessments. Similarly, credit may be taken for correct flight crew performance of the periodic

checks required to demonstrate compliance with CS 25.1309(b) provided overall flight crew workload during the time available to perform them is not excessive and they do not require exceptional pilot skill or strength. Unless flight crew actions are accepted as normal airmanship, they should be described in the approved Aeroplane Flight Manual."

3.3 OPTION FOR APPLYING THE METHOD IN THE AIRCRAFT DESIGN PROCESS.

This section explains option 1—the design process.

3.3.1 Option 1: The Method as a Design Evaluation Tool.

3.3.1.1 Description.

The method could be used during the design stage as a means to assess the quality or acceptability of the design of the flight deck and aircraft systems. The manufacturer may hold flight deck design review meetings during which the flight deck design, flight-warning computer design, and flight deck annunciation and procedures are discussed with respect to failure handling. In such a setting, the method provides a structure for evaluating and benchmarking the flight deck design and system failure cases. It helps the manufacturer to focus on a set of key design characteristics in the review process. By using this method as a sort of design guide for effective flight crew intervention, the key characteristics that play a role in crew intervention can be identified, checked, and discussed in a transparent and structured way. Moreover, applying the method enables a comparison with other system failure cases on the same or other aircraft types. Applying the method helps to document the results of such analyses in a structured way. Besides such a review session, the applicant still has to comply with the certification requirements in 25.1302 and 25.1309.

3.3.1.2 Discussion.

During the design process, when the aircraft and system architecture are evolving, changes in the design of the flight deck can be made relatively easily. In this type of application, the method is used as a tool for systematically addressing relevant aspects of the flight deck design for each potential failure condition. Identified weaknesses can result in redesigning the flight deck or aircraft systems. Design and evaluation is an iterative process.

In this type of application, the individual results for each of the 14 design characteristics are the most valuable, as they indicate possible weaknesses in the design. The method is most valuable if the users are also required to document the rationale for each individual result of the evaluation of characteristics. The overall numerical outcome of the method is of lesser significance.

3.3.1.3 Proposed Next Steps.

If this option is to be pursued further, additional evaluation sessions with aircraft and systems manufacturers should be organized, similar to the sessions that were held with Embraer, Gulfstream, and Fokker. Involving equipment manufacturers, such as Airbus, Boeing,

Bombardier, Cessna, and Honeywell, will expand the feedback and foundation of the method. The follow-up step is to develop the method into a "toolkit" that can be further tailored to the needs of the industry. The result could be published as a manufacturing best practice, similar to other manufacturing best practice documents that have been developed and provided by the FAA and/or aerospace industry manufacturers, associations, organizations, and working groups. The practices outlined in these documents are designed for voluntary use by anyone in the aviation community. The objective is to promote continuous improvement.

3.4 OPTIONS FOR APPLYING THE METHOD IN 25.1309.

This section explains options 2 through 4—the certification process.

3.4.1 Option 2: Evaluation Method to Determine if Full or Zero Credit can be Taken for Flight Crew Intervention.

3.4.1.1 Description.

The current AMC 25.1309 allows the applicant to take either full or zero credit for successful flight crew intervention. The method could be used as a means of compliance to determine whether the applicant can take full or zero credit. The advantage is that the method enables a structured analysis of the design characteristics and helps to specify the general qualifications listed in AMC 25.1309, such as complexity of the action, exceptional pilot skills, and information provided.

The method provides a structured analysis of the flight deck's ability to support the flight crew in coping with a failure condition. The outcome of the method is a numerical value. This value in combination with a pass-fail criterion will determine whether the applicant can take full credit for flight crew intervention. The pass-fail criterion is a (minimum) score that must be obtained to take full credit (1), otherwise the applicant should take zero credit (0). For example, the pass-fail criterion could be set at a value of 0.6. If for a certain failure case, the numerical outcome of the method is equal to or greater than 0.6, the applicant is allowed to take full credit for the correct flight crew intervention. If the numerical outcome is less than 0.6, the applicant cannot take credit. The applicant should then decide whether this is acceptable or the design should be modified to obtain a result greater than 0.6 to be able to take full credit for flight crew intervention.

3.4.1.2 Discussion.

The biggest issue with this type of application of the method is the question of defining the value for the pass-fail criterion. This is basically the regulator's task as it should reflect the level of safety of current and future designs that is demanded by society. Nevertheless, from a research perspective, several suggestions can be provided on how to define such a pass-fail criterion. These are given below.

Intuitively, it seems logical to link the value of the pass-fail criterion to the criticality of the situation that would result from the absence of crew intervention or from incorrect crew

intervention. Failures where absent or incorrect crew intervention has a catastrophic effect should then obtain a higher score than failures that only result in minor effects, as shown in table 1.

Table 1. Relationship Between Failure Severity and Minimum Score

System and Crew Failure Condition Criticality	Minimum Required Score*
No safety effect	Not applicable
Minor	>0.2
Major	>0.4
Hazardous	>0.6
Catastrophic	>0.8

*Values are shown here for illustration purposes only.

Applying the method implies that only the overall result is important and not how it was obtained with respect to which of the characteristics scored a 1 on an individual basis (all characteristics are equal weight). The development of the method was initially based on the detection-decision-action framework, and it could be argued that all three phases are equally important. To avoid the possibility that absence of good features related to failure detection can be compensated with features related to action, it would be necessary to have three pass-fail criteria; one for each phase. A distinction should then, obviously, be made between key characteristics related to detection, decision, and action so that different scores can be calculated for each phase and compared with the pass-fail criteria. A disadvantage of this approach is that it further complicates the method.

3.4.1.3 Proposed Next Steps.

If this option is to be pursued further, a threshold value needs to be defined. Essentially, this decision is made by the regulator, not the researchers. However, to lay the foundation for such a decision, a baseline score, i.e., a score that is representative for the state of the art, should be determined. This baseline score can be determined by analyzing the correlation between inappropriate and appropriate crew response on one hand, and the credit score for the corresponding system failure on the other hand. This can be done by reviewing accidents, incidents, and occurrences involving various system failures, while determining the appropriateness of the flight crew response and the corresponding score.

3.4.2 Option 3: Evaluation Method in Support of Discussion Between Applicant and Certification Authorities.

3.4.2.1 Description.

The method can be used by the applicant to claim certification credit for a successful flight crew intervention in a discussion with the certification authorities. The difference between option 2

and option 3 is that there is no formal, defined pass-fail criterion in option 3. In this case, the method is not intended to provide a final judgement on a particular design, but serves as a starting point for the discussion between the applicant and the certification authorities. Nevertheless, it is considered beneficial if the discussion between the applicant and certification authorities is as efficient as possible and focuses on the important issues. At a minimum, the method serves as guidance to identify the issues that need to be addressed in the evaluation. The certification authorities determine—based on expert judgement and certification experience—whether the applicant has sufficient basis for claiming full flight crew intervention credit.

3.4.2.2 Discussion.

This type of application takes place during the design process and, as such, is somewhat similar to option 1. The difference is that, in option 3, the results are shared with the regulator, whereas in option 1, the results are kept internally by the system developer or aircraft integrator. The list of key characteristics is used in option 3 as a list of "agenda items" in the discussion between the applicant and the certification authorities. The benefit lies in streamlining the discussion and evaluation along the list of key characteristics. The method provides structure, supports documentation, and enhances reproducibility of the results. The potential impact of this type of application depends on the moment in the design process that the applicant decides to meet with the regulator. If the architecture is still evolving, the effect of applying the method is potentially larger than at a late stage in the design when the architecture is frozen and changes, at this point, would be expensive.

In this application, as with option 1, the numerical overall score is less relevant. The value of the method lies in the structured evaluation of the characteristics and the scores for the individual characteristics, because the latter will show potential areas for improvements for each characteristic.

3.4.2.3 Proposed Next Steps.

If this option is to be pursued further, evaluation sessions should be organized with the certification authorities (FAA and EASA) to assess the usability of the method in the context of certification. The certification authorities should apply the method to a number of test cases to assess the definition of characteristics and explanatory text and to determine to what extent the method matches with the certification process. This activity will provide feedback on the usability of the method by the certification authorities in discussions with the manufacturers on issues related to 25.1309.

3.4.3 Option 4: Evaluation Method to Determine Quantitative Probability of Flight Crew Intervention.

3.4.3.1 Description.

Option 4 uses the method to evaluate and express the ability of the flight crew to cope with a failure condition in the form of a flight crew intervention probability. The probability of (in)appropriate flight crew intervention can be inserted in the quantitative analysis, e.g., in a fault

tree analysis. The combination of failure condition probability and flight crew intervention probability should then meet the required quantitative criteria defined in 25.1309 for a specific severity effect, see table 2.

Table 2. Relationship Between Probability and Severity of the Failure Condition (AMC 25.1309)

Effect on Aeroplane	No effect on operational capabilities or safety	Slight reduction in functional capabilities or safety margins	Significant reduction in functional capabilities or safety margins	Large reduction in functional capabilities or safety margins	Normally with hull loss
Effect on Occupants excluding Flight Crew	Inconvenience	Physical discomfort	Physical distress, possibly including injuries	Serious or fatal injury to a small number of passengers or cabin crew	Multiple fatalities
Effect on Flight Crew	No effect on flight crew	Slight increase in workload	Physical discomfort or a significant increase in workload	Physical distress or excessive workload impairs ability to perform tasks	Fatalities or incapacitation
Allowable Qualitative Probability	No Probability Requirement	<---Probable---->	<----Remote---->	Extremely <----------> Remote	Extremely Improbable
Allowable Quantitative Probability: Average Probability per Flight Hour on the Order of:	No Probability Requirement	<----------> $<10^{-3}$ Note 1	<----------> $<10^{-5}$	<----------> $<10^{-7}$	$<10^{-9}$
Classification of Failure Conditions	No Safety Effect	<-----Minor------>	<-----Major------>	<--Hazardous-->	Catastrophic
Note 1: A numerical probability range is provided here as a reference. The applicant is not required to perform a quantitative analysis, nor substantiate by such an analysis, that this numerical criteria has been met for Minor Failure Conditions. Current transport category aeroplane products are regarded as meeting this standard simply by using current commonly-accepted industry practice.					

A classical approach in human error quantification, as applied in nuclear engineering, is to use basic error probabilities for general types of tasks. Examples of generic tasks include selecting a control, operating a control, etc. The values of these basic error probabilities are obtained from standard tables or reference books that are agreed upon within the industry. The basic error probabilities are then modified to account for specific circumstances or contexts, such as the presence of workload, the operator's experience, and the quality of the man-machine interface. These modification factors are called performance-shaping factors (PSF). Two of the most commonly used techniques that follow this approach are Human Error Assessment and

Reduction Technique (Kirwan, 1994) and Technique for Human Error Rate Prediction (Swain and Guttmann, 1983).

In analogy with this classical approach for human reliability assessment, the numerical outcome of the method could be used to determine the PSF's value for flight crew intervention probability. This approach would take into account that the flight deck design is one of the many (performance-shaping) factors that influence appropriate flight crew intervention.

Table 3 shows an example of linking the score from the method with a PSF that represents the quality of the flight deck that supports the flight crew in coping with the system failure.

Table 3. Linking the Score With a PSF[1]

Flight Crew Intervention Score	PSF for Flight Deck Design[2]
< = 0.2	5
>0.2 <= 0.6	3
>0.6	1
etc.	etc.

[1]This table only shows an example of the mechanism. Establishing an acceptable link between each score and the PSF values will require additional research and consultation with industry and regulatory authorities.
[2]Values are shown here for illustration purposes only.

A slightly different approach would be to assume that, in the context of flight operations, the quality of the flight deck design is the dominant factor for flight crew intervention probability. In that case, the flight crew intervention probability can be linked directly to the numerical outcome of the method, which would result in the scheme shown in table 4.

Table 4. Linking the Score With Flight Crew Intervention Probability[1]

Flight Crew Intervention Score	Probability of Unsuccessful Flight Crew Intervention[2]
< = 0.2	1×10^{-2}
>0.2 <= 0.4	5×10^{-3}
>0.4 < =0.6	1×10^{-3}
>0.6	5×10^{-4}
etc.	etc.

[1]This table only shows an example of the mechanism. Establishing an acceptable link between each score and a flight crew intervention probability will require additional research and consultation with industry and regulatory authorities.
[2]Values are shown here for illustration purposes only.

3.4.3.2 Discussion.

Development of a human reliability method for flight crew action requires determining basic error probabilities for generic flight crew tasks and determining PSF to account for specific circumstances (not only the quality of the flight deck design but also the quality of training, experience, the level of workload, etc.). In addition, the link between the numerical score from the method and the PSF needs to be established (i.e., table 3). The whole concept must be agreed upon by industry and the aviation authorities. This is a tremendous task and certainly not feasible in the short term. Some human factors experts believe that this approach is inherently impossible.

Linking the score directly to a flight crew intervention probability is based on the assumption that the flight deck design is the dominant factor in flight crew error likelihood. This assumption may be questioned, as other factors might be equally important.

Applying the method to determine a PSF for the quality of the flight deck design only makes sense if there is an acceptable method for human reliability assessment in which this PSF can be applied. Currently, such a method does not exist for the flight crew. According to the current AMC 25.1309, quantitative assessments of human error probability are not considered feasible. In view of the transition from traditional compliance-based regulations to performance-based regulations that takes places in several areas of certification and air traffic management, further research into this option is a step in line with current developments.

The discussions with the aviation authorities and industry revealed that linking the numerical outcome of the method to a flight crew probability, directly or indirectly, via a PSF, is a concept that creates some resistance. The main objection regards the feasibility of providing quantitative assessments of flight crew intervention and, in particular, the validation of such a concept.

3.4.3.3 Proposed Next Steps.

If this option is to be pursued, research should be conducted into the feasibility of defining levels of flight crew intervention probability for different scores. This option would initiate a major research effort, which would involve a literature review on existing studies of flight crew error probabilities, simulator experiments, airline operations observations, and validation in a number of test cases.

3.5 OPTION FOR APPLYING THE METHOD IN 25.1302.

This section explains option 5—acceptability of design.

3.5.1 Option 5: Assessment Method for Acceptability of the Design.

3.5.1.1 Description.

The method can, with some modifications, be used for assessing the design acceptability of installed systems and equipment used by the flight crew. Modification of the method would comprise extending the list of design characteristics to characteristics that are important to flight crew performance, in general, and not limited to flight crew performance in response to system failures.

3.5.1.2 Method of Evaluation.

Each section of 25.1302 would be represented by one or more design characteristics on the extended list. The applicant may use the extended list of design characterises as a formal method to demonstrate to the authorities that all sections of 25.1302 are properly addressed.

3.5.1.3 Proposed Next Steps.

If this option is to be pursued further, the list of characteristics needs to be complemented with characteristics that influence flight crew action in normal operations. Each of the characteristics needs to be linked with the relevant paragraph of 25.1302.

3.6 OPTIONS FOR APPLYING THE METHOD IN ACCIDENT INVESTIGATIONS AND FLIGHT TEST PILOT SCHOOL CURRICULUM.

This section explains options 6 and 7—accident investigations and flight test pilot school.

3.6.1 Option 6: Diagnostic Tool for Accident Investigators.

3.6.1.1 Description.

The method allows accident investigators to systematically analyze flight crew response to system failure cases to identify the flight deck design characteristics and the procedures that may have contributed to accidents and incidents involving system failures or system conditions.

3.6.1.2 Proposed Next Steps.

FAA accident investigators should evaluate the use of the method in accident investigations. This will enable further refinement of the method and the explanatory text. The follow-up step is to develop the method into a toolkit that will be tailored to the needs of the FAA accident investigators.

3.6.2 Option 7: Implementing the Method Into the FAA Flight Test Pilot School Curriculum.

3.6.2.1 Description.

The method could be introduced into the curriculum of the FAA Flight Test Pilot School. Flight test pilots are an integral part of the certification process, e.g., by evaluating aircraft handling and flight deck design in response to system failures. The method can become one of the tools used to identify potential flight deck design improvements.

Currently, there is no standardized, systematic procedure for flight test pilots to evaluate the flight deck design. The lack of consistency in flight deck design evaluation is currently seen as a handicap. The method could be used to fill this gap, similar to the way in which the Cooper-Harper rating scale is used for the subjective evaluation of flight-handling characteristics. In similar fashion, the method could be introduced as part of the system development engineers curriculum.

3.6.2.2 Proposed Next Steps.

Meetings should be held with the FAA and its flight test pilot school to discuss implementing the method into the curriculum. Applying the method to a few cases during the flight test and certification process will enable further refinement of the method and explanatory text. The follow-up step is to develop the method into a toolkit that will be tailored to the needs of the FAA flight test pilots.

4. CONCLUSIONS AND RECOMMENDATIONS.

4.1 CONCLUSIONS.

The initial research was to develop a method that would yield different levels of certification credit between 0 and 1 that could be used in a quantitative safety assessment. From the evaluations with manufacturers and discussions with the Federal Aviation Administration (FAA) and the European Aviation Safety Agency (EASA), it is believed that the value of the method lies more in the structured approach for evaluating flight deck design characteristics than in the determination of a quantified credit level. Although the overall score is not relevant in the current safety assessment, the scores for the individual characteristics will help the FAA identify potential areas in the design that need attention during the certification and discussions with the applicant. Furthermore, the method could be used by the FAA flight test pilot school and accident investigators and perhaps lead to the development of best practices.

Based on the results of the analysis described in this report, the following conclusions were drawn:

- The evaluation of the method by aircraft manufacturers provided recommendations for improving the method, e.g., with respect to the definition of characteristics and the explanatory text.

- The manufacturers considered applying the method in the design and definition phase beneficial since it provides a structured guidance tool to support the analysis and discussion on the design and human factors aspects of the flight deck. It helps to identify flight deck design and human factors issues that may need further analysis or improvement.

- From the evaluation, it was concluded that the method is easy to use and the effort to apply the method is reasonable (1 hour per failure case) once the users are familiar with the construct and list of characteristics.

- Seven options for implementing the method into the certification and design process were identified. These options comprise using the method during accident investigations and as part of the curriculum of flight test pilots. Each application area sets particular requirements or directions for further research and development activities.

4.2 RECOMMENDATIONS.

Based on the results of the analysis described in this report, the following recommendations were considered important:

- It is recommended to improve the list of design characteristics according to the suggestions and recommendations made by the manufacturers during the evaluation.

- As a result of the proposed change in the 25.1322 rule[2] about Flight Crew Alerting, it was recommended to review the method to determine the effect of the new rule on the design characteristics dealing with annunciation. If there is an impact, modifications of the affected characteristics should be considered.

- It was recommended that the FAA formulate a viewpoint on the method and determine how to proceed with the development of this method. The FAA should decide which application area(s) will be the focus of future research and development. Based on experience and feedback, options 1, 3, and 7 are the recommended areas for future research. Although this deviates from the original goal for the research project, it is believed that these options would support FAA most in the certification process and certification flight testing, as these applications contribute to good human factors design practice.

[2] FAA Notice of proposed rulemaking, published July 9, 2009, Docket No. FAA 2008–1292; Notice No. 09–05.

- It was recommended to disseminate the findings of the present study to interested parties, such as other branches within the FAA (e.g., small airplane directorate), EASA, and the European Human Factors Advisory Group.

5. BIBLIOGRAPHY.

Kirwan, B., 1994, *A Practical Guide to Human Reliability Assessment*, Taylor and Francis, London, United Kingdom.

Kratchounova, D., 2009, "Flight Crew Intervention Credit in System Safety Assessment Validation," Report GER-8204, Gulfstream, Savannah, Georgia, USA.

Roelen, A.L.C. and Wever, R., 2004, "Key Design Characteristics That Influence Flight Crew Intervention in Case of System Failures," NLR-CR-2004-480, NLR, Amsterdam.

Roelen, A.L.C., Huson, W.J., and de Reus, A.J.C, 2005, "Key Design Characteristics Related to Inappropriate Response Events," NLR-CR-2005-587, NLR, Amsterdam.

Swain, A.D. and Guttmann, H.E., 1983, "Handbook of Human Reliability Analysis With Emphasis on Nuclear Power Plant Applications," Final report, NUREG/CR-1278-F, SAND80-0200, Sandia National Laboratories, Albuquerque, New Mexico, USA.

Wever, R. and Roelen, A.L.C., 2009, "Analysis of Multi-Function Display Failures and Landings With Inoperative Thrust Reverser on A320 and B737 Aircraft," NLR-2008-627, NLR, Amsterdam.

APPENDIX A—RECOMMENDED IMPROVEMENTS AND COMMENTS ON THE LIST OF DESIGN CHARACTERISTICS

1 Is the system failure reliably, explicitly and timely indicated to the flight crew by means of an aural and/or visual and/or tactile alert?

☐ Yes
☐ No

Comments:
- What is the impact of the NPA 25.1322 on this characteristics?
- This characteristic should reflect more than just whether or not alerting is present; it should be able to provide credit based on the way the alerting has been implemented and should promote good human factors design in alerting design.
- This characteristic should perhaps distinguish between reliable, timely and explicit so that each of those can be rated separately.

Characteristics 1 and 5 seem to have much in common. They could perhaps be combined.
- The scope of 'system failure' should be defined.
- 'Equipment' should be considered as an alternative to the word 'system'.

2 Does the system failure manifest itself to the flight crew through inherent cues and / or aircraft motion cues?

☐ Yes
☐ No

Comments:
- Some inherent cues are good because they unambiguously guide the flight crew in failure detection and diagnosis. These inherent cues should be seen as positive. It should be considered if characteristic 2 should distinguish between ambiguous and unambiguous inherent cues.
- The difference between inherent cues and secondary cues should be explained in relation to characteristic 2.
- It should be made clear if for instance the worst possible scenario should be regarded, or the most likely scenario.

3 Does the system establish suitable priorities, does it vary the level of urgency and does it inhibit alerts depending on the flight phase and current aircraft condition and configuration?

☐ Yes
☐ No

Comments:
- It is insufficiently clear if this characteristic 3 refers to the warning logic in general or specifically to the prioritisation of alerts for the system failure that is analysed.

4 Does the system remove incorrect or potentially misleading information?

☐ The failure does not produce incorrect or potentially misleading information.
☐ Yes, incorrect or potentially misleading information is completely removed.
☐ No, incorrect or potentially misleading information is not or only partly removed.

Comments:
- None

5 Are unambiguous system condition message(s) presented to the flight crew?

☐ Yes
☐ No

Comments:
- It should be clarified if this characteristic also addresses the presentation of (un)available functions.
- The characteristic should be rephrased by adding ‘in relation to the system failure condition’.

6 Is the corrective or compensatory procedure presented to the flight crew?

☐ Yes, the procedure is completely and automatically presented on a display
☐ No, the required procedure is prompted but not automatically presented
☐ No, the procedure neither prompted nor automatically presented

Comments:
- Use of the word ‘prompting’ in the description of characteristic 6 caused confusion. This should be better explained in the text and the accompanying example.
- It should be made clear how to account for ‘memory’ items that are part of the procedure?

7 How is the attention-drawing feature deactivated?

☐ The attention-drawing feature can be deactivated by the flight crew.
☐ The attention-drawing feature will be deactivated automatically after a certain time span.
☐ The attention-drawing feature will be deactivated automatically only after crew action has been successfully completed.

Comments:
- This characteristic should be reviewed in the context of NPRM 25.1322.
- It should be explained that this characteristic refers to the attention-getter and not to the failure message.

8 Does the system provide guidance with respect to the appropriate sequence of actions in the procedure?

☐ Yes
☐ No

Comments:
- This design characteristic should be rephrased. There is a distinction to be made between the presence of a hand guidance feature in the flight deck design and the guidance (by the design) of the pilot along the appropriate sequence of actions.
- It should be made clear that, for the purpose of this characteristic, a procedure is not considered to provide 'guidance'.

9 Does the system provide immediate feedback to the flight crew after an inappropriate action is made?

☐ Yes
☐ No

Comments:
- The phrasing should be improved. It should be made clear that this refers to an explicit indication that the flight crew has made a mistake, instead of an indication that can be used by the crew to infer they have made a mistake.
- In the description of characteristic 9 it should be made clear that omission of (a step in) the procedure is also regarded as 'inappropriate action'.
- For characteristic 9 it was suggested to cater for the criticality of possible inappropriate response. Response from the flight crew can result is a situation that is better, similar or worse than the situation before the crew action.
- Answer "not applicable" could be suitable, in case an inappropriate action is not possible with a certain control or input?

10 Does the procedure contain an evaluation of a particular condition and a subsequent flight crew decision based on the result of that evaluation?

☐ Yes
☐ No

Comments:
- The explanatory notes on characteristic 10 should be expanded. It should not only address if-then statement but also cross-checking and pilot judgement.

11 Does the procedure contain one or more action(s), which require a continuous operation and/or monitoring of controls to such a degree that these action(s) jeopardise other operating and/or monitoring tasks?

☐ Yes
☐ No

Comments:
- None

12 Does the system provide feedback with respect to the system condition during and after completion of the corrective or compensatory procedure?

☐ Yes
☐ No

Comments:
- The explanatory note on characteristic 12 should indicate that ‘feedback on the system condition’ refers to e.g. a synoptic.

13 Does the system automatically prompt deferred actions at the appropriate moment and does it automatically display information that must be remembered as a result of the failure?

☐ Yes
☐ No

Comments:
- It is suggested to explain better what in meant with ‘deferred actions’ in characteristic 10. ‘Deferred actions’ refers to envelope restrictions, unavailable systems, or specific actions but not to obvious and basic airmanship.
- It should be made clear that this characteristic refers to deferred action items that are directly a result of the failure and associated procedure.

14 Is feedback on the control input provided?

☐ Yes
☐ No

Comments:
- None

APPENDIX B—CREDIT LEVEL PER CHARACTERISTIC

Table B-1. Credit Levels for Design Characteristics (Full (1), Partial (0.5), or No (0) Credit)

Characteristic	Credit		Characteristic	Credit
1 Annunciation			**8 Sequential guidance**	
Yes, alert provided	Full		Yes, guidance provided	Full
No alert	No		No guidance	No
			Not applicable	Full
2 Inherent cues			**9 Feedback on inappropriate action**	
Yes, inherent cues present	No		Yes, feedback provided	Full
No inherent cues	Full		No feedback	No
3 Prioritisation			**10 If-then statements**	
Yes, priorities established	Full		Yes, procedure contains if-then	No
No prioritisation	No		No if-then	Full
Not applicable	Full			
4 Removal misleading data			**11 Task interference**	
No misleading data	Full		Yes, task interference	No
Misleading data removed	Full		No task interference	Full
Misleading data not removed	No			
5 Unambiguous			**12 Feedback task completion**	
Yes, unambiguous data	Full		Yes feedback after completion	Full
No, ambiguous data	No		No feedback	No
6 Procedure presentation			**13 Deferred actions**	
Procedure presented	Full		Deferred action prompted	Full
Procedure prompted	Partial		Deferred action not prompted	No
Procedure not presented	No		Not applicable	Full
7 Alert deactivation			**14 Feedback on input**	
Manual deactivation	Full		Yes, feedback on control input	Full
Automatic deactivation	No		No feedback on control input	No
Deactivation after completion	Partial		Not applicable	Full
Not applicable	Full			