Designing a Tool to Assess the Usability of Electronic Flight Bags (EFBs)

DOT/FAA/AR-04/38

DOT-VNTSC-FAA-04-12

Divya C. Chandra

Michelle Yeh

U.S. Department of Transportation

Research and Special Programs Administration

John A. Volpe National Transportation Systems Center

Cambridge, MA 02142

Vic Riley

User Interaction Research and Design, Inc.

Point Roberts, WA 98281

Office of Aviation Research

Washington, DC 20591

September 2004

Notice

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

Notice

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2004	Final Report, September 2004

4. TITLE AND SUBTITLE
Designing a Tool to Assess the Usability of Electronic Flight Bags (EFBs)

5. FUNDING NUMBERS
FAE2-AB083

6. AUTHOR(S)
Divya C. Chandra, Michelle Yeh, Vic Riley*

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
John A. Volpe National Transportation Systems Center
Research and Special Programs Administration
Cambridge, MA 02142-1093

8. PERFORMING ORGANIZATION REPORT NUMBER
DOT-VNTSC-FAA-04-12

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation Federal Aviation Administration
Office of Aviation Research, Human Factors Research and Engineering Division
800 Independence Avenue, SW
Washington, D.C. 20591
Program Manager: Dr. Tom McCloy

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
DOT/FAA/AR-04/38

11. SUPPLEMENTARY NOTES
* User Interaction Research and Design, Point Roberts, WA

12a. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the National Technical Information Service, Springfield, VA 22161

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The Federal Aviation Administration (FAA), system designers, and customers all recognize that Electronic Flight Bags (EFBs) are sophisticated devices whose use could affect pilot performance. As a result, human factors issues have received considerable attention from the EFB community. In addition, the FAA's Advisory Circular (AC) on EFBs (AC 120-76A) identifies a need for evaluating EFBs from a human factors perspective, and contains a list of human factors considerations for review. However, the AC does not specify how to perform EFB human factors evaluations.

This research was directed at developing a tool that could be used by FAA Aircraft Certification Service specialists in the field to conduct structured and comprehensive, yet practical, EFB usability evaluations. Two tools were developed for initial tests, with the expectation that a single tool would eventually emerge. The tools were refined over the course of several tests with prototype commercial EFB systems. In the end, we found that both tools are valuable, but in different ways. In this report, we describe both tools, our procedures for testing the tools, and our methods of processing the resulting data into feedback for the manufacturer.

14. SUBJECT TERM
Electronic Flight Bag, EFB, usability, evaluation, human factors tool, design, system design

15. NUMBER OF PAGES
157

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

This page left blank intentionally.

PREFACE

This report was prepared by the Operator Performance and Safety Analysis Division of the Office of Safety and Security at the Volpe Center. It was completed under the Division's Flight Deck Technology Human Factors program under the sponsorship of the FAA's Human Factors Research and Engineering Division. We would like to thank our FAA sponsor, Tom McCloy, as well as Bill Kaliardos, Colleen Donovan, and the many other FAA staff who have given us feedback and suggestions. Many thanks also to the manufacturers who volunteered their units for testing, and to manufacturers who provided expert feedback. And finally, thanks to the participants for their time and valuable suggestions about our evaluation tool and methods. Vic Riley's involvement was funded by the Volpe Center (DTRS57-03-P-80181).

The views expressed herein are those of the authors and do not necessarily reflect the views of the Volpe National Transportation Systems Center, the Research and Special Programs Administration, or the United States Department of Transportation.

Feedback on this document can be sent to Divya Chandra (Chandra@volpe.dot.gov) or Michelle Yeh (Yeh@volpe.dot.gov). Further information on this research effort can be found at http://www.volpe.dot.gov/opsad/efb.

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

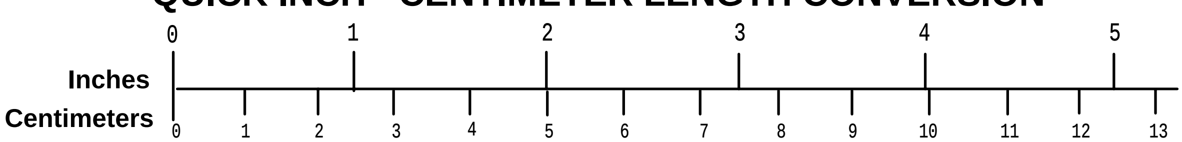
VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSION

°F	-40°	-22°	-4°	14°	32°	50°	68°	86°	104°	122°	140°	158°	176°	194°	212°
°C	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°

For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures.

Price $2.50 SD Catalog No. C13 10286 Updated 6/17/98

Table of Contents

List of Figures

Executive Summary

In March 2003, the Federal Aviation Administration (FAA) issued Advisory Circular (AC) 120-76A on Electronic Flight Bags (EFBs). This AC puts forth a streamlined field approval process for EFBs, contributing to the rapid growth of the EFB industry. In turn, the FAA is being contacted by more applicants seeking approval via the process specified in AC 120-76A.

The goal of this research was to develop a tool to aid FAA Aircraft Certification Service specialists in conducting structured and comprehensive EFB usability evaluations in the field. A significant portion of AC 120-76A (Section 10) addresses human factors considerations for EFBs, but the procedure for performing this evaluation is not specified. The AC also refers to a lengthy and comprehensive document on EFB human factors considerations by Chandra, Yeh, Riley, and Mangold (2003). However, the depth and breadth of Chandra et al. make it cumbersome to use in field evaluations.

In this report, we describe two promising paper-based tools that were developed to assist field evaluators: a short high-level tool that can be used in a brief evaluation, and a long detailed tool that could be used in a more comprehensive evaluation. The two tools were developed initially with the expectation that a single tool would eventually emerge. However, we found that both tools are valuable in their own way. We expect that the tools will benefit the FAA, as well as designers and operators, by providing a structure for EFB human factors evaluations. Both tools were designed for use by evaluators who are not human factors experts, although human factors experts may also find the tools helpful.

The evolution of the tools, our procedures for testing the tools, and our methods of processing the resulting data into feedback for the manufacturer are described. The tools were developed and refined over the course of several tests with prototype commercial EFB systems that were volunteered by vendors for these tests. The test methods and feedback reporting process were adapted from industry-standard techniques for usability assessment.

The content of these tools does not represent a coordinated FAA policy; the content reflects only the views of its authors. In addition, there is no requirement for either the FAA or industry to use these EFB usability-assessment tools. These products will stand on their own merit; If they are not useful, they will not be adopted. The next step is to introduce more potential users, especially those in the FAA, to the tools and methods to determine if these products are useful in practice.

1 Introduction

The Federal Aviation Administration (FAA) defines an Electronic Flight Bag (EFB) as any "electronic display system intended primarily for cockpit/flight-deck or cabin use" in the March 2003 Advisory Circular (AC) on EFBs, AC 120-76A [10]. This AC puts forth a streamlined field approval process for EFBs, contributing to the rapid growth of the EFB industry. In turn, the FAA is being contacted by more applicants seeking approval via the process specified in AC 120-76A.

In practice, the term "EFB" describes a wide variety of devices. The devices come in different hardware configurations, from laptops to fully installed systems, and they can support a variety of capabilities, such as electronic documents, electronic charts, and flight performance calculations. Other available functions include cabin video surveillance and surface moving map displays. (See [10] for a more complete definition and more examples of EFB capabilities.) In effect, EFBs are just general purpose computers. It is this flexibility and customizability that is so appealing to customers.

EFBs may look like familiar equipment in some ways, but they are new devices from a flight deck perspective because of their multiple flexible configurations and their open-ended ability to host multiple applications. In this sense, they are more like computers than like traditional avionics. Where traditional avionics combine specially designed hardware and software into a product with relatively fixed functionality, EFBs are multi-purpose computing platforms, generally with open architectures, onto which manufacturers (and possibly users) can install a wide variety of software applications. This makes a single, comprehensive definition of EFBs elusive, and it requires that an EFB evaluation tool must also be open-ended enough to address all of the uses to which an EFB may be put. In addition, many EFBs make extensive use of graphical user interfaces, which are also relatively new to the flight deck, and they can support multiple new functions, some of which may impact operating procedures. Ultimately, EFBs could play a central role in the future of flight deck information management [21] and they could be used in ways that we cannot even foresee today.

The FAA, system designers, and customers all recognize that EFBs could affect operating procedures and pilot performance [10, 21]. In fact, a significant portion of AC 120-76A (Section 10) addresses human factors considerations for EFBs [10]. As a result, human factors issues have received considerable attention from the EFB community. Some of the specific issues called out in AC 120-76A include user interface consistency, legibility, error potential, and workload.

AC 120-76A [10] refers to a comprehensive report on human factors considerations for the design and evaluation of EFBs [2]. That document was updated and superseded by Chandra, Yeh, Riley & Mangold [6], which has been distributed widely and is actively in use by EFB developers and customers. Chandra et al. contains considerations that apply to any EFB system, and considerations for four specific functions: electronic documents, electronic checklists, flight performance calculations, and electronic charts. The report helps designers make informed choices, and helps evaluators understand human factors considerations for EFBs that they may need to review. Established user interface design principles are described, recommendations are made, tradeoffs are described, and sources for more information are referenced.

Chandra et al. is valued as a comprehensive and readable reference. However, it does not satisfy one important FAA need: that of aiding an evaluator who is doing a field human factors evaluation of an EFB. The need for this type of evaluation is identified in AC 120-76A [10], but the procedure for doing this evaluation is not specified. In fact, translating the general human factors guidance in Chandra et al. [6] into a thorough yet practical EFB evaluation is a non-trivial task. Chandra et al. is not well suited for use during EFB evaluations because of its length and depth.

As a result, our research is directed at developing tools that could be used by FAA Aircraft Certification Service specialists to conduct structured and comprehensive, yet practical, EFB usability evaluations in the field.[1] Our goal is to make Aircraft Certification field human factors evaluations of EFBs more structured and standardized, while still being broad and comprehensive. Both EFB-specific issues and general user interface topics are covered.

We expect that the tools will benefit the FAA, and, secondarily, designers and operators, by providing a structure for EFB human factors evaluations. Because the products of this research are publicly available, system manufacturers could also use the tools and the recommended evaluation procedure during in-house design reviews to anticipate the general results of a future FAA evaluation.

1.1 Purpose

The purpose of this report is to document the progress to date on constructing usability-assessment tools for EFBs.[2] Below, we present our assumptions for designing the tools (Section 1.2) and provide an overview of the project history (Section 1.3). We then cover what the tools look like to date, how they were developed and tested, and how they could be used to assess and track EFB usability (Section 2). We also present our test procedure and methods of processing the resulting data into feedback for the manufacturer (Sections 3 and 4). In addition, Appendices A, B, and C of this report contain old versions of the tools from key points in the research, to illustrate how the tools have evolved based on usability tests and feedback from FAA Aircraft Certification.

In the next several months, we plan to introduce more potential users, especially the primary intended users in the FAA, to the tools to determine if they are useful in practice. Note that the tools described here may change based upon new feedback. The content of these tools does not represent a coordinated FAA policy; the content reflects only the views of its authors. In addition, note that there is no requirement for either the FAA or

[1] Note that some types of EFBs will be evaluated by FAA Flight Standards Service, not Aircraft Certification. While Flight Standards personnel are becoming aware of this research and its potential application to their needs, this research was based on Aircraft Certification processes and needs. In particular, the typical framework for an Aircraft Certification evaluation was assumed.

[2] Note that this report will be of use to researchers who are interested in our research steps and logic. While readers can also learn about the intended use of the tools, this report goes beyond a "user’s manual" for the tool. For users who only want to know how to use the tool, this report provides more detail than necessary.

industry to use these EFB usability-assessment tools. These products will stand on their own merit; If they are not useful, they will not be adopted.

1.2 Assumptions Regarding FAA Approval Process for Avionics

This research was based on assumptions derived from our understanding of the typical FAA Aircraft Certification process for conducting human factors evaluations for avionics approval. The goal of Aircraft Certification is to identify any major system weaknesses that need to be addressed prior to approval. The formality of these evaluations varies based primarily on the system complexity. In some cases, data from formal human factors testing (e.g., using simulators and pilots) must be presented to the FAA. Our research does not address this type of evaluation; the tools that have been developed do not provide the kind of quantitative data that can be obtained from in-depth studies. [3]

In a less formal procedure, evaluations are conducted in the form of multiple field visits to the manufacturer's site by small teams of FAA representatives (usually 2 to 4 people). These are the types of evaluation to which our research applies. As many as four or five reviews may be held over the course of system development. The evaluations are conducted in an office setting and they are brief—just 2 to 4 hours long—but thorough. The *equipment* is the focus of this evaluation, as opposed to its installation in the flight deck or training issues. In other words, evaluators focus on the hardware and software that they see in the office setting. Issues are tracked across evaluation sessions to ensure that they are addressed appropriately before the system is approved.

The evaluators in this setting may or may not be human factors experts, but they are familiar with a wide range of systems and operational environments. In addition, each evaluator brings his/her own unique perspective and experience to the evaluation (e.g., systems engineering, pilot/operator expertise, or human factors expertise). These individual differences will be reflected in the evaluation of the device. While a tool can facilitate the evaluation process by ensuring that many aspects of the interface are considered, using the tool does not eliminate individual differences between evaluators, and does not erase the application of individual judgment and ultimate authority.

During the on-site evaluation, the manufacturer gives a short presentation on the device to the FAA, and then each evaluator can ask questions and try using the device. Evaluations may take place at various stages of product maturity. In early evaluations, the device could be a relatively simple mockup, and over the course of the evaluation period, the device will mature into the real system. The idea is that everyone benefits from early interaction and exchange of information. Problems that are uncovered in the beginning stages cost less to resolve than problems discovered when the system is ready for release.

All parties take notes at the meeting, and the manufacturer can begin addressing any significant problems right away. However, one member of the FAA team is given the

[3] Further information on the FAA approval process for avionics can be found online at http://www.faa.gov/certification/aircraft/ under the column labeled "Design Approvals," items labeled "Original Design Approval Process" and "Avionics Approval Guide." (Website accessed September 22, 2004.). Also see FAA Order 8110.4B (April 24, 2000) [11] and AC 21-40 [12] on obtaining supplemental type certificates.

responsibility of coordinating the FAA notes into written feedback for the manufacturer. This lead FAA author drafts feedback from his/her notes and the notes of other evaluators. The draft feedback is coordinated among the evaluators who attended the meeting, and others with relevant expertise if necessary, until a final written review is accepted by all. Thus the manufacturer receives the written FAA feedback approximately two weeks after the onsite evaluation.

Note that the goals of an Aircraft Certification evaluation are different from the goals that a system designer or potential customer may have. The designer or customer may care about whether the system yields a measurable performance improvement relative to an alternative process or system (i.e., to choose between two methods or systems). This comparative type of evaluation can help EFB customers who are trying to determine how long it will take to get a return on investment, or designers who are trying to determine whether a particular design produces better or worse performance on particular functions. In contrast, FAA evaluators are concerned with questions such as, "Does the system meet its intended function without introducing any undue difficulty or additional risk?" In other words, does the system work, and does it work safely? Does it have any critical weaknesses? These are questions that are not relative; they apply to any one system that is being evaluated. The tools discussed in this report are geared to addressing these safety questions; they are not designed for making comparisons between systems.

1.3 Project History

We developed and refined the EFB usability assessment tools and evaluation methods over the course of several evaluations with realistic systems that were volunteered by vendors for the purpose of trying out the test procedure and draft assessment tools. Four different systems were formally evaluated between November 2002 and January 2004 (see Figure 1).[4] The formal evaluations included multiple sessions with outside evaluators and test facilitators. Both the test procedures and the EFB assessment tools were modified over the course of these evaluations.

In the initial study (November 2002), we evaluated the EFB using three standard techniques from the usability-analysis field [17, 19, 20]. The first was an "expert review," which consisted of a comprehensive examination of the device, its applications, features, tools, user interface, and physical form by a professional reviewer. The second technique was a "co-discovery" procedure, which is a loosely structured, but comprehensive, evaluation of the device's interface and applications by a team. Multiple team evaluations

---

[4] In addition, there were two other units that were tested informally during this timeframe. One informal test was conducted as part of an effort funded by the FAA SafeFlight21 Program Office. That evaluation focused on a surface moving map application running on an early prototype EFB. The format of the tool developed for use in that evaluation (but not its content) was developed based on earlier tools used for the EFB effort funded by the FAA Human Factors Research and Engineering Division. The second informal test was of a checklist application. That informal test was conducted by the project sponsors, as a way for them to understand the use of the tools better. Feedback from both of these informal tests was incorporated into the tool design.

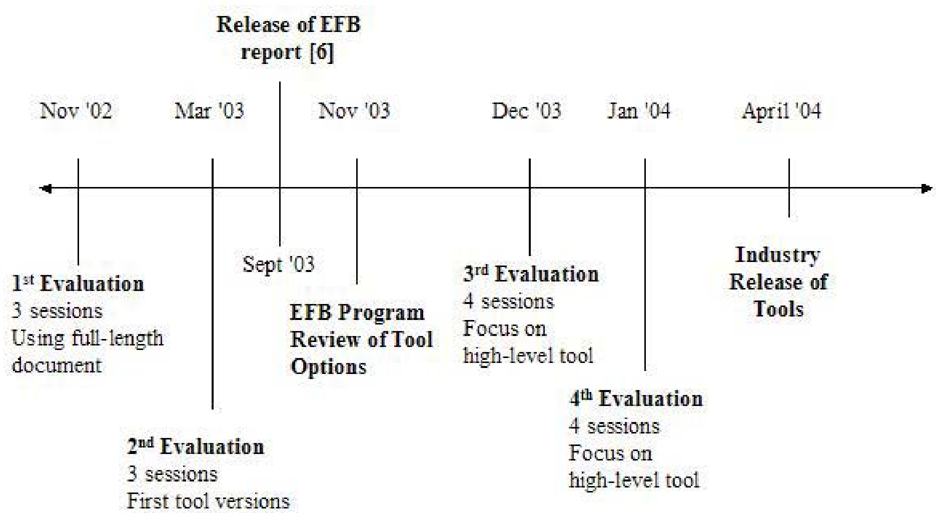


Figure 1. Project history.

were run, each with a different aviation human factors expert working together with the same facilitator.

The third and final technique used in the November 2002 study was a "heuristic review." Specifically, the device was evaluated by the project team against every consideration listed in a draft version of Chandra, et al. [6]. Each of the requirements and recommendations in that document represents a usability principle that should be satisfied (i.e., a *heuristic* in the usability literature)—hence the name of this technique, "heuristic review." In the usability field, heuristic reviews are usually conducted by usability experts. Our review was conducted by human factors experts familiar with the contents of Chandra et al. The task took one full day, much too long for a typical FAA evaluation. Clearly, a different type of assessment tool was needed; one that does not take such a long time to complete, and one that could be performed by field staff who are not human factors experts.

After the initial test in November 2002, use of the tools replaced use of the full Chandra et al. document in the evaluation, and the expert-review component was dropped. The team-evaluation format was retained and improved (see Section 3.1) for use in a natural setting, where the evaluation is not likely be moderated by a facilitator.

The first versions of two usability-assessment tools were constructed for the second evaluation (March 2003). The most recent versions of these tools are described in Section 2 below. Many versions of the tool were tested over the past several months. To document how the tool evolved, the appendices of this report contain three old versions of the tools, as well as the latest edition of the tools. The earliest tools, from March 2003, and our findings on their usability, are contained in Appendix A. In November 2003, several alternative designs for the tools were discussed at a program review meeting. These proposed versions are contained in Appendix B. Based on the discussion in November, an updated version of the tools was developed for two formal evaluations conducted in the December 2003/January 2004 timeframe. Those versions of the tools are provided in Appendix C. Finally, the latest versions of the tools, based on the outcomes of all the tests, and further discussion with our sponsors, are contained in Appendix D. The tools in Appendix D were distributed to industry for comment in April 2004, along with a presentation explaining their use. That industry presentation is reproduced in Appendix E.

2 EFB Usability-Assessment Tools

In this section, we first describe briefly how the tools were developed (Section 2.1) and their general use and content (Section 2.2). Next, each tool is described in detail (Sections 2.3 and 2.4); this is followed by a brief comparison of the tools in Section 2.5.

2.1 Tool Development

In developing the EFB usability-assessment tool, we started with three sources of information. One, of course, was Chandra et al. [6]. Second, we considered our expertise in human factors and the general principles of user interface design, which are well documented [17, 20]. Finally, we reviewed other checklist tools that were designed for assessing usability within an aviation context [1, 16]. An early decision was made to focus on a paper format so that the tool could be used easily in the field, even without a laptop computer.

The tool development process took us in two different directions. One tool was derived by condensing Chandra et al. [6]. This is called the "detailed tool" because it is long, consisting of over 200 specific items. The other tool was developed from a short list of high-level user interface design principles; it consists of about 20 items. Samples from these two tools and some alternative formats are provided in Appendices A, B, C, and D, as well as Chandra, 2003 [4].

Our original aim was to pit the detailed and high-level tools against each other, to determine which would form the best foundation for a single, all-purpose tool. Therefore, in the March 2003 evaluation, pairs of evaluators tried to use both tools in three sessions. The order of the tools was alternated, and the evaluators provided comments on both tools. Detailed results of these tests are provided in Appendix A. In brief, we found the key issues to be that the language of the short-tool was difficult to understand, and the length of the detailed tool was such that it could not be completed in a brief evaluation.

In preparation for the next evaluation, we first refined the language of the short tool, to address this known weakness. Second, we tried to use the short tool as the foundation for a single tool that could be used "in layers" (i.e., at either a high-level or detailed way, depending on the time available [4]). We tried to combine the two tools by mapping groups of topics from the detailed tool to the high-level topics in the short tool. The idea was that different users would be able use the tool in whatever way they wanted. For example, human factors experts might use just the high-level topics alone, while evaluators without human factors expertise could use the more structured and specific topics from the detailed tool. Alternate proposals for the combination tool are provided in Appendix B. In the end, the most promising options were the second alternative "high-level tool" shown in Section B.2, and the improved version of the detailed tool (Section B.4). These two tools were refined several times throughout the remainder of the tests, with the expectation that a single tool would eventually emerge. In the end, though, we found that both of the tools are valuable, but in different ways. Therefore, both tools are presented in this report.

2.2 General Tool Content and Use

As mentioned earlier, the tools will be used to evaluate EFB *equipment*, i.e. hardware and software that can be tested in an office setting. In other words, the tools only address the EFB human factors/pilot interface. Issues related to aircraft installation or crew training/procedures are not addressed. In the language of Chandra et al. [6], the tools address "equipment" topics. In the full EFB report [6], each statement is also categorized based on the type of information it contains. Some statements are considered to be "requirements," while others are recommendations, suggestions, or design tradeoffs.[5] In constructing the detailed tool, we limited the items in the tool to equipment requirements and recommendations from Chandra et al., as shown in Figure 2.

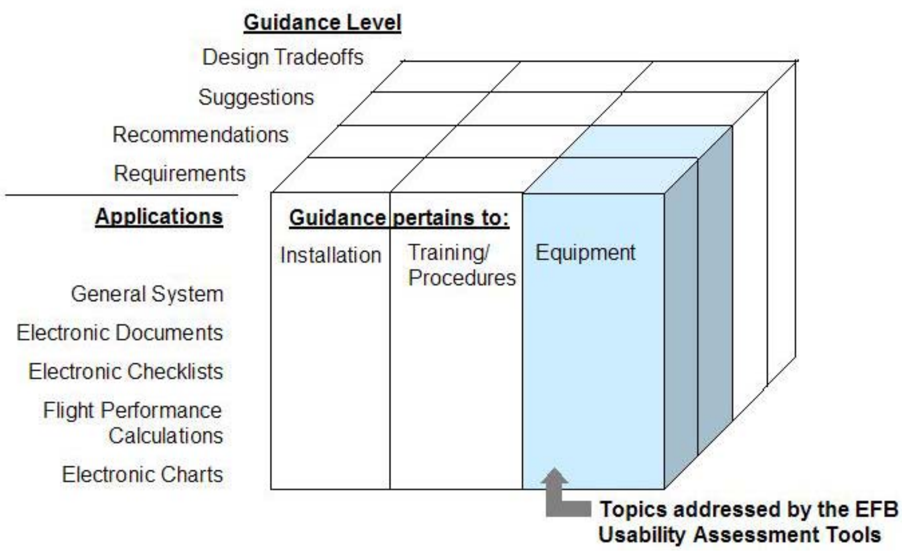


Figure 2. Relationship between content of Chandra et al. [6] and the EFB usability-assessment tools.

The tools are primarily intended for use by FAA field evaluators (e.g., engineers or test pilots), but they could also be used by others such as EFB system developers (e.g., software engineers), or even EFB operators (e.g., airline personnel). Note, however, that some experience and familiarity with the tools and the original document [6] is helpful for getting the full benefit. Human factors experts may also use the tool, of course. They may find the tools helpful for structuring their reviews, as long as they match reasonably well with their internal expectations of what to look for in a user interface; if the tools do not match their expectations, the expert may find the tools' structure helpful, but somewhat constraining.

Two important features of any assessment tool are the ability to record comments, and to assign severity ratings. After trying several alternatives, we opted for flexibility; the tools do not have space designated for recording evaluator comments or ratings. Instead the evaluator can choose how to record comments. Notes can be recorded directly into an

[5] Note that we use the term "requirement" here in a *non-regulatory* sense.

electronic file using a word processor, handwritten on a paper copy of the tool, or written into a separate notebook. Similarly, the tools do not have a designated rating scheme, giving the evaluator more flexibility in deciding how to designate severity ratings and accommodating individual rating preferences and styles. Severity ratings can also be assigned post-hoc based on evaluators' notes on the impact, frequency, and persistence of problems (see Nielsen's chapter in [19] and Section 4.2).

A numerical or categorical scale for problem severity could be used. An example numerical scale might be a 3-, 4-, or 5-point acceptability scale. An example categorical scale might be "showstopper" (i.e., important problem that must be resolved), "concern" (i.e., an issue that could turn into a showstopper), or "area for optimization" (i.e., an area that does not appear to pose a safety concern but where the interface could be improved). These scales are especially easy to use with the concrete, item-by-item format of the detailed tool, but they could also be assigned for problems uncovered with the high-level tool.

In general, we found that using a problem severity scale during the evaluation increased the amount of time required, especially if the ratings are assigned as a team (i.e., by a negotiated process) in real time. In addition, we expect that higher-resolution scales will provide designers with more detailed information about the quality of the system, but they will increase the time required for the evaluation significantly. For the purposes of a brief FAA evaluation, the rating scale could be relatively coarse, e.g., "acceptable" or "not acceptable," or a categorical scale such as the one described above (showstopper, etc.). A coarse acceptability scale would reduce the time needed for the evaluation, and it would allow FAA evaluators to focus their attention on the most important problem areas.

Note that issues can be tracked across multiple evaluations with either tool after problem areas are identified. In other words, the evaluators can use the first evaluation to identify significant problem areas, and then they can assess whether the problems were addressed appropriately in subsequent evaluations. Areas that were not problems in the first evaluation could be examined in terms of changes since the previous evaluation, to determine if any new problems have arisen.

2.3 High-level Tool

Appendix D contains the latest full version of the high-level EFB usability assessment tool. The short (1-page) version of this tool, for a generic EFB system, is provided in Figure 3. The 2.5-page version in Appendix D contains additional customized items for four applications (electronic documents, electronic checklists, electronic charts, and flight performance calculations), based on guidance in Chandra et al. [6].

The high-level tool is simply a relatively short list of usability topics to consider for the evaluation.[6] Notice that there is some overlap between the items; the items are not

---

[6] For readers who are familiar with the usability literature, an interesting exercise is to compare the Volpe high-level tool with Nielsen's top ten list of usability heuristics [18, p. 30]. The Volpe tool actually provides a more specific and detailed taxonomy than Nielsen, although there is considerable correlation between the two.

EFB Usability Assessment Tool

HARDWARE CONSIDERATIONS

- Physical Ease of Use
  — Input devices and display, accessibility of controls
- Labels and Controls
- Lighting Issues (day vs. night use)
  — Brightness adjustment, illumination of labels
- Amount of feedback, potential for errors

SOFTWARE CONSIDERATIONS

Symbols and Graphical Icons
- Clarity of intended meaning, confusability
- Legibility and distinctiveness

Formatting/Layout
- Fonts (size, style, case, spacing)
- Arrangement of information on the display
  — Consistency with user expectations and internal logic

Interaction (Accessing functions and options)
- Home pages and ease of movement between pages
- Number of inputs to complete a task
- Ease of accessing functions and options
- Feedback (system state, alerts, modes, etc.)
- Responsiveness
- Intuitive logic

Error handling and prevention
- Susceptibility to error (mode errors, selection errors, data entry errors, reading errors, etc.)
- Correcting errors (e.g., cancel, clear, undo)
- Error messages

Multiple Applications
- Consistency and compatibility across applications
- Identifying current position within system
- Ease of switching between applications

Automation (if any)
- Is there enough? Too much?
- Is it disruptive/supportive? Predictable? User control over automation? (e.g., manual override)

General
- Consistency of controls/elements; are they distinctive where appropriate?
- Visual, audio, and tactile characteristics
- Use of color (especially red and amber) and color-coding
- Amount of feedback (system state, alerts, modes, etc.)
- Clarity and consistency of language, terms, and abbreviations
- End-user customization (if any)

WORKLOAD
- Problem areas

OTHER

Figure 3. High-level tool for a generic EFB system.

mutually exclusive. For example, "Feedback" is listed under *Hardware Consideration*, *Interaction*, and *General topics*. Similarly, error handling and prevention is listed as "potential for errors," under the Hardware topic, and as "confusability" (under the Symbols and Graphical Icons topic), and it is called out as a more general topic overall.

The overlap between items in the high-level tool occurs for various reasons, but in the end, the multiple occurrences of similar topics feels natural and unobtrusive to the tool user. In the case of "feedback" issues, the overlap occurs because we, the tool designers, could not know in advance where feedback would be most relevant for the given EFB. Feedback could be a prominent issue for the hardware in one EFB system, or for the software in another system. In the case of error handling, the topic was called out more than once because we were specifically interested in capturing detailed comments about error potential, which may be especially important to regulators. Another explanation for the overlap is simply that the topics are not (and cannot be) mutually exclusive. Each topic is inherently multi-faceted. Regardless of the justification, the overlap between items in the high-level tool helps to ensure that all significant issues are found, regardless of the specific characteristics of the device. In some cases, a single issue will be identified from multiple avenues, but that is better than missing an issue entirely.[7]

Evaluators go through the high-level tool commenting (aloud) about each item, with a note-taker recording the comments. The comments, which could be either positive or negative, can actually be more valuable to a designer than problem severity ratings because they give the designer insight into the cause of any difficulties. In some cases, topics in the tool will not be relevant, but it is important to consider every item to ensure a thorough evaluation. As the evaluators comment on each item, they provide supporting examples, and, if they choose, preliminary assessments of problem severity.

Through tests of the tool against real systems, we fine-tuned its content, item order, and language. As mentioned earlier, the content of the tool was generated from a generic high-level list of user interface dimensions (see Chandra [4] for the full initial list). This list was fleshed out by adding items that represented themes (i.e., groups of items) in the detailed tool generated from the full-length EFB document [6]. The net result is a good blend of high-level and somewhat more specific topics.

Our philosophy for item order was to go from concrete to abstract or local to global. Because evaluators may still be familiarizing themselves with the system early in the evaluation, we expect that they will find it easier to start by commenting on concrete aspects of the design (e.g., icons and formatting). As they build up experience with the system, they will be able to comment on more abstract, potentially global, aspects of the design (e.g., error potential, consistency across applications, or workload).

Another significant issue mentioned earlier is that evaluators had difficulty with the language in earlier versions of the tool. Terms that were familiar to some human factors

---

[7] From a theoretical perspective, it may be worth noting that the only way of avoiding either missing issues or identifying issues redundantly is when there is a perfect mapping between the structure of the evaluation tool and the structure of the user interface. Given the range of possible EFB user interfaces, such a perfect mapping is unlikely, so building some redundancy into the tool is the best way to minimize the possibility of missing issues entirely.

experts were not always intuitive for nonexperts; some terms were not even clear among human factors experts. Our most recent tests show that the current language is understandable, or at least not distracting, to evaluators. During the evaluations, if the evaluators did not understand a term, they were asked to guess at its meaning and then use their own definition to complete the tool. This allowed us to determine how the evaluators were interpreting unclear terms, provided input into how we could clarify the term, and also supported the overall objective of using the tool to motivate thinking and exploration.

In summary, the high-level tool is valuable because it covers the breadth of a system regardless of the system's complexity. Evaluators can use the tool effectively, especially if the main goal is to catch "showstopper" problem areas. Our tests have shown that a key advantage of this tool is that evaluators can easily cover all the items in under 1 hour. For such a brief evaluation, the tool gives surprisingly good results in terms of catching problem areas, even when the evaluators are not human factors experts, making it suitable for use in brief FAA evaluations. For evaluators who are human factors experts, the tool's relative open-endedness gives them the opportunity to apply their own expertise and address the issues they consider most important. One limitation of this tool is that its results are more subjective than objective, because the topics are merely reminders of what to examine without specifying a clear criterion to be met. Evaluators apply their own (internal) criteria to the topic at hand. A second limitation of the high-level tool is that the results do not necessarily help to identify ways of fixing the problems that are uncovered. To accomplish that goal, more structured usability tests are necessary.

2.4 Detailed tool

Chandra et al. [6] provides the foundation for the detailed tool. Specifically, the detailed tool is a reformatted version of that report's Appendix B, which is an 11-page summary of roughly 100 pages of equipment requirements and recommendations.

The original format of Appendix B from Chandra et al. [6], and sample items, are shown in Figure 4. Each item is a paraphrased version of guidance from the main document. When Chandra et al. [6] is viewed electronically, links to the full topic description are active in Appendix B. The items in this tool are most useful to readers who are already familiar with the structure and general content of the full document. Non-regulatory "requirements" are indicated with a ❖ symbol, and recommendations are indicated with a ❑ symbol. The difference between "requirements" and "recommendations" is based only on the authors' assessment of the potential importance of each item to safety; the document is not regulatory.

The format of the detailed usability assessment tool is shown in Figure 5. This format is tighter than that used for Appendix B of Chandra et al., but it is still significantly longer than the high-level tool. All of the content of that appendix is retained in the detailed tool. The items have been grouped and re-ordered, however, for closer correspondence with the high-level tool (see D.1 and D.2). Note that, as the order of the items in the high-level tool evolved, item order in the detailed tool was adjusted accordingly. For

Section	**Topic**	**Guidance**
2.1.1	Workload	❑ Flight crew workload and head-down time should be minimized (AC 120-76A, Section 10.c)
2.1.5	Legibility—Lighting Issues	❖ Automatic brightness adjustment should operate independently for each EFB
		❑ Screen brightness should be adjustable in fine increments or continuously ❑ Buttons and labels should be adequately illuminated for night use

Diamond bullets represent non-regulatory "requirements."

Square bullets represent "recommendations."

Figure 4: Format of EFB summarized equipment requirements and recommendations from Chandra, Yeh, Riley & Mangold, 2003 [6]. The structure is intended to support quick review with pointers to more detail when needed.

2.4.3 General Use of Colors

- ❖ Red and amber should be reserved for highlighting *warning* and *caution* level conditions respectively (AC 120-76A, 10.d(1))
- ❖ Color should not be sole means of coding important differences in information; color should be used redundantly
- ❖ Color-coding scheme should be interpretable easily and accurately
- ❑ Each color should be associated with only one meaning
- ❑ No more than six colors with assigned meanings should be used in a color-coding scheme
- ❑ EFB colors should not conflict with flight deck conventions
- ❑ For Part 121 and 135, default colors that represent different types of data should be customizable only by an appropriately authorized administrator
- ❑ If colors are customizable, there should be an easy way to return to default settings

Figure 5. Format of an EFB usability assessment tool based on Appendix B (Summary of Equipment Requirements and Recommendations) in Chandra, Yeh, Riley & Mangold, 2003 [6].

example, all the "Hardware" considerations appear at the beginning of the detailed tool; the high-level tool begins with hardware topics as well. The next item on the final versions of both tools is Software Considerations for Symbols and Graphical Icons, and so on. For a generic EFB system, the detailed tool is five pages long, and including the topics for specific applications adds another five pages.

The latest version of the detailed tool is provided in Appendix D. There are over 200 items in this tool, and although the items are specific (i.e., in the form of do's and don'ts), some of them still require the evaluator's judgment. If a system complies with the statement, the result is a "more usable" system relative to those that do not comply. The formatting for the bullets in the detailed tool is identical to that used in Appendix B of Chandra et al. [6].

To use the detailed tool, the evaluator would go through one item at a time, and decide whether the system met the item, and, if a rating is desired, to what extent. As discussed earlier (Section 2.2), the rating scale can be tailored to the time and goals of the assessment. Manufacturers could use a higher-resolution acceptability scale to learn where to optimize the design, or a coarse scale may be used for a time-constrained FAA review.

In summary, the detailed tool was developed from Chandra et al. [6]. It contains a long list of usability principles that are highly specific to EFBs. Users who are familiar with the full document will be able to use the detailed tool more quickly and efficiently. However, we found that the detailed tool takes much longer to complete than the high-level tool, so it is less suitable for a brief evaluation. None of our evaluation teams were able to complete the detailed tool, even when given well over one hour.

2.5 Comparison of High-Level and Detailed Tools

Both the high-level and detailed tools help evaluators articulate positive and negative aspects of an EFB user interface. They promote a comprehensive review of the interface, aiding evaluators in catching many (but not all) of the same issues. The detailed tool tends to catch more EFB-specific issues; the high-level tool appears to be better at capturing problems that are more pervasive throughout the design (as opposed to problems that are self-contained, being encountered only under specific circumstances).

Completing the entire detailed tool takes significantly more time than completing the high-level tool. This is primarily because the detailed tool has many more individual items than the high-level tool. Although the depth and time of the evaluation can be adjusted based on time constraints and the objective nature of some items from the detailed tool make them quick to complete, covering the full list of items generally takes more than a few hours. In addition, some items require complex assessments, either objective or subjective. Finally, achieving group consensus on each and every item's problem-severity rating is especially time consuming. Higher-resolution severity-rating scales, which could be used to identify ways of optimizing the system (e.g., reduce workload or training time), also add to the time required.

Both tools present items in similar orders. However, there is not a simple mapping between topics in the high-level tool and items in the detailed tool. Most items in the high-level tool map to multiple items in the detailed tool. Some items in the detailed tool are also cross-referenced to more than one topic area in the high-level tool. (See the detailed tool in Section D.2 for examples.)

Another difference between the tools is that the high-level tool provides only topics, whereas many items in the detailed tool bring up a general topic area and specify a criterion that should be met. As a result, evaluator ratings for these types of items from the detailed tool are generally more objective. With the high-level tool, the topic area stimulates the evaluator to identify strengths and weaknesses in the design, but the evaluator needs to apply their own (internal) criterion to perform evaluations for each topic. Therefore, EFB evaluations using the high-level tool are more subject to individual differences.

3 Test Procedure

In this section, we focus on the test procedure developed during the research program to evaluate the EFB usability-assessment tools. In particular, we focus on the mature protocol used in the December 2003 and January 2004 tests of the tools. First, a typical session is described. Next, in Section 3.2, the test methods are discussed in more general terms. In Section 3.3 we provide generic advice for those who wish to conduct their own, similar EFB tests. Finally, in Section 3.4, we provide advice for those who will perform the functions of an observer/note-taker in the tests.

3.1 Typical Session

Each test session lasted 3 to 4 hours total, and we ran three or four sessions per device, each with two or three evaluators. That is, 6 to 8 evaluators generally reviewed a given product overall. This is a typical number for usability tests (see Nielsen's chapter in [19] and 20]. As noted in Nielsen [19], the vast majority of usability problems are uncovered with 6 to 8 evaluators. The costs of running additional evaluators are usually not justifiable.

The evaluators in our tests were typically researchers with an aviation and/or human factors background. Some were licensed pilots and/or experienced system designers, but they were not FAA personnel or air transport pilots (the intended end users of the systems). To give them a sense of the FAA perspective, we sent evaluators materials in advance, including copies of AC 120-76A [10], Appendix B from Chandra et al. [6], AC 25-11 [9], and a draft copy of the assessment tools. FAA staff would be familiar with the two ACs and may have seen the assessment tools before the evaluation as well. In addition, it is helpful for tool users to (a) have enough general knowledge of user-interface components to be able to articulate their impressions of a device, and (b) expect that they will encounter problems and realize that these problems are not their "fault."

The tests were conducted with a facilitator in an office setting. The test consisted of the following stages, which are described in more detail below:

1) Introduction (15 min)
2) Task-Based Exploration (1 to 1.5 hour)
3) Tool-Based Review
    a. High-Level Tool (up to 1 hour)
    b. Detailed Tool (up to 1 hour)
4) Feedback on tools and wrap-up (15 min)

3.1.1 Introduction to the Test

During the introduction, the facilitator explained that the purpose of the test was to evaluate the EFB usability assessment tools and gave a brief introduction to the EFB system and applications that would be reviewed. The brief introduction provided context on the application(s), the system, and their intended use.

The introduction was not intended to mimic formal system training. As a result, the evaluators were possibly less prepared than EFB end users might be with the system. However, we felt that this was an appropriate worst-case scenario to consider for two reasons:

1. EFB end-users may see some system features infrequently, or under stressful conditions, where intuitiveness could be an important factor in actual performance.

2. The typical FAA evaluator may not receive full training with the system prior to reviewing it.

In addition, manufacturers strive to build systems that require minimum training. Our protocol put this theory to the test.

3.1.2 Task-Based Exploration

The task-based exploration phase was effectively a self-paced familiarization period with the system. Participants stepped through a set of tasks that included common functions that typical users would want to accomplish. Some were generic tasks, but others were specific to the system if the manufacturer, or project team, decided there were areas of particular interest (e.g., areas identified as problematic from earlier evaluations). In these cases the task descriptions steered users towards exercising particular functions. In general, the tasks were conducted in a natural flow, proceeding directly from one task to the next rather than "resetting" users to a designated starting point.[8]

Every task was designed to have a beginning state and a desired goal. They were open-ended enough that users could digress for a while, but specific enough that participants knew when they had successfully satisfied the goal. An example of an appropriate task for an electronic chart application might be, "Find and display the airport diagram for Boston Logan (KBOS)." Note that tasks should not be over-specified or under-specified. A task is over-specified if it provides specific steps for accomplishing the goal, such as those in a user manual (e.g., "Select the airport from the Airport Menu and choose Boston Logan (KBOS)"). A task is under-specified if it does not provide users with a clear goal (e.g., "Use the chart application as you would during an actual flight"). More generic task examples are provided in Appendix F.

It is important to let evaluators perform the tasks without assistance, even if they stray from the manufacturer's intended path toward a goal during this phase. Unintended digressions can help evaluators develop an internal model of the interface structure, which can help identify where the user interface structure is non-intuitive or inefficient. Recovery from digressions also helps users build confidence in their understanding of the system.

Participants were asked to think aloud as they performed the tasks, stating their expectations and rationale for the steps they tried. These spoken comments were transcribed by a note-taker/observer in real-time. (In some tests, the facilitator and note-taker were the same person; in other tests, they were separate individuals.) The notes captured the entire discussion, including any dead-ends that the evaluators encountered. The note-taker/observer could also ask for clarifications and/or examples as needed. In general, however, the evaluators were not interrupted.

3.1.3 Tool-Based Review

The high-level tool and the detailed tool were described earlier (see Section 2). A sample version of the high-level tool is shown in Figure 3, and a sample from the detailed tool is shown in Figure 5. The high-level tool was typically completed in well under one hour. The detailed tool, however, took longer, especially if the system consisted of multiple applications. In earlier evaluations, we varied the order of the tools [4], but in the latter two evaluations, we always presented the high-level tool first because it was the main tool we were assessing. Evaluators were given only one hour to work on the detailed tool,

[8] Occasionally, evaluators become confused and the facilitator asked them to go back to a familiar page and begin the task from that point again.

even if they had not finished. Again, evaluators were asked to talk aloud as they worked through the tool, and a note-taker transcribed their comments. Clarifications and/or examples were solicited as necessary.

3.1.4 Feedback on Tools and Wrap-up

The last step in the test was to obtain feedback from the participants on our tools and methods. We used a written questionnaire to structure the comments. Responses to the questionnaire helped us to identify changes to be made to the tool. Some changes, such as clarification of terms, were implemented prior to the next test session; aggregate results from the questionnaires are not meaningful so they are not presented here.

3.2 Discussion of Test Format

In the usability literature, "co-discovery" describes the technique where two or more users work together to evaluate the device [8]; we use the term "team evaluation" in this report for the same concept. In our tests, the evaluators spoke aloud the whole time; this is called the "thinking aloud" protocol in the usability literature [15]. The evaluators were asked to talk continuously about what they saw, what their expectations were, whether they were confused, their hypotheses about why they might be confused, etc. Both of these techniques add to the quality of an evaluation and fit well with the current FAA process.

Co-discovery is useful because it is a good way to collect a lot of subjective data in a short time [22]. The team-evaluation format creates more opportunity for discussion and exploration and, as a result, evaluators working in teams often discover more about a system in one session than evaluators working alone do. When team evaluations were compared directly with single-user evaluations, researchers found that the team evaluations were "more efficient in extracting high value information without noticeable differences in performance or subjective impression of the software" [15].

There are potential drawbacks to the team format (see [22]), two of which may be particularly relevant to an FAA human factors evaluation. First, the evaluators' interaction may be unrealistic if the task is not typically performed by a team. So if the EFB is designed to be used by only one crew member, having two use the system may not be optimal. Second, the team format could potentially introduce bias into the evaluation, in that the team may express negotiated opinions of the system, hiding underlying confusion or disagreements. This type of bias could be more or less likely depending on the pre-existing relationship between the evaluators. Evaluators who already know each other and are assertive, as most FAA evaluators would probably be, are less likely to defer to one another, mitigating the risk of hidden confusion. In the end, of course, the team evaluation is the process used by the FAA, so our tests replicate a more realistic situation than highly-controlled usability testing. Whatever bias this introduced in the tests is likely to exist in the real process as well.

Thinking aloud during the evaluations serves several important purposes. First, it allows an observer to take more detailed notes about what the evaluators did and thought during the evaluation. Second, thinking aloud gets evaluators to articulate their impressions of the system (good or bad). If they get confused, the evaluators talk through their attempts at understanding what is happening. If the evaluators identify a problem, then they describe what the problem is, under what circumstances they find that it occurs, and perhaps even hypothesize about the source of the problem (e.g., perhaps a button was labeled unclearly, or a perhaps they find that the task requires too many steps). All of these verbalizations provide system manufacturers with more information than ratings of problem severity alone do.

The observers in our tests played a key role in capturing the discussion and comments of the evaluators; they were often more aware of the difficulties than the users themselves. In addition to taking notes, the observer managed time during the evaluation, making sure that the evaluators completed at least the full high-level tool. The observer also asked for clarifications/examples from the evaluators when they were identifying specific characteristics of the device. If necessary, the observer also reminded the evaluators to think aloud; some evaluators find it easier to talk continuously than others.

Having a dedicated person to observe and transcribe the evaluation is not standard in an FAA evaluation, in which evaluators generally take their own notes. Our early tests revealed, however, that evaluators were distracted by having to take their own notes, and their note-taking disrupted the flow of the open-ended discussions in progress. Also, notes taken by the evaluators tended to be incomplete and not especially useful to anyone but their author (if that). In contrast, notes from a dedicated note-taker were relatively complete transcripts of the sessions, recorded directly into an electronic document in our tests (see a sample set of notes in Appendix G). Using the observer's detailed notes, it was much easier to review the data and to construct written feedback for the manufacturer about the device. In the FAA Aircraft Certification process, the person who is responsible for creating the written feedback to the manufacturer would probably be the best candidate for taking detailed notes of the session.

In addition to preparing a written report providing feedback to the manufacturer, we discussed the feedback with the manufacturer on the phone in a follow-up call. In discussing the problems with the manufacturers, we often found that the problem reports were not a surprise. We found this to be evidence of the value of the tool because we spent far less time with the devices than the manufacturer and still uncovered their most difficult usability issues. So, although the evaluation was brief, it did uncover key usability problems.

The tools described in Section 2 could be used with a procedure quite different from what was used in our research tests. For example, a single evaluator could sit down with the device and the tools, and simply mark whether the device was "acceptable" or not for each topic in the high-level tool, or for each item in the detailed tool. While this use of the tools is possible, it is far from optimal. Although we did not test this type of tool usage, our hypothesis is that the quality of such an evaluation will be far below that of an evaluation of the type described in Section 3.1, because of the lack of a detailed record of the evaluation, and because of the lack of input from multiple evaluators.

3.3 Steps for Running an Evaluation

The steps listed below summarize what the research team did for an EFB test. The list can be used as a brief systematic guide to doing EFB human factors assessments in a manufacturer, field, or research setting. FAA evaluations may be less controlled than the tests we describe, but will follow a similar process. More detail on these steps, from an industry point of view, is provided in Appendix E. For the research tests, the whole process took approximately three weeks, from the time we received the test equipment, to the time that we had drafted the manufacturer feedback. The steps for running an evaluation are:

1) Review the user interface. The goal is to understand what functions can be performed, and how they are accessed. For example, a flow diagram could be prepared to ensure that the whole product (or application) is reviewed. Any specific areas of interest should be identified.

2) Select the benchmark tasks. What are some of the common and less common functions that a user will try to do with the device? Evaluators will attempt these tasks in order to learn the system. Remember not to over-specify or under-specify the tasks, as discussed in Section 3.1.2. Some tasks could be developed to steer users toward particular areas of the user interface.

3) Choose the evaluators. Ideally, evaluators should not be too familiar with the device, or they may not be able to see it from the point of view of a new user. While the evaluations can be run with single evaluators, we prefer to have two or even three evaluators at one time to take advantage of the co-discovery procedure and to simulate a typical FAA evaluation. Discussions among the evaluation team take time but add depth to evaluation.

4) Decide on the time/depth desired. As noted in Section 2.2, the tool can be used in different ways depending upon the time available. Scope the evaluation by choosing what applications/functions will be reviewed, and how in-depth the evaluator will review them (e.g., by selecting an appropriate rating scale if desired).

5) Provide a copy of the tool to evaluators in advance and set expectations for the test. Evaluators need to be familiar with the tool structure and items before the evaluation. They should also know in advance how much time the evaluation will take and/or how much breadth and depth the evaluation will cover. Being familiar with the tools and having reasonable expectations about the test both help to limit digressions. In addition, evaluators will find it easier to shift mindsets between items and they will spend less time reading during the evaluation if they understand what is expected of them in advance.

6) Select a quiet location for the evaluation. Evaluators should not be interrupted during the test.

7) Select a note-taker/observer. Qualifications for this job include an ability to document what the evaluators are saying without inserting any analysis or interpretation during the actual evaluation. In other words, the observer should be able to record the session without bias. The observer should be somewhat familiar with the system, and have a plan for taking notes in real time.[9] If multiple sessions are run, try to have the same observer

[9] In more formal usability tests, the sessions are videotaped for review later. This does not happen in FAA evaluations. Analyzing data from videotapes is also time consuming; we found that a good note-taker can capture the relevant information quite efficiently for our evaluations.

across the different evaluation sessions to make the synthesis task easier. (See Section 3.4 for more advice for observers.)

8) Run the evaluation(s). Time should be managed carefully to ensure that none of the evaluation steps are shortchanged. Evaluators should get some brief initial training to familiarize them with the available functions and purpose of the EFB. After the initial training, the evaluators try to accomplish the tasks using the device, and then complete the assessment tools, referring back to the device as desired.

9) Synthesize the data and write up feedback. The synthesis and write-up of feedback for the manufacturer is critical for identifying patterns and systemic interface issues that may not be obvious during the actual evaluation(s). This step is discussed further in Section 4 below. In this step, a lead author (often the observer in the tests) reviews the data and tries to identify and document problem areas in the interface. A report on the tests is written to the manufacturer describing the tests and results in a constructive manner, using language that makes sense to the system designers. A follow-up phone conversation about the written feedback is helpful for ensuring that the problem areas are conveyed clearly to the designers. During the conversation, designers can ask a lead evaluator for clarification of, or more information about, the findings and evaluation process.

3.4 Advice for the Observer

The task of the observer is to record the session without bias. This is not easy at first, but with some practice (one or two sessions), the job becomes quite manageable. In addition, the note-taker/observer is the logical person to lead the synthesis effort, especially if he/she observed more than one test because, by watching others use the system, the observer may have developed a deeper understanding of the device than evaluators who participate in only one test.

For novice observers, preparation is especially helpful. Specifically:

1) Be familiar with the system. The observer should be somewhat familiar with the system before the evaluation, so that he/she can observe difficulties without becoming involved in the problem solving. In particular, the observer should explore the system before the test to understand the underlying structure of the user interface.

2) Be familiar with the tasks. The observer should be familiar with the tasks that evaluators will perform. Ideally, the observer will be the one who developed the tasks if they were customized for the test.

3) Be familiar with the tools. The observer should be familiar with the assessment tools and their use, so that he/she can manage time effectively, and ensure that all the necessary tests components are completed.

4) Have a system for taking notes. The observer should have a plan and system for taking notes. For example, the observer should be able to recognize and name pages on the system using a personal naming convention, if necessary. If participants go to an unfamiliar page, the observer should plan to draw it or note key features so that the page can be found later. Any personal observations should be distinguished from evaluator's comments within the notes. The observer should also have a plan as to how the notes will be taken (e.g., electronically vs. handwritten), and have a template in mind for what the notes will look like (e.g., straight text in a word processor vs. table format, etc.). Personalized templates could be developed, but what works for one observer may not work for another. A sample of our observer notes is provided is Section G.1.

5) Transcribe the session. During the evaluation, the observer should transcribe the session, recording as much of the of the evaluators' dialogue as possible. (Again, see G.1 for a sample set of notes.) He/she should enter what the participants say, without either interpreting or drawing conclusions, or influencing the discussion. All discussions should be recorded because, while some may be irrelevant, others may lead to key findings. By recording all discussions, the observer will be sure to capture any difficulties or patterns as they happen.

6) *Interrupt only when necessary.* Interruptions disrupt the flow of the evaluation and should be minimized. Appropriate interruptions would be to (a) ask evaluators to provide examples to support their impressions of the device; (b) encourage the evaluators to think aloud, if they are too quiet; (c) help evaluators go back to an earlier point in the discussion, to re-attempt a task in which they became confused or lost. In particular, the observer should watch the evaluators struggle to resolve problems on their own and record what solutions were attempted without providing a solution unless the evaluators are unable to find it after several attempts.

4 Data Analysis, Synthesis, and Reporting

In this section, we discuss the importance of the data analysis of the usability tests (Section 4.1) and the process that we used to create meaningful and constructive feedback for the manufacturer (Section 4.2). A more comprehensive and general discussion of writing feedback for developers is given by Jeffries in [19]. Note that, given the nature of the FAA evaluation, our feedback focuses on identifying the weaknesses of the interface, not on its strengths. Also, the focus is on describing and justifying the problem, not on suggesting and justifying solutions.

4.1 Purpose

Synthesizing data from the evaluation into feedback for system developers is nontrivial to accomplish, but of great value. It is through this process that evaluators reflect upon the interface as a whole (e.g., in terms of consistency, or other high-level issues) to produce more accurate diagnoses of underlying design issues. Identifying relationships among issues makes it easier to track problems and promotes development of more global solutions because synthesis can reveal root causes for difficulties that may seem unrelated initially. For example, changing the task flow of a user interface may be a better global solution than changing button labels to improve navigation through user interface. Another reason that synthesis is important is that initial suggestions for system designers may need to be revised as problems are understood better. Mature feedback, with more accurate conclusions and recommendations, will be more meaningful, accurate, and useful because the focus is on the root problem, not multiple surface symptoms of that root problem.

The goal for the synthesis is to provide feedback that is actionable, specific, and based on objective data as much as possible. By "actionable," we mean that improvements should be feasible, and the manufacturer should be able to prioritize, address, and track progress on each issue. Specific and objective examples of where users became confused or inefficient while attempting to use the system are critical for both demonstrating to system designers that the issue is real, not just a matter of opinion, and for illustrating more general problems that may be found elsewhere in the interface. Comments from the evaluators are invaluable for this purpose.

4.2 Process

We collected many pages of notes from each evaluation session (see sample in Appendix G). The notes were in two separate sections, one from the task-based exploration phase, and the other from the tool-based review. Either of these sets of notes could be analyzed independently, but the tool-based review produced notes that were easier to use as a starting point.

We first collated the notes from the tool-based review across the different evaluation sessions. This produced a file that used the section headings from the tool, with comments on each aspect from every evaluation team below. In practice, evaluators did not proceed through the tool items in order; they often started from one topic and then mentioned other associated topics. However, this was not a problem for the data analysis because the overall quantity of data collected from the tool was relatively small (but dense), and related issues could easily be identified.

After collating the notes across sessions, the next step was to clean up the notes by deleting incomplete thoughts, repeated comments, and any other uninformative material.

Once the notes were clear, a list of problems was created. This list captured specific difficulties encountered when using the device. Ideally, the list contains only *unique* problems (i.e., each problem is listed only once); generating this list takes some effort, and possibly some skill as well, because the analyst must decide which problem reports are unique, and which are actually duplicates. Note that although each problem is listed only once, the number of times the problem is observed or reported should be tracked.

The next step was to identify relationships between the problems and to gather related problems under a single topic heading. Problems were eventually classified as one of the following: (a) *core problems*, i.e., the topic headings; (b) *examples* of core problems, i.e., specific instantiations of the core problem; (c) *observations*, which were items we felt that developers would want to know about, but that did not have a clear human factors impact. Observations often had to do with preferences that might be specific to the intended end user.

During the preparation of the report, we iterated on the first step of classifying problems, revising and refining the core problems as we searched for the right level of abstraction. For example, a set of error-related issues may appear at first to be unrelated, but may all arise from a single root cause. When the classification step was completed, the number of core problems was minimized; a list of dozens of problems could usually be rewritten as a list of 12 to 15 core problems, with several examples and observations under each.

The last step of data analysis was to determine the severity of each of the core problems by noting its impact, frequency of occurrence, and persistence. [10] Problem severity would definitely be high if there was a violation of FAA regulations and/or policy, but it could also be high if there was a significant anticipated performance impact. The problem's frequency of occurrence should be considered as well. Is the problem systemic (found throughout the system) or local (self-contained)? Is the problem common across evaluators (i.e., many users have the same trouble, not just a few)? If so, that would indicate that a better system design is needed. The persistence of the problem within a participant is another reason to consider elevating the problem severity rating. Some problems may exist only for new users, but the intuitiveness of the interface impacts training time, so it is important to note whether participants learn to ignore the problem, or whether they have trouble throughout the test. If the user has trouble repeatedly over the course of the evaluation, the user interface should be improved.

Over the course of these evaluations, we developed a standard format for written feedback to the manufacturer (see Appendix G for a sample). It included an overview of the evaluation protocol and purpose, and a table of contents, which provided an overview of the core problems. The individual core problems were assigned high, medium, or low priorities (i.e., one of three problem severity ratings). High priority core problems were those that either violated known FAA regulations and/or policy, or were global and, in our opinion, had a potentially significant performance impact. Low priority core problems were areas we felt could use improvement, but did not appear to have a significant performance impact. The bulk of issues were neither high nor low priority, and so were given a default label of "medium" priority.

In the feedback, the titles for the core problems were carefully crafted to be indicative of the problem, not the solution. In addition, the titles were specific to that EFB. In other words, it was not always appropriate to use generic headings from the tool; doing so can produce feedback that is not always specific enough to act upon. We recommend that feedback be given in terms of functional user interface components. Each topic area began with a general statement of the difficulties encountered, along with information about the frequency of occurrence. *Specific* examples were provided, and, where possible, quantitative results (e.g., 4 of the 6 participants had this problem). Examples of where users became confused or inefficient are critical for demonstrating to system designers that the issue is real, not just a matter of opinion, and for illustrating more general problems that may be found elsewhere. Where appropriate, we made suggestions for design changes that could address the issue. Often, these were suggestions made by the evaluators during the session, but sometimes they were suggestions based on the synthesized findings across evaluation sessions. As mentioned earlier, however, the focus of the report was on describing and justifying the problem, not on suggesting and

---

[10] One manufacturer used a slightly different technique to identify problem severity. Rather than starting from the list of unique problems, she started from the unedited list of problems and simply counted the frequency of each unique problem, to be as objective as possible. She considered the most frequent problems to be of highest severity. Because frequency was a main driver in the researchers' problem severity ratings, her technique produced severity ratings that correlated well with the researchers' severity ratings. However, certification personnel may apply other criteria, such as perceived potential safety impact.

justifying solutions. Related observations were also included under the heading of the core problem.

5 Summary

This research was focused on developing tools that could be used by FAA field evaluators from the Aircraft Certification Service in conducting structured and comprehensive, yet practical, EFB usability evaluations. Two paper-based tools were developed for use by evaluators who are not human factors experts: a short high-level tool that can be used in a brief evaluation, and a longer detailed tool that can be used in a more comprehensive evaluation. The high-level tool consists of a set of topics that evaluators are asked to consider; every topic on the list should be covered to ensure a thorough evaluation. The detailed tool consists of over 200 items and is based upon the content of a lengthy document on EFB human factors considerations [6]. The tools were developed iteratively, over the course of several tests with volunteered prototype commercial EFB systems. The methods employed in the tests were adapted from standard techniques in usability engineering.

The tools add value to the evaluation process because they help evaluators without human factors expertise to ask the right questions. In addition, anecdotal evidence from tests of the tools against realistic commercial EFB systems shows that evaluators do uncover key usability issues with the tools. Note, however, that the tools developed for this research were not tested against any alternative tools, and no formal test of their value was performed.

In the course of this research, we learned that *language* is especially important for evaluating EFB usability, both in the design of the tool and in providing feedback to the manufacturer. The language of the tools must be understood by a wide range of users in order to maximize their utility. We fine-tuned the language of the high-level tool so that it is usable by those without human factors expertise. However, having a basic understanding of general user interface terms (e.g., select and click) is helpful.[11] The language of the detailed tool is best understood by those who are familiar with the full document on which it is based [6]. In addition, the language of the feedback report to the system developers must be clear, constructive, and meaningful to the designers. The report needs to present issues that are actionable and on which progress can clearly be tracked.

Another important lesson from our tests is that having detailed notes of the evaluation, and then using these notes in a synthesis step, enhanced the quality of the feedback significantly. The notes should capture, reasonably well, a real-time transcript of the session, without bias or personal interpretation. An observer who has experience taking notes and basic prior experience with the system is the best person to do this job. The synthesis step helps evaluators to step back and think about the problems they encountered at a higher level, promoting the development of more global problem descriptions, which in turn produce more global problem solutions.

[11] One source for the definitions of common user interface terms is the Glossary in [18].

6 Areas for Further Research

The EFB usability assessment tools have matured significantly over the past year. Still there are small areas for improvement in the design of the tools and in the selection of the evaluation tasks, and there are some broad areas of research that have not been addressed by this effort.

In terms of improving the tool design, we have a few relatively minor open issues. First, we are considering whether and how to incorporate references to relevant FAA policy within the high-level tool. These policy references are included in the detailed tool already. Second, we are considering how to improve the utility of the detailed tool using a searchable electronic format. For example, a data base version of the 200-plus items could be searched via query for specific types of items (e.g., "extract all items that are relevant to electronic documents that are called out in FAA policies such as AC 120-76A").

In order to facilitate EFB evaluations, another area for improvement is to improve the generic task list given in Appendix F. That list was created by collecting all the tasks that we used in our tests and generalizing the instructions. However, there are many other tasks given in Chandra et al. [6] that could also be considered as a starting point for a generic task list.

Some broad research tasks were not addressed by this effort. First, the tools were not formally validated. By validation, we mean that we did no formal tests to document whether more issues were uncovered with the tool versus without the tool. In fact, quantifying the benefit of this type of usability-assessment tool is a difficult task because of the high variability of task performance and significant individual differences. We know that some evaluators find more usability issues than others, but it is not clear why, or whether there is even a consistent pattern between evaluators [19]. We also did not formally study how a background knowledge of human factors affects how the tool is used. For example, we do not know how human factors experts might benefit from the tool, versus how those who are not human factors experts might benefit. Anecdotally, we found that human factors experts used the tool more as a reference than as a step-by-step guide to conducting the evaluation, as those without human factors expertise did.

In addition, the tools were developed for the needs of FAA; They do not satisfy all of the usability testing needs of manufacturers. In particular, there are three shortcomings of the usability assessment tools and process from a manufacturer's point of view. First, the design of an EFB, from the manufacturer's viewpoint, must satisfy not only the FAA but also their intended customers. Therefore, testing with end users is necessary, and many operational complexities (e.g., regarding work flow) must be understood in order to optimize the design. Second, the tools do not provide detailed quantitative results (e.g., time to complete a task, number of steps, or number of errors made). The only quantitative results from the high-level tool come in the form of frequency of problem occurrence. Quantitative results from usability testing are important for justifying the cost of resources to address problems, and could also be important data for more formal FAA evaluations (e.g., for installed EFB systems [10]). Finally, results from the Volpe tools highlight problems in the design but do not always specify solutions. Additional usability tests will be necessary to choose between design options.

Even with the present scope, the Volpe EFB usability-assessment tools may be quite useful to manufacturers. In fact, some manufacturers are beginning to try out the tools and are considering how to fit the tools into existing design and development processes. For example, some are using the detailed tool as an aid to system designers and developers. Using the tools and test methods described in this report could be a relatively inexpensive way for manufacturers to catch significant problems early on and to track progress on addressing these problems. Preliminary feedback from manufacturers suggests that this is the case.

The immediate next step for this research effort, however, is to introduce more potential users, especially those in FAA Aircraft Certification, to the tools and methods to determine if these products are useful in practice as designed, or with minor modifications. Later, we would also like to explore the utility of the tools to other FAA users, such as field evaluators in FAA Flight Standards.

7 Conclusions

This report describes research towards developing tools to aid FAA Aircraft Certification field evaluators in conducting human factors evaluations of EFBs. This is an important step towards bringing a high-level knowledge and understanding of human factors issues down to the level at which that knowledge is actually used to help make decisions about the usability of these devices.

Two tools were developed, each with different strengths. The short high-level tool is particularly promising for use by the FAA in brief onsite evaluations. This tool helps to remind evaluators about the many different aspects of the user interface to review. The simple list format of the tool is easy to use, but yields a relatively thorough evaluation because each item on the list triggers the evaluator to explore the user interface in depth. Using the high-level tool in a team evaluation is a particularly good way of collecting a large amount of data about the usability of the device quickly. The lengthy detailed tool is valuable in checking the system for specific characteristics and functionality. It is especially useful in identifying highly specific issues (e.g., appropriateness of font size) because many items are concrete and fairly objective. The detail provided on EFB-specific functions make the detailed tool particularly relevant to devices that perform those functions. Because of the length of the detailed tool, however, its use is limited in short evaluations.

From a theoretical perspective, the most important aspect of this research has been to construct a new taxonomy of user interface qualities that is designed explicitly for use by nonexperts, i.e., the high-level tool. Although formal validation of the high-level tool remains to be done, the results of tests to date are quite promising because we have found that nonexperts can identify key usability issues in a brief evaluation timeframe using the high-level tool. In the long term, we may even find that the applicability of the high-level tool, and the general components of the detailed tool, extend beyond EFBs; Although these tools were designed for EFB evaluations, they contain a substantial amount of generic material that could be applied to any flight deck system.

8 References

1. Cardosi, K.M. and Murphy, E.D., 1995, Human factors checklist for the design and evaluation of Air Traffic Control systems,. (Report No. DOT-VNTSC-FAA-95-3.1), Cambridge, MA: USDOT Volpe Center.
2. Chandra D. C. & S.J. Mangold. (2000). *Human factors considerations in the design and evaluation of electronic flight bags (EFBs) Version 1: Basic functions*. (Report No. DOT-VNTSC-FAA-00-22). Cambridge, MA: USDOT Volpe Center.
3. Chandra, D. C. (2002). Human Factors Evaluation of Electronic Flight Bags. *Proceedings of HCI–Aero 2002*, pp. 69–73. Menlo Park, California: AAAI Press.
4. Chandra, D. C. (2003). A tool for structured evaluation of electronic flight bag usability. *Proceedings of the 22nd Digital Avionics Systems Conference*, pp. 13.E.2.1–13.E.2.8. Madison, Wisconsin: Omnipress.
5. Chandra, D. C., Yeh, M., & V. Riley (in press). Designing and Testing a Tool for Evaluating Electronic Flight Bags. To appear in *Proceedings of HCI–Aero 2004.*
6. Chandra, D., Yeh, M., Riley, V., & S.J. Mangold (2003). *Human factors considerations in the design and evaluation of Electronic Flight Bags (EFBs), Version 2*. DOT-VNTSC-FAA-03-07 and DOT/FAA/AR-03/67. Cambridge, MA: USDOT Volpe Center. Available at www.volpe.dot.gov/opsad/efb.
7. Chandra, Divya C. & Mangold, Susan J. (2000) Human factors considerations for the design and evaluation of Electronic Flight Bags. Proceedings of the 19th Digital Avionics Systems Conference. 10-12 October 2000, Philadelphia, PA.
8. Dumas, J.S. and J. Redish (1999). *A practical guide to usability testing.* Exeter, England: Intellect.
9. Federal Aviation Administration, Advisory Circular AC 25-11. July 16, 1987. *Transport Category Airplane Electronic Display Systems*.
10. Federal Aviation Administration, Advisory Circular AC 120-76A, March 17, 2003. *Guidelines for the certification, airworthiness, and operational approval of electronic flight bag computing devices.*
11. Federal Aviation Administration Order 8110.4B, "Type Certification Process," April 24, 2000. Available from http://www.faa.gov/certification/aircraft/, under the "Original Design Approval Process" heading of the "Design Approvals" column.
12. Federal Aviation Administration Advisory Circular (AC) 21-40, "Application Guide for Obtaining a Supplemental Type Certificate," May 6, 1998. Available from http://www.faa.gov/certification/aircraft/
13. Federal Aviation Administration Policy Statement ANM-99-2, *Guidance for Reviewing Certification Plans to Address Human Factors for Certification of Transport Airplane Flight Decks*. (1999). Available from http://www.faa.gov/certification/aircraft/. Go to the "Policy Statements" link under the "Aircraft Certification Software" heading of the "Representatives of the Administrator (Designees)" column (September 22, 2004).
14. Federal Aviation Administration Policy Statement ANM-01-03, *Factors to Consider When Reviewing an Applicant's Proposed Human Factors Methods for Compliance for Flight Deck Certification.* (2001). Available from http://www.faa.gov/certification/aircraft/. Go to the "Policy Statements" link under the "Aircraft Certification Software"

heading of the "Representatives of the Administrator (Designees)" column (September 22, 2004).

15. Hackman, G.S., and D.W. Biers (1992) Team usability testing: Are two heads better than one? In *Proceedings of the Human Factors Society 36th Annual Meeting*, pp. 1205-1209.
16. Huntley, S., J.W. Turner, C.M. Donovan, E. Madigan, 1995, FAA Aircraft Certification Human factors and operations checklist for standalone GPS receivers (TSO C129 Class A), (Report No. DOT-VNTSC-FAA-95-12), Cambridge, MA: USDOT Volpe Center.
17. Mayhew, D., 1999, The Usability Engineering Lifecycle: A practitioner’s handbook for user interface design, San Francisco, CA, Morgan Kauffman.
18. Microsoft Corporation. *The Windows™ Interface Guidelines for Software Design.* (1995). Redmond, Washington: Microsoft Press.
19. Nielsen, J. & R.L. Mack (Eds.) (1994). *Usability Inspection Methods*, New York, John Wiley & Sons, Inc.
20. Nielsen, J., 1994, Usability Engineering, San Francisco, CA, Morgan Kauffman.
21. Shamo, M. (2000). What is an electronic flight bag, and what is it doing in my cockpit? *Proceedings of HCI–Aero 2000.* Toulouse, France.
22. Wilson, C. (1998). (Blostein, J, Ed.) Usability Techniques: Pros and cons of co-participation in usability studies. Usability Interface, 4(4). Retrieved June 22, 2004 from http://www.userforum.dk/content/artikler/utech.html

Appendix A:
Tools from March 2003

A.1 EFB Checklist from March 2003

This assessment tool consists of a comprehensive listing (183 items) of requirements and recommendations taken from Chandra, Mangold, and Riley (2002). The checklist consisted of a filtered and condensed paper version of the full Volpe EFB Human Factors report draft of September 2002. To produce the list, only items that could be addressed in a bench test were included. In other words, all items pertaining to aircraft installation and crew training were not in the checklist tool. The remaining topics were then paraphrased and grouped loosely. Furthermore, all Requirements were placed in one section, and Recommendations in another section. Each of these sections was further divided by function, as follows:

- System Considerations
- Electronic Documents
- Electronic Checklists
- Flight Performance Calculations
- Electronic Charts

Instructions to participants on how to use the checklist and the checklist itself are provided in the next pages. The format was designed for efficiency; items were labeled with a topic so the evaluator could quickly determine if an item applied to a design, key words within the items were bolded so that participants could get the gist of the item without reading the full text. The response scheme was designed with an FAA evaluator in mind. With this response scheme, an evaluator could keep track of items that he/she had reviewed and identify a level of concern or note an absence of concern for each item. An "NE" (not evaluated) column was provided so that evaluators could mark that an item could not be evaluated at the time. The "-" column represented some level of concern about the item. The suggested response code scheme for the "-" column was:

P = problem that must be addressed
C = concern that should be addressed (or it may become a "problem")

O = an area for optimization (or it could become a "concern")

For the "√" column in the checklist, the following codes were proposed:

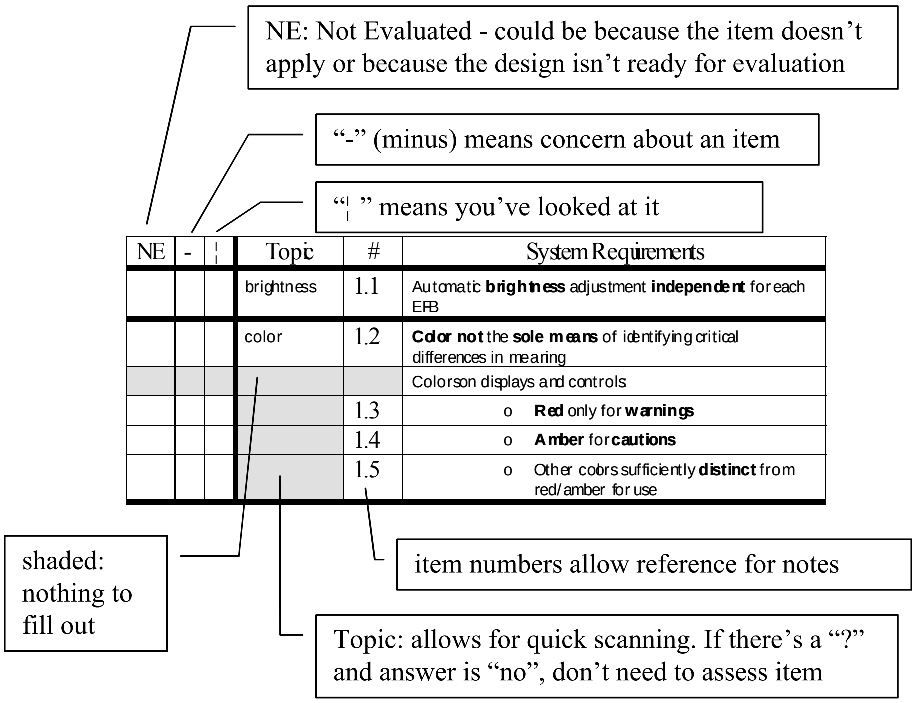
checkmark = okay
+ = system strength

The checklist was provided in paper form to the participants, and an observer recorded notes on an electronic version as the teams completed the tool. Participants were encouraged to write notes on the paper checklist in order to provide examples to support their ratings, or to enable the evaluator to return to the document at a later time to further consider a rating, issue, or concern.

A.1.1 Instructions to Evaluators

This checklist is intended for the evaluation of Electronic Flight Bag products. It contains the requirements and recommendations from Human Factors Considerations in the Design and Evaluation of Electronic Flight Bags (EFBs) (US-DOT, Volpe Center, 2002).

The following graphic shows the layout and intended use of the checklist:



You may want to enter your own codes into cells rather than just checkmarks. For example, in the "-" column, you may wish to record different levels of concern. One possible method is to use:

P = problem that must be addressed

C = concern that should be addressed (or it may become a "problem")

O = an area for optimization (or it could become a "concern")

Similarly, in the "Checkmark" column, you may want to indicate positive reactions:

checkmark = okay

+ = system strength

However, feel free to use any method you prefer.

A.1.2 Full EFB Checklist Tool from March 2003 Test

NE	-	√	Topic	#	System Requirements
			brightness	1.1	Automatic **brightness** adjustment **independent** for each EFB
			color	1.2	**Color not** the **sole means** of identifying critical differences in meaning
					Colors on displays and controls:
				1.3	o **Red** only for **warnings**
				1.4	o **Amber** for **cautions**
				1.5	o Other colors sufficiently **distinct** from red/amber for use
			alerts	1.6	For installed systems, **alerts** and **reminders** meet CFR 23.1322, 25.1322, 27.1322 or 29.1322 as appropriate.
			feedback	1.7	Provide **feedback** for **every** user **input**
				1.8	System **busy** indicator
			typefaces	1.9	**Typefaces appropriate** for viewing **distance** and **lighting** conditions
			labels	1.10	All physical **controls** properly **labeled** for their intended function
				1.11	All soft **function keys labeled** for their current intended function.
				1.12	**Inactive** soft **function keys** either **not labeled**, or the label uses a visual convention to indicate that that function is **not** currently **available**
			version info	1.13	EFB **provides** latest **revision information** to crew upon request
			default settings	1.14	If app can be customized, easy way to **return** to **default settings**
			accuracy	1.15	Databases checked for **accuracy**, **currency**, and **corruption** prior to installation

NE	-	√	Topic	#	Electronic Document Requirements
			consistency	2.1	Logical **structure** of an electronic document **consistent** with the **hard copy** version of that material, if a hard copy exists.
				2.2	Electronic version of a **figure** able to **display all** of the **content** of the paper version. User able to view the **overall figure** at one time, even if not all details are readable, in order to get an overview of the figure. User also able to read **all** the **details** in the figure, although not all of the figure may be visible when the details are readable.
			off screen content	2.3	**If** document segment is **not visible** in its entirety in the available display area, existence of **off-screen content** clearly **indicated** in a consistent way.
			active regions	2.4	**Active regions** clearly visually **highlighted**
				2.5	Means to **activate** and **deactivate** regions apparent.
			multiple open documents? indicate active	2.6	**Indicate** which document is **active**, and display that indication continuously. Also, under non-emergency, normal operations, user able to **choose** which of the open documents is currently **active**.
			animation?	2.7	User able to **start** and **stop** animated segments. User able to stop animation at any time, even if segment has not ended. Also, **supporting text** provided to describe and support animation. Text available even if animation not currently running.

NE	-	√	Topic	#	Electronic Checklist Requirements
			completeness	3.1	**All** checklists belonging to a supported category must be **available**
			reference to unsupported checklist?	3.2	If need access to an **unsupported checklist**, indicate **location** in paper document
			titles	3.3	Each checklist must have a **constantly visible title**
			normal sequence, accessibility	3.4	Normal checklists accessible in **normal sequence of use**. Also **individually accessible** at all times.
			non-normal accessibility	3.5	**Access** to individual **non-normal checklists** supported at all times
			closed loop? alert/display checklist	3.6	**Alert** crew that a checklist applies to a detected non-normal condition. Only **call up** appropriate checklist **when commanded** by crew.
			closed loop? master list	3.7	**All checklists** required to manage multiple non-normal conditions listed together in **one master list**
			multiple open checklists?	3.8	Able to **access other checklists** without having to close the currently displayed checklist first
			multiple open checklists?	3.9	Able to **choose** which checklist is currently active.
			multiple open checklists? multiple functions?	3.10	If more than one unrelated checklist can be open or if EFB supports multiple functions that can interrupt checklist completion, **placeholder** required to remind user which item was active prior to leaving checklist
			parent/child checklists?	3.11	Able to **choose** whether **parent** or **child** checklist is active
			active item	3.12	Indicate **active** item
				3.13	**Moving** active-item **pointer** to next checklist item requires only **simple action**
			track item status? move backward	3.14	Able to move **backward** to previous item without changing any item's status
			track item status? close incomplete	3.15	If user tries to close incomplete checklist, **indicate** checklist **not completed**. User can close incomplete checklist after **acknowledging** indication.
			track item status? complete deferred	3.16	Before checklist declared complete, user must **return** to **deferred** checklist items and **complete** or **override** them

NE	-	√	Topic	#	Electronic Checklist Requirements (cont.)
			user indicates completion?	3.17	Action to indicate completion **simple** and **distinct** from action of moving to next item
			user indicates completion? change status	3.18	If active item marked "complete," **changing** item to **different status** or **return** it to **uncompleted status** must be **simple** to accomplish
			ECL indicates status? clear indication	3.19	Clear **visual indication** of item status (active, deferred, overridden, uncompleted, closed-loop sensed) provided
			multi-screen checklist? look ahead	3.20	User can **look ahead** in a multi-screen checklist without changing active item. If user makes change to an active item that is out of view, that active item must be **brought into view**.
			calculation worksheets?	3.21	Easy access from checklist item to corresponding **worksheet** for both initial calculation and subsequent review and modification of calculated value. Values calculated in linked worksheet appear in corresponding checklist **location**. Corresponding checklist fields blank prior to insertion of calculated value.
			task reminders?	3.22	Reminders **displayed constantly** once in progress, **attract** the pilot's **attention** when action should be performed. If multiple task reminders possible at one time, crews able to determine **how many** are in progress and to **what tasks** they refer.
			branching?	3.23	When checklist branches based on key decision, selected **branch clearly indicated**. User able to **back up** to decision step and choose another branch
					Flight Performance Calculations Requirements
			units	4.1	**Units** of each variable clearly labeled
					Electronic Chart Requirements
			scale shown? accuracy	5.1	Chart **scale** information must be **accurate**. If accurate scale information not available, scale must **not** be displayed.
			own position	5.2	Own-aircraft position must **not** be shown on non-georeferenced or not-to-scale charts (e.g., arrival and departure procedures)

NE	-	√	Topic	#	System Recommendations
			brightness	6.1	Able to **adjust** screen **brightness** in **fine increments** or **continuously**
			labels	6.2	Labels **brief**, **clear**, and **consistent**
				6.3	Accompany icon with **text label**
				6.4	Able to access **help** information about icons with more detail than text label
				6.5	Labels should be **separated** from the graphical **drawing** area or visually highlighted
				6.6	Adequate **labeling** of **buttons** for **night** use
			guidance	6.7	Complies with industry-**standard guidance** material
			style guide	6.8	All apps follow **common style guide**, preferably consistent with that aircraft
				6.9	Style guides for **third party developers**
			conventions	6.10	Follows **personal computer conventions**, except where clearly inappropriate for flight deck
			consistent controls and graphics	6.11	**Consistent** set of **controls** and **graphical** elements for **similar** functions
			distinctive controls	6.12	Controls with **different** functions visually **distinct**
			feedback	6.13	**Progress** indicator
			consistent color and formatting	6.14	**Consistent** conventions for **color** and other **formatting**, standard **help** facility, standard **labeling** and **menus**, same means of accessing **common actions**
			colors	6.15	Colors do **not violate** flight deck **conventions**
				6.16	Color **not** used as **sole means** of identifying critical differences between information.
				6.17	Colors codes **not** be **user-customizable**
				6.18	**Green** is generally assigned to indicate a **safe** condition
				6.19	**Magenta** and **white** may also be **reserved** in some flight decks
			colors customizable?		If colors are **customizable**, then
				6.20	o Easy way to return to **default**
				6.21	o **Restrict** users from setting colors that **conflict** with flight deck **convention**
				6.22	o Red **not** be used for **informational** purposes
			applications	6.23	User able to **select** the **active** app
				6.24	Easy **identification** of **active** app
				6.25	Easy to **switch** between apps
				6.26	When returning to a background app, should be in **same state**, other than completion of background processing
				6.27	System **responsiveness** does not suffer, even if all apps are running simultaneously
				6.28	Able to **switch** apps, even with **pending activities** after user acknowledgement
				6.29	Non-flight-related apps do **not** have **reminder message**, and require **extra confirmation** step to launch

NE	-	√	Topic	#	System Recommendations (cont.)
			applications	6.30	**Anchor location** between apps
				6.31	**Easy access** to anchor location from any app
				6.32	Able to keep **track** of how to move between **different topics** easily
				6.33	Able to **return** to **start point** easily
			system busy	6.34	**Storage** of user **entries** while system is busy for later processing (when able)
			icons	6.35	Minimize icon **training**, maximize **cross-cultural intuitiveness**
				6.36	Graphical icons and symbols **interpretable** based on **shape alone**, without needing secondary cues such as color
				6.37	**Legible** on **min** expected display **resolution** viewed from **max** intended viewing **distance**
			alerting	6.38	**Positive indication** of any full or partial application **failure**
				6.39	**Immediacy** of indicator **appropriate** to the function that failed
				6.40	Alerts **integrated** or **compatible** with other flight deck alerts
				6.41	Messages **prioritized**, and prioritization scheme documented and evaluated
				6.42	Both visual and auditory **messages inhibited** during **high workload** phases of flight
				6.43	**No interruptions** during **high workload** phases except those indicating failure/degradation of current app
				6.44	**Avoid flashing** text or symbols
			audio	6.45	Audio has **user-controlled volume**
				6.46	**Visual indicator** of accompanying audio (either graphic or text)
				6.47	Able to **stop** supplemental audio at any time
			accessibility	6.48	**Consistent philosophy** for accessing **different** types of information
				6.49	**Similar** types of information accessed in the **same way**
				6.50	Functions accessible in proportion to **frequency** of use and **criticality** to mission
			screen	6.51	**Active areas** on the screen sized to permit **accurate selection** with pointer/cursor device under all operating conditions
				6.52	Display easy to **clean**
			keyboards	6.53	Keyboards **appropriate** for the given **task**
				6.54	**Keyboards** provide appropriate **tactile feedback**
				6.55	Able to **rest/stabilize hand** to use the **keyboard**, especially during turbulence
				6.56	**Physical function keys** provide **tactile feedback**
				6.57	**Key repeats** filtered by software if too close together
				6.58	**Soft function keys** assigned **consistently** to functions

NE	-	√	Topic	#	System Recommendations (cont.)
			cursor control	6.59	Able to **rest / stabilize hand** when using **pointer** or **cursor control device**
			text	6.60	Typeface highly **legible**
				6.61	P/R, B/D/E, E/G/O/C, I/1, and Z/2 **not confusable**
				6.62	**Uppercase** used **sparingly**, avoid italic
				6.63	**Constant stroke widths**, **sufficient contrast** between character and background
				6.64	Typeface at least **1 / 200** of viewing distance
				6.65	o For 30" viewing distance, 0.15" tall (15 pixels at 100 pix/inch)
				6.66	o For 25" viewing distance, 0.125" tall
				6.67	o For 35" viewing distance, 0.175" tall
				6.68	Text **larger** at **lower illumination**
					Character **height to width ratios** (especially important for documents)
				6.69	o <80 char per line, 1 to 0.7 up to 0.9 (15 pix tall, 10.5 to 13.5 pix wide) for monotype fonts
				6.70	o >80 char per line, at least 1 to 0.5 (15 pix tall, 7.5 pix wide)
				6.71	o 1:1 for M and W in a proportional font
				6.72	**Horizontal spacing** between characters of 0.01 character height (15 pix tall, 1.5 pix spacing)
				6.73	**Space between words** is one character for monotype fonts, or width of "N" for proportional fonts
				6.74	**Line leading** (vertical spacing) at least two stroke widths or 0.15 of character height (15 pix tall, 2.25 pix leading), whichever is greater
				6.75	**Line lengths** appropriate for text content

NE	-	√	Topic	#	Electronic Document Recommendations
			OS style guide	7.1	Document interface **consistent** with **operating system style** guide
			consistent interface	7.2	Document **interface**, including text, color, and formatting, internally **consistent**
				7.3	Similar **types** of information in similar **locations**
			brevity	7.4	Data formatted into **short segments**, each communicating **one clear point**
			min display	7.5	Documents formatted to **comply** with manufacturer's **minimum display area** requirements
			visibility	7.6	**Necessary** information **not off-screen**
				7.7	If information is off-screen, shows how **far** into the document the currently displayed segment is and how **long** the document is
			descriptive text	7.8	Figures have **descriptive text**
			links?	7.9	Table of content and index entries **linked** to corresponding **locations** in text
				7.10	**Cross references linked** to each other, both within and across documents
			search	7.11	Supports **multiple search techniques**, such as by key word, hyperlinks, and header/footer information
			bookmark	7.12	Able to **bookmark** selected text locations
			history	7.13	Electronic document application keeps track of **most recently visited locations** in the document, allows user to select from this list to **return** quickly to a **recent location**
			cancel	7.14	Able to **cancel** a movement by returning to previous location in **one step**
			multiple open documents?	7.15	Master list of **all open documents** available
			animation	7.16	*Animation **not overused***

NE	-	√	Topic	#	Electronic Checklist Recommendations
			combined checklist	8.1	In place of parent-child checklists, create a **single checklist** that incorporates both
			multiple open checklists?	8.2	**Master list** of all **open** checklists provided
				8.3	**Master list** of checklists **to be completed** for managing a non-normal condition indicates the status of each checklist (e.g., pending, or completed)
			track item status?		Able to:
				8.4	• **Move** from uncompleted item to next item, changing status of uncompleted item to "**deferred**"
				8.5	• **Move** to next item **automatically** after completing an item
				8.6	Easy **undo** available to correct erroneous completion check
			delayed action	8.7	Non-normal checklist items to be performed at a later time **automatically integrated** into **subsequent** checklists
				8.8	**Reminders** for tasks that require a delayed action to ensure that the task is completed at the appropriate time
			multi-screen		While **multi-screen** checklist in use, the following information should be constantly be available:
				8.9	• How **long** the whole checklist is
				8.10	• How **far down** the checklist the currently **displayed** information is
				8.11	• How **much** of the checklist has been **completed**
				8.12	**Positive indication** that checklist as a whole, as well as each item in that checklist, is **complete**
			related information	8.13	**Links** to information related to individual items provided when that information is also part of the EFB. Links could direct users to **additional information** about that item, about the **system** addressed by the item, and/or to **MEL** information for that system.
				8.14	**Single step** action to **return** to the item from the related information
				8.15	Related information appears in a **single window** or **area** of the screen.
				8.16	Activating hyperlinks within the related information **updates** the information in that one **window** (or area), rather than opening a separate window with the new content
			unselectable items	8.17	Items that are not on the selected branch are **not selectable**
			modifications	8.18	ECLs easily modifiable by **airline personnel**

NE	-	√	Topic	#	Flight Performance Calculations Recommendations
			invalid entries	9.1	Does **not** accept user-entered data that are not of the correct **format** or **type** needed by the application. Error message provided that communicates **which entry** is suspect and **what type of data** is expected.
				9.2	When incorrect item is identified, **only** that item discarded, not the whole set of entries related to the task in progress
			consistency between sources	9.3	Labels, formats, and units used in software **match** the labels, formats, and units available to the user from **other sources** (e.g., paper reports)
			related information	9.4	Related information, such as data stored within the application or computations for cross-checking, is **in view** or **easily accessible**
			default values	9.5	Default values based on **most conservative** parameters for that calculation

NE	-	√	Topic	#	Electronic Chart Recommendations
			logical structure	10.1	**Logical structure matches** that of **paper charts** (information grouped in paper charts also grouped in electronic charts)
			visual structure	10.2	**Visual structure**, including symbology and general layout, **compatible** with paper charts (exact copy not necessary)
			updates	10.3	Corrections/updates made **electronically**, not via paper notifications to flight crews.
			hard copies	10.4	EFB generated hard copies printed at **same size** as paper chart prepared by that manufacturer, or larger. If more than one option available, **user selects** chart size.
			color?	10.5	EFB generated **hard copies** also in **color**, or **all information preserved** and **visible**
			scale	10.6	Accurate chart **scale visible**, especially if display can be zoomed
			zoom?	10.7	Visual **edges clearly marked**, only used when no further information outside that area
				10.8	**Panning** also supported. (Similarly, if chart application supports panning, it should also support zooming.) User always knows which way to move to bring more of chart into view.
				10.9	Easy way to **return to default** zoom settings
				10.10	**Accuracy** of the **ownship** display **maintained** across zoom levels
				10.11	**Cannot** zoom beyond the point where information is **no longer accurate.** (For example, if a symbol is accurate within several hundred feet, it should not be possible to zoom in so close that accuracy within a few feet is implied.)
			track-up and north-up?	10.12	Visual format differences between the two must be **salient**, cannot confuse the two
				10.13	**Text** and individual **symbols** (e.g., navaid or waypoint symbols) **not rotated** on track or heading-up charts (difficult to draw rotated text/symbols cleanly and difficult to read rotated text)
			title	10.14	**Title** of currently selected chart always in view
			pre-select	10.15	Able to pre-select charts for later **quick display** (e.g., departure airport)
			error management	10.16	Promotes **good error management** (helps crew select correct chart, allows common corrections to be made quickly)
			search	10.17	Supports **multiple search methods** (e.g., by name, geographical region, present position). Search results ordered intuitively (best guesses at top, least likely at end).
			alternate runways	10.18	Selection of alternate runways as **easy as possible** during approach
			precision	10.19	Ownship depiction and **precision** of chart elements **appropriate** for the **accuracy** of the **data**, does not imply more precision than is merited

NE	-	√	Topic	#	Electronic Chart Recommendations (cont.)
			declutter?	10.20	**Cannot remove safety critical information**, such as terrain, obstructions, or special use airspace without knowing that it is suppressed. (If such information can be de-cluttered, should not be possible for pilot to believe that it is not shown because it is not there.)
			workload	10.21	Managing display configuration does **not** cause significant workload. Routine display configuration changes minimized.

A.2 Quick Assessment Tool (QAT)

The Quick Assessment Tool (QAT) is an 18-item set of usability topics that was constructed for more general usability evaluations. These topics represent dimensions along which the quality of the user interface can be assessed. The full set of terms is listed in Table A-1 below.

1. Functions And Options Visibility	10. Functional Or Navigation Logic
2. Modes	11. Steps
3. Feedback	12. Task Match
4. Perceptual	13. Colors
5. Icons/Symbols	14. Ergonomic
6. Destructive Adjacencies	15. Error Penalties
7. Automation	16. Formatting
8. Distinctiveness	17. Terminology
9. Responsiveness	18. Control/Display Relationships

Table A-1: List of Terms in Quick Assessment Tool

As Table A-1 shows, a variety of topics are included in order to ensure a systematic and thorough review. For each topic, the reviewer is asked for an overall rating, and for further comments. The comments that are generated are expected to be more valuable than the ratings per se, so we attempted to make it as easy as possible to record comments. We expected that the reviewers would provide specific examples, as well as general observations in their comments.

The QAT was tested to determine if the format and content was appropriate and applicable for identifying user interface issues in an EFB and whether it would be usable by both human factors experts and non-human factors experts. Note, however, that (a) the tool was originally designed as an aid for one experienced human factors expert, so the terms are not necessarily intuitive to evaluators in general, and (b) the topics/dimensions are not orthogonal with each other, meaning that there may be some overlap between them. For example, a problem getting from one place to another in the interface could be reflected in "Functions and Options Visibility" (i.e., the user could not find the option that would take him/her to the right place) or "Functions or Navigation Logic" (i.e., the logic used in the interface did not match the user’s mental model), or "Steps" (i.e., there were too many steps involved in getting from one place to another).

Two versions of the QAT were used. One was an electronic form prepared in Microsoft Excel, and the other was a paper version. A sample of the electronic version is shown in Figure A.1 below.

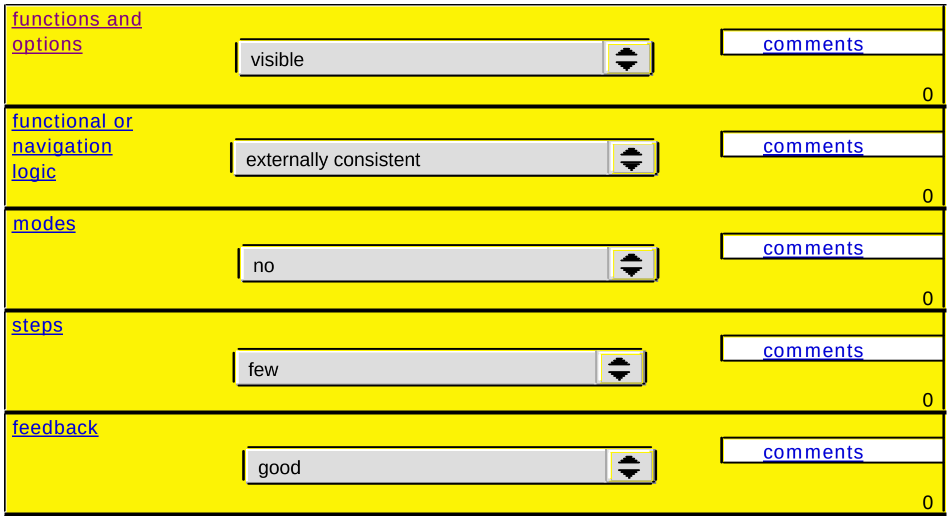


Figure A-1: Electronic format of QAT

To use the electronic form, the evaluator selects a rating for each item by pulling down a menu and selecting their response. For example, the response options for "Functional and Options Visibility" are "visible," "guided," "findable," "hidden," and "invisible." Definitions for the terms could be called up by clicking on the item name. Users could also type in comments within the electronic QAT, on a separate spreadsheet.

There was one small difference between the electronic and paper forms of the QAT. In the electronic form, each option was assigned a numeric value, and the rating worksheet kept track of the total "score" for the system. The total score was the sum of the numeric values for each response. The spreadsheet was designed such that higher numeric values were given to "worse" ratings. In other words, higher scores indicated a more problematic user interface. Although these scores were available, we felt that it was difficult to assign any meaning to them. Also, we did not explore the reliability and validity of these scores. As a result, these data were not evaluated. If desired, the development of numerical scores for rating EFB usability could be pursued as a separate area for future research.

The paper version is shown on the following pages. To use the paper form, the user reads the name of the term at the top left and its definition to its right. The rating scale is given in the rows below the term, with definitions of each rating to their right. A generous area for comments is available, and more comments can be written in below the table on the same page if needed.

functions and options visibility	Refers to how easy it is to find the features, functions, and options that the product makes available to the user.	Please circle your rating:	Definition:
Comments:		**visible**	all functions and options are visible at the top layer of the interface
		guided	the interface guides you to the functions or options you want
		findable	some functions or options are not where you expect, but you can find them in a reasonable amount of time
		hidden	some functions or options are not where you expect and they're difficult to find
		invisible	the product provides some functions or options that are not visible anywhere in the interface (for example, push-and-hold buttons)

More Comments:

functional or navigation logic	Refers to how easy it is to learn and remember the underlying logic of accessing and operating functions and navigating around the interface.	Please circle your rating:	Definition:
Comments:		**externally consistent**	the logic is consistent with an existing user mental model (for example, the desktop metaphor that makes PCs easy to use)
		internally consistent	the logic isn't consistent with an external mental model, but is internally consistent with itself
		inconsistent	the logic has some rationale behind it, but the rationale is different for different functions or pages, or is applied inconsistently
		arbitrary	the logic is inconsistent for arbitrary reasons
		externally inconsistent	the logic conflicts with existing user mental models or cultural conventions (this is the most likely to cause errors)

More Comments:

modes	Modes exist where the same user action (input) can have different outcomes depending on the state of the system, such as where the same control can access multiple functions.	Please circle your rating:	Definition:
Comments:		**no**	every control does the same thing regardless of state
		simple	the product has relatively few modes and the transition logic is simple and straightforward
		complex	the product has many modes and/or the selection/transition logic is potentially complicated

More Comments:

steps	Describes the number of steps needed to accomplish tasks. However, a tradeoff often exists between efficiency and ease of learning. Minimizing steps may require making the steps obscure and hard to learn, while guiding the user through a process may require more steps. When evaluating this, consider whether steps are necessary for ease of learning or ease of use.	Please circle your rating:	Definition:
Comments:		**few**	the design is relatively efficient - no wasted steps
		some	there are some inefficiencies, unnecessary steps in some cases
		many	there are many unnecessary steps

More Comments:

feedback	This has to do with the response of the system to user inputs. More loosely, it can also refer to the system's indications of its own status.	Please circle your rating:	Definition:
Comments:		**good**	the product provides visible response to every user action, and these responses are consistent and appropriate
		inconsistent	the product provides inconsistent responses for user actions or inconsistent status indications
		poor	the product doesn't provide adequate responses for user actions or adequate status information

More Comments:

task match	This refers to the extent to which the interaction logic and interface match the logic of the task, and the user's response is compatible with the information coming into the task.	Please circle your rating:	Definition:
Comments:		**good**	the way the user interacts with the product is consistent with the logic of the task and the information coming to the user from the task environment
		inconsistent	the user may have to make mental estimations, transformations, or calculations, or reorganize the information coming into the task before making a response
		poor	the user has to translate from the logic of the task to the logic of the system while using the interface

More Comments:

perceptual	This refers to all of the perceptual aspects of the design: display and control visibility; display clarity, resolution, contrast; audio clarity, etc. Consider all of the operational environments in which the product may be used, and all potential users.	Please circle your rating:	Definition:
Comments:		**good**	there are no apparent difficulties with reading and understanding displays, labels, indicators, etc.
		inconsistent	there may be some difficulties
		poor	there are obvious difficulties

More Comments:

colors	This refers to the use of color on displays, on the product faceplate (if applicable), controls, labels, etc.	Please circle your rating:	Definition:
Comments:		**good**	color is used effectively and consistently, and does not violate any applicable standards or user expectations
		inconsistent	mostly good but some concerns
		poor	color is not effective and/or consistent, but still does not violate any applicable standards or user expectations
		incompatible	some uses of color violate applicable standards and/or user expectations (most likely to cause error or distraction)

More Comments:

icons/symbols	This refers to the ease of interpreting icons and symbols.	Please circle your rating:	Definition:
Comments:		**clear**	it's immediately obvious what icons and symbols are supposed to mean
		obscure	it's difficult to interpret icons or symbols
		error-prone	icons or symbols are potentially misleading, suggesting that they do something different than what they actually do

More Comments:

<table>
<tr><td>ergonomic</td><td>This refers to all of the physical properties of the product: posture while using, ease of accessing and using controls, ease of viewing displays, force required, repetitive motions, etc.</td><td>Please circle your rating:</td><td>Definition:</td></tr>
<tr><td colspan="2" rowspan="3">Comments:</td><td>good</td><td>everything is accessible and there are no concerns about undue force, repetitive motions, posture, etc.</td></tr>
<tr><td>inconsistent</td><td>there are some concerns</td></tr>
<tr><td>poor</td><td>the product has major ergonomic problems</td></tr>
</table>

More Comments:

destructive adjacencies	This refers to the presence of two controls, function keys, options, etc., that are located next to each other in such a way that the user may select one when intending to select the other, AND where doing so would cause a serious problem or inconvenience. For example, a DELETE button located next to an EDIT button.	Please circle your rating:	Definition:
Comments:		**no**	there are no instances of this
		yes	there is at least one instance

More Comments:

error penalties	This refers to how serious the consequences of making an error are, to safety, user time or convenience, etc.	Please circle your rating:	Definition:
Comments:		**none**	the product handles errors well and gives the user the option to reverse the effects of any error, or errors are essentially inconsequential
		low	some errors may have minor consequences
		high	some errors that are difficult or impossible to reverse may have major consequences

More Comments:

<table>
<tr><td>automation</td><td>If the product doesn't have any automation, choose "supportive". If it does, this refers to the extent to which any automation supports the user's goals, objectives, and task strategies.</td><td>Please circle your rating:</td><td>Definition:</td></tr>
<tr><td colspan="2" rowspan="3">Comments:</td><td>supportive</td><td>there is no automation, or automation supports the user's goals, objectives, and task strategies</td></tr>
<tr><td>intrusive</td><td>automation can intrude in user tasks, or using automation is difficult or inconvenient</td></tr>
<tr><td>preemptive</td><td>automation can lock the user out of tasks and/or prevent the user from doing the task they way they want to</td></tr>
</table>

More Comments:

formatting	This refers to all aspects of the display design, control layout, etc., including page layout, menu design, dialog boxes, soft control positioning, organization, clustering, white space, text formats (font type and size, emphasis, etc.), and so forth.	Please circle your rating:	Definition:
Comments:		**externally consistent**	the formatting is consistent with an existing user mental model (for example, corresponding paper documents, similar familiar computer applications, etc.)
		internally consistent	the formatting isn't consistent with an external mental model, but is internally consistent with itself
		inconsistent	the formatting has some rationale behind it, but the rationale is different for different functions or pages, or is applied inconsistently
		arbitrary	the formatting is inconsistent for arbitrary reasons
		externally inconsistent	the formatting conflicts with existing user mental models or cultural conventions (this is the most likely to cause errors)

More Comments:

distinctiveness	This refers to the extent to which controls, displays, and indicators with dissimilar functions have appropriately dissimilar appearances. If they look too much alike, the user may select one while intending to select the other.	Please circle your rating:	Definition:
Comments:		**good**	controls, displays, and indicators with dissimilar functions have appropriately distinctive appearances
		inconsistent	some may be confused with each other
		poor	many may be confused with each other

More Comments:

terminology	This refers to the words that are used for the interface, including controls, indicators, user options and functions, alerts, dialog boxes, menus, etc. A common problem is that the designer will often select terms that reflect how he or she thinks about the product rather than how the user thinks about the task.	Please circle your rating:	Definition:
Comments:		**externally consistent**	all terms are consistent with the task environment, other materials used in the task, standards and conventions, and user expectations
		internally consistent	there are no appropriate external entities that the design should be consistent with, so the design is internally consistent with itself
		inconsistent	there are no appropriate external entities that the design should be consistent with, and there are some internal inconsistencies in terminology
		arbitrary	different terms are used for the same function or option in some cases
		externally inconsistent	some or many terms are not consistent with applicable external materials used in the task, standards, conventions, or user expectations

More Comments:

responsiveness	This refers to how quickly the system responds to user inputs. Feedback to acknowledge user inputs should seem instantaneous. Processes that require extended processing time should be indicated with system busy indicators and, if particularly long, progress indicators.	Please circle your rating:	Definition:
Comments:		**fast**	system responses are fast enough to avoid slowing the user down, prevent the user from making unneeded double entries, and satisfy user expectations
		moderate	some responses take more time than desired
		slow	system responses are overall slower than desired, may slow the user down and cause user to attempt inputs multiple times

More Comments:

control/display relationships	This refers to the extent to which any relationships between controls and associated displays are clear and obvious, and the operation of a control is consistent with the behavior and response of the associated display. An example is the speed with which values change in response to turning a knob, and the tradeoff between precision and speed.	Please circle your rating:	Definition:
Comments:		**externally consistent**	control/display relationships are consistent with other systems the user is familiar with, and display behaviors are appropriate for given control inputs
		internally consistent	control/display relationships are consistent across the design, and display behaviors are appropriate for given control inputs
		inconsistent	there are some relationships and/or behaviors that are inconsistent with others in the design
		arbitrary	there are some instances where the relationship between controls and displays is obscure and apparently arbitrary
		externally inconsistent	some control/display relationships are inconsistent with other systems the user is familiar with (most likely cause of errors)

More Comments:

A.3 Evaluation and Results

A.3.1 Method

Six aviation/human factors researchers worked in pairs to evaluate an EFB with the two assessment tools after a period of guided and free exploration of the system. Participants were encouraged to refer to the unit while they completed their reviews. Each participant completed the Quick Assessment Tool independently, but the EFB Checklist was completed as a team, with negotiated responses to each item. This co-discovery process appeared to elicit comments about the terminology, scope, and applicability of each of the Checklist items. After using the first tool, participants completed a survey regarding the ease-of-use, appropriateness, clarity, and usefulness of that tool before using the next tool. The questionnaire asked the participants about whether the briefing was adequate, whether they found the tool useful, and what specific aspects of the tools were confusing, unclear, or could use improvement. In addition to the questionnaire results, we also learned much from observing the co-discovery sessions and discussing the tools with the participants.

We varied the order of the presentation of the two tools to determine whether and how much completing one instrument affected completing the other. Two of the three teams worked on the EFB checklist first. Also, two of the three teams used an electronic version of the QAT.

As part of the evaluation, we also asked one system designer to use the tools to get his feedback on the tools. This evaluation was conducted at the manufacturer's site. This single-session was primarily useful in terms of getting feedback on specific topics in the Checklist tool. We were not able to combine that data with the data obtained from the co-discovery sessions held at Volpe.

Results are presented in two sections below. First, we present data from the co-discovery sessions. Second, we present our thoughts on the synthesis stage of the evaluation, because this turned out to be a critical step in our evaluation process.

A.3.2 Results

Our main goal was to evaluate the human factors assessment tools. The surveys that were completed by and large confirmed our more subjective observations and discussions. Discussions with the participants, and observations of their use of the tools also provided a wealth of information. These data taken as a whole provide a set of goals for the redesign of both the tools and procedures that we used. General participant comments are presented first below, and then each tool is discussed. The tools are then briefly compared with each other.

A.3.2.1 General Participant Comments

A common observation was that the participants felt they were not able to do a thorough evaluation in the four hours they were given. Some wanted more information from the manufacturer briefing, some wanted more time for free exploration, and some wanted more time to complete the assessment tools considering every item carefully. This sense of being rushed was especially pronounced amongst the four participants who had human factors expertise, who clearly felt they were leaving tasks undone.

Participants were also aware that their evaluations could be performed at different levels, from a general assessment to a detailed thorough review. Here it seems that our participants were not representative of our target audience in that they did not have any experience conducting these types of evaluations. Our participants were struggling to find a balance between the two types of review (general versus detailed), doing a task unlike most others they are used to doing. This sometimes led to confusion and/or frustration. The task was demanding and required full concentration for the duration of the evaluation.

Some participants had good suggestions about how the tools could be altered to promote faster completion, or good insights about why using a particular tool seemed to take a long time. One suggestion was to have more concrete test scenarios and more concrete response scales (e.g., a scale from agree to disagree for a given statement such as "I was able to accomplish X easily."). Another suggestion was to use a Cooper-Harper-like response scale, which would lead you to a rating by asking successively more specific questions. These suggestions should be considered as alternative formats for the next version of our assessment tools.

A.3.2.2 EFB Checklist Tool

The EFB Checklist was designed to elicit a comprehensive review of an EFB's attributes. Overall, four of the participants agreed that the Checklist added value to the process. However, one participant said that it was too detailed, and one was neutral. While all participants also stated that it was easy to answer the items, their responses to more detailed questions about the Checklist tool pointed out multiple weaknesses in its design. For example, four participants felt they could not complete the checklist in a reasonable time; there were too many items on the list and each had to be read carefully. Also, four participants said they found the wording acceptable with exceptions but two participants said they found the wording confusing. Three participants found the response scheme confusing, while three found it acceptable. Participants did agree in general though that the format (e.g., topical organization, bolding, numbering) was helpful.

On the positive side, many of the items on the EFB Checklist are objective (i.e., based on direct observation) and, as a result, both human factors and non-human-factors experts were able to come to an agreement on their ratings. Also, the co-discovery technique we used, in which the participants discussed items and responses among themselves, generated comments that might not have emerged had each participant completed the Checklist independently. However, these comments were more often recorded by the observers rather than the participants.

Overall, the main drawback of the EFB Checklist was that it took too long to complete; none of the teams actually completed the entire Checklist within the one hour allotted. We observed several factors that contributed to the overall length of time it took to complete the Checklist. First, participants often appeared to spend a long time reading individual items because this was their first exposure to the items. Some of the wordings on individual items could have been improved, but overall, it seemed to be lack of familiarity with the item, not the specific wording, affected the time it took to understand the item. Second, a good deal of the time taken to complete the Checklist was actually used to discuss ratings for individual items or to follow a chain of associations spurred by the item. Digressions seemed to occur more frequently early on, and issues that were brought up in the later items were sometimes discussed before the items were reached. Again, familiarity with the item list could help to constrain the task, because the evaluators would know that their discussion pertained to a later item. A third factor that contributed to the length of the Checklist evaluation was that it took some time for participants to shift mindsets between items. One of the more problematic shifts that participants experienced was that they had to evaluate both applications and the overall system together. Sometimes, they did not know which level to focus upon, the application or the overall system. A potential fourth factor is that the items themselves were relatively brief sentences or sentence fragments. While brevity might help an evaluator who was already familiar with the items, it could have hindered an evaluator who was unfamiliar with the item. Longer items with more detailed text might have been easier for first-time evaluators to understand.

Another significant problem with the Checklist was that there was some confusion over the rating scale Improvements could be made to reduce or eliminate this confusion about the response codes. For example, participants pointed out in both the survey and the discussions that the "–" and "√" columns should be combined; there was no need for two separate response columns. Also, the "NE" column should have two response codes, either NA for "not applicable because the product will not support this feature/item", or NE for "not evaluated because this feature is not yet ready for evaluation." In addition, participants sometimes had significant debates over whether an item deserved a "Problem," "Concern," or "Optimization" rating. It is not clear whether this was a desirable effect or not. It is likely that experienced regulatory evaluators would have less trouble than our participants had with this scale.

Finally, the number and nature of comments recorded from the Checklist tool varied substantially for different participants. This may have been due to several factors. First, some participants may be better at "generating" comments than others. Some may have been pre-occupied in their own observations and in trying to understand the system and the Checklist. Second, there is no obvious area for writing comments on the paper version, and it was sometimes difficult to match comments to the appropriate topic statement after the fact. In general, the most useful comments came from the experimenters who took notes as the participants discussed the Checklist items. This suggests that to get the full benefit of using the Checklist might require a formal, or perhaps even dedicated, note taker. This would allow other participants to focus on the Checklist topics without the additional burden of keeping their own notes.

To summarize, we have several recommendations on improving the format of the Checklist to improve its utility. Specifically:

- Simplify and clarify the response/coding scheme.
- Clarify some items and perhaps expand some items to ensure that they can be understood quickly by both experienced and inexperienced evaluators.
- Group the Checklist items better and even eliminate some items, if possible, to shorten the list.

- Provide a more obvious area on the Checklist form for notes, and possibly for examples and exceptions.

In addition, there are process improvements that could be made in use of the EFB Checklist. Specifically:

- Provide a copy of the Checklist to evaluators in advance so that they can be familiar with its structure and individual items before the evaluation.
- Designate someone as the note-taker during the Checklist completion, either one of the participants, or a separate observer. The note-taker could also be the person designated to synthesize all the individual observations noted during the evaluation.

A.3.2.3 Quick Assessment Tool

All participants stated that the QAT added value to the evaluation process. In addition, the QAT was faster to complete than the Checklist; all participants felt that it could be completed in a reasonable time. There are several explanation for the speed of completing the QAT. One, of course, is that there were far fewer items on the QAT (18) than the Checklist (over 100), and each item had a closed set of responses. Another likely factor is that this tool was completed separately by each participant; ratings were not negotiated, and discussions were not instigated. However, because the QAT was completed individually, it was generally done in silence, so there was no opportunity for an observer to capture any issues.

The most problematic aspect of the QAT was its terminology. Four participants rated the wording as "acceptable with exceptions," and two felt it was "acceptable." Two terms were cited as particularly confusing, "task match" and "destructive adjacencies. " Although the written definitions helped, verbal explanations were often necessary.

There are at least two reasons why the QAT terminology may have been unclear. The first is that the terms were specific to the mindset of the tool's designer alone. Even some of the other usability experts on the test team found the language unclear, so it is not surprising that the participants had some trouble with it as well. The second reason that the terminology may have been unclear is, as mentioned earlier, that the different topics did overlap somewhat. As a result of this overlap, some participants noted that it was sometimes difficult to decide under which topic area to note a particular comment. In fact, the elicitation of the comments was more important than where they were placed, but the participants did expend effort on placing comments as well as they could. Familiarity with the tool could help reduce these problems, but significant training time might be required.

Aside from the terminology, some participants had difficulty using the electronic QAT, primarily in selecting the hyperlinks. Also, one evaluator pointed out that the items in the QAT could be applied at both the system and application level (e.g., navigation could be rated relative to movement *between* applications and movement *within* applications). This is an important distinction that should be clarified in our next evaluation.

When viewed as a whole, we found that consensus among participant ratings for the QAT varied from topic to topic. (Recall that the QAT ratings were produced individually, unlike the Checklist ratings.) This could also be a reflection of the fact that the topic definition or the rating definitions were not clear, or it could reflect the fact than one individual caught a problem that changed his/her rating while another individual did not catch that same problem. For example, in our test, all raters gave "Functional and Navigational Logic" the same rating. However, the ratings for "Error penalties" were spread widely, covering the spectrum of responses. If the QAT were completed as a team, some of these ratings disparities could be resolved, but that may not solve the problem about clarity of the term.

The QAT did generate many comments from participants, particularly when it was used before the Checklist Tool. In our test, the bulk of comments were obtained on the areas that were identified as problematic from the expert reviews as well, so there was good correlation between the two methods of assessment in that regard. However, there were several areas where no comments were given, and it is not clear whether that reflected a lack of time, or actual lack of comments.

Another problem that the observers noted is that participants sometimes wanted to say that the interface deserved one rating "in general," but they had seen just one or a few specific examples of problem areas. One suggestion for capturing this type of observation is to include an "examples' or "exceptions" area within, the comments section. Breaking out comments in this way may also help evaluators to think about examples and exceptions more clearly.

To summarize, our recommendations for improving the QAT are:

- Translate the QAT terms and rating options into simpler language for both human factors experts and nonexperts. Try to clarify the overlap between the topic areas, if any. Also, the rating scale should clearly identify "best" to "worst" ratings at each end of the scale.
- Structure the participant comments by separating out three types of comments: General, Examples, and Exceptions.

A.3.2.4 Comparison of Assessment Tools

The two assessment tools highlighted usability issues from quite different perspectives. More specifically, as some participants noted in the questionnaires, the QAT was useful for addressing high level user interface issues, while the Checklist was useful for addressing specific detailed user interface issues. Not surprisingly, however, the length of time required for the evaluation appeared to be directly correlated with the depth of the review, with the in-depth review demanding a lengthy evaluation time. Ideally, either a more flexible tool, which could be customized for the time available and depth required for the evaluation, should be developed, or a tool should be designed so that it addresses issues at a medium level of detail.

Note, however, that all the teams in this study completed both the Checklist and the QAT, so we were not able to assess the utility of the tools individually, only in combination. In other words, we cannot say definitively whether one tool proved to be more or less useful than the other. Instead, our sense is that experiences gained through completing the first tool were considered when completing the second tool, raising the general quality of the evaluation. Also, anecdotally, our impression was that individual differences in factors such as personality (e.g., big-picture thinker versus detail-oriented person), expertise (human factors or not), or note-taking preferences (paper versus electronic) may have affected the utility of the tools for different participants. These individual differences could be explored in future studies, if desired.

Appendix B:
Tools from November 2003 FAA meeting

Three options for the high level tool, varying in detail, and a redesign of the detail tool are presented. Option 1 for the high level tool (Section B.1) provides only an overview of the human factors issues to look for. This tool is intended to be used by evaluators who are familiar with the content of the tool, and only need reminders of topics to consider when conducting an evaluation. Columns are provided for each application so evaluators can provide ratings on how the human factors issues are addressed within an application.

Options 2 and 3 in Section B.2 and B.3 show a more detailed version of the high level tool. These versions provides evaluators with details for each topic regarding what to look for during an evaluation for each of the topic headings.

The detailed tool presented in Sections B.4 maps the guidance from Appendix B of the EFB Version 2 document to general topics in the high level tool. Appendix B is an 11-page summary of roughly 100 pages of equipment requirements and recommendations. Each item is a paraphrased version of guidance from the main document.

B.1. Option 1 for High Level Tool

HIGH LEVEL TOOL (Level 1)	General EFB System	Electronic Documents	Electronic Checklist Systems	Flight Performance Calculations	Electronic Charts
Workload					
Hardware Considerations					
Software Considerations General Principles					
Platform: Consistency and Compatibility					
Information Structure/Formatting/Layout					
Symbols					
Interactions: Accessing functions and options					
Error handling and prevention					
Automation					
Other					

B.2. Option 2 for High Level Tool

HIGH LEVEL TOOL with Detail (Level 2)	General EFB System	Electronic Documents	Electronic Checklist Systems	Flight Performance Calculations	Electronic Charts
Workload					
Hardware Considerations • Physical Ease of Use • Labels and Controls • Language, terms, and abbreviations					
Software Considerations General Principles • Visual, audio, and tactile characteristics • Use of color • Feedback (system state, alerts, modes, etc) • Language, terms, and abbreviations					
Information Structure/Formatting/Layout					
Symbols and Icons					
Interaction: Accessing functions and options • Movement between pages • Number of inputs to complete a task • Ease of accessing functions and options • Feedback (system state, alerts, modes, etc) • Responsiveness • Intuitive logic					
Error handling and prevention • Susceptibility to error (mode errors, selection errors, data entry errors, reading errors, etc.) • Error recovery					
Consistency and Compatibility Across Applications					
Automation • Disruptive/Supportive • Predictable • Can override or reverse					
Other					

B.3. Option 3 for High Level Tool

	General EFB System	Rating
Workload	2.1.1 Workload	
Hardware	2.1.5 Legibility—Lighting Issues 2.2.4 Kneeboard EFBs 2.4.1 User Interface—General Design 2.5.1 Pointing and Cursor Control Devices 2.5.2 Hardware Controls 2.5.3 Display 2.5.5 Keyboards	
SOFTWARE General Principles	 2.4.1 User Interface—General Design 2.4.3 General Use of Colors 2.4.8 Alerts and Reminders 2.4.14 Supplemental Audio	
Platform: Consistency and Compatibility	2.4.2 Application Compatibility and Style Guides	
Information Structure/Formatting/Layout	2.4.10 Legibility of Text—Characters 2.4.11 Legibility of Text—Typeface Size and Width 2.4.12 Legibility of Text—Spacing for Readability	
Symbols	2.4.4 Graphical Icons 2.4.13 Non-Text Display Elements	
Interactions: Accessing Functions and Options	2.4.5 Multi-Tasking 2.4.6 Responsiveness 2.4.7 Anchor Locations 2.4.18 Links to Related Material	
Error Handling and Prevention	2.1.7 Failure Modes 2.4.9 Display of System Status 2.4.15 Ensuring Integrity of EFB Data 2.4.17 Crew Confirmation of EFB Software/Database Approval 2.4.19 User-Interface Customization	
Automation		
Other		

	Electronic Documents	Rating
Workload		
Hardware Considerations		
SOFTWARE General Principles	3.5.1 Printing 3.5.2 Animation	
Platform Consistency and Compatibility		
Information Structure/Formatting/Layout	3.2.1 Consistency of Information Structure 3.3.1 Visual Layout and Structure 3.3.2 Minimum Display Area and Resolution 3.3.3 Off-Screen Text 3.3.4 Active Regions 3.3.6 Figures	
Symbols		
Interactions: Accessing Functions and Options	3.4.1 Moving to Specific Locations 3.4.2 Managing Multiple Open Documents 3.4.3 Searching	
Error Handling and Prevention		
Automation		
Other		

	Electronic Checklists	Rating
Workload		
Hardware Considerations		
SOFTWARE General Principles	4.2.1 Checklists Supported by the ECL System 4.5.3 Task Reminders	
Platform Consistency and Compatibility		
Information Structure/Formatting/Layout	4.2.2 Information and Visual Layout/Structure of Electronic Checklists 4.3.2 Managing Checklists 4.3.3 Managing Non-Normal Checklists 4.3.6 Closing All Checklists 4.4.2 Displaying Item Status 4.4.4 Specifying Completion of Item 4.5.4 Checklist Branching	
Symbols		
Interactions: Accessing Functions and Options	4.3.1 Accessing Checklists 4.3.2 Managing Checklists 4.3.4 Lengthy Checklists 4.3.5 Closing or Completing a Checklist 4.4.2 Displaying Item Status 4.4.3 Moving Between Items Within a Checklist 4.4.4 Specifying Completion of Item 4.5.1 Links Between Checklist Items and Related Information 4.5.2 Links to Calculated Values 4.5.4 Checklist Branching	
Error Handling and Prevention		
Automation		
Other		

	Flight Performance Calculations	Rating
Workload		
Hardware Considerations		
SOFTWARE General Principles	5.1.1 Default Values 5.1.3 Support Information for Performance Data Entry	
Platform: Consistency and Compatibility		
Information Structure/Formatting/Layout		
Symbols		
Interactions: Accessing Functions and Options	5.1.5 Modifying Performance Calculations	
Error Handling and Prevention	5.1.2 Data-entry Screening and Error Messages	
Automation		
Other		

	Electronic Charts	Rating
Workload		
Hardware		
SOFTWARE General Principles	6.2.1 Transition from Paper to Electronic Charts 6.2.3 Hard Copy Backups of Electronic Charts	
Platform Consistency and Compatibility		
Information Structure/Formatting/Layout	6.2.7 Orientation of Electronic Charts	
Symbols		
Interactions: Accessing functions and options	6.2.5 Basic Zooming and Panning 6.2.9 Access to Individual Charts 6.2.11 De-cluttering and Display Configuration	
Error Handling and Prevention	6.2.2 Updates to Electronic Charts 6.2.4 Scale Information 6.2.10 Knowledge and Display of Own-Aircraft Position	
Automation		
Other		

B.4. Redesigned Detailed Tool

2 General EFB System

Workload

2.1.1 Workload

- ❑ Flight crew workload and head-down time should be minimized (AC 120-76A, Section 10.c)

Hardware

2.1.5 Legibility—Lighting Issues

- ❖ Automatic brightness adjustment should be independent for each EFB (See AC 25-11)
- ❑ Screen brightness should adjustable in fine increments or continuously
- ❑ Buttons and labels should be adequately illuminated for night use

2.2.4 Kneeboard EFBs

- ❖ Kneeboard EFB should be easily removable

2.4.1 User Interface—General Design

- ❑ User interface should have a consistent set of controls and graphical elements (see also General Principles)
- ❑ Controls used for different functions should be visually distinct
- ❑ Graphic elements and controls should follow personal computer conventions, except where clearly inappropriate for flight deck environment (see also General Principles)

2.5.1 Pointing and Cursor Control Devices

- ❑ Input devices should be selected and customized based on the type and complexity of the entries to be made and flight deck environmental factors that affect its usability
- ❑ Performance parameters should be tailored for the intended application and for the flight deck environment
- ❑ Users should be able to rest and/or stabilize their hand when using the pointer or cursor control device
- ❑ Active areas should be sized to permit accurate selection with the pointer/cursor device under all operating conditions

2.5.2 Hardware Controls

- ❖ All controls should be properly labeled (14 CFR 23.1555, 25.1555, and 27.1555)
- ❖ All soft function keys should be labeled
- ❖ Inactive soft function keys should not be labeled or should use a visual convention to indicate that the function is not available
- ❑ Physical function keys should provide tactile feedback when pushed
- ❑ Key repeats should be filtered by the software if they occur too closely together
- ❑ Soft function keys should be drawn in a reserved space outside the main content area
- ❑ The same function should appear on the same function key, whenever possible
- ❑ Labels should be consistent
- ❑ Labels should be clear and brief
- ❑ Labels may use standard abbreviations; ambiguous abbreviations should be avoided
- ❑ Labels should be located near the controls they identify and should not be confusingly close to other labels or other controls
- ❑ Labels should be drawn in horizontal text
- ❑ Physical controls should be collocated with the display

- ❑ The most frequently used controls should be placed at the most accessible locations
- ❑ Controls presented in a small space may need to be grouped according to function and/or order of use
- ❑ Controls should be designed to deter inadvertent activation

2.5.3 Display

- ❑ The physical nature of the display screen should minimize the likelihood that information will be obscured

2.5.5 Keyboards

- ❑ Keyboard type should be appropriate for the given task
  - QWERTY type keyboards should be used for text entry
  - Numeric keypads are best suited for significant numeric entries
- ❑ Keyboards should provide appropriate tactile feedback
- ❑ Users should be able to rest/stabilize their hand to use the keyboard, especially during turbulence

SOFTWARE

General Principles

2.4.1 User Interface—General Design

- ❑ User interface should have a consistent set of controls and graphical elements (see also Hardware)
- ❑ Graphic elements and controls should follow personal computer conventions, except where clearly inappropriate for flight deck environment (see also Hardware)
- ❑ Menu functions should be accessible in proportion to frequency of use and criticality to mission

2.4.3 General Use of Colors

- ❖ Red and amber should be reserved for highlighting *warning* and *caution* level conditions respectively (AC 120-76A, 10.d(1))
- ❖ Color should not be sole means of coding important differences in information; color should be used redundantly
- ❖ Color-coding scheme should be interpretable easily and accurately.
- ❑ Each color should be associated with only one meaning
- ❑ No more than six colors with assigned meanings should be used in a color-coding scheme
- ❑ EFB colors should not conflict with flight deck conventions
- ❑ For Part 121 and 135, default colors that represent different types of data should be customizable only by an appropriately authorized administrator
- ❑ If colors are customizable, there should be an easy way to return to default settings

2.4.8 Alerts and Reminders

- ❖ Alerts and reminders should meet 14 CFR Part 23.1322, 25.1322, 27.1322 or 29.1322 as appropriate. Their intent should be generalized to the use of colors on displays and controls (AC 120-76A, 10.d (1))
- ❖ Red should be used only for warnings (AC 120-76A, 10.d (1))
- ❖ Amber should be used only for cautions (AC 120-76A, 10.d (1))
- ❖ Other colors should be sufficiently distinct from red/amber for use (AC 120-76A, 10.d (1))
- ❑ Alerts and reminders should be consistent with AC 25-11, 14 CFR Part 23.1311a, AMJ 25-11
- ❑ Alerts should be integrated or compatible with other flight deck alerts (AC 120-76A, 10.d (1))
- ❑ Messages should be prioritized and prioritization scheme should be documented and evaluated (AC 120-76A, 10.d (1) and AC 120-76A, 10.d (2))
- ❑ Strong attention-getting techniques (e.g., flashing or bright text) should be avoided (AC 120-76A, 10.d (1))
- ❑ During high workload phases of flight:

(a) Required flight information should be continuously present and unobscured, except those that indicate failure or degradation of the EFB application (AC 120-76A, 10.d (1))
(b) Messages should be inhibited, except those that indicate failure or degradation of the EFB application (AC 120-76A, 10.d (1))

2.4.14 Supplemental Audio

- ❑ Supplemental audio should be avoided in flight
- ❑ Users should be able to control the volume
- ❑ Users should be able to turn off the supplemental audio
- ❑ Objects with supplemental audio should be coded so the user knows of the associated audio before activating it
- ❑ Supplemental audio that is solely audio should have text description available
- ❑ Users should be able to stop the supplemental audio at any time

Platform: Consistency and Compatibility

2.4.2 Application Compatibility and Style Guides

- ❑ All applications should follow a common style guide, preferably specific to that aircraft
- ❑ Color and other formatting should be internally consistent across applications (AC 120-76A, Section 10.b (1))
- ❑ Help facility, if available, should be standardized across applications
- ❑ Soft key labels and menus should be consistent across applications
- ❑ Common actions allowed on multiple applications should be performed in the same manner (see also Interactions: Accessing Functions and Options)
- ❑ Manufacturers should prepare style guides for third party developers

Information Structure/Formatting/Layout

2.4.10 Legibility of Text—Characters

- ❑ Typeface should be highly legible. HFDS recommends:
  - Spare use of upper case text (8.2.5.8.2)
  - Mixed upper and lower case for continuous text (8.2.5.8.4)
  - Serif fonts for high resolution displays (8.2.5.7.5)
  - Sans serif fonts otherwise (8.2.5.7.6)
  - Character contrast between 6:1 and 10:1 (8.2.5.6.12)
  - Characters stroke width 10 to 12% of character height (8.2.5.6.14)
- ❑ Individual characters should not be easily confused with other characters
- ❑ Slanting or italic text should be avoided

2.4.11 Legibility of Text—Typeface Size and Width

- ❖ Typeface should be appropriate for viewing distance, lighting conditions, and text criticality
- ❑ The FAA HFDS recommends that:
  (a) Minimum character height should be 1/200 of viewing distance, e.g., for 35” viewing distance, 0.175” tall (17.5 pixels at 100 pix/inch) (8.2.5.6.6)
  (b) Preferred character height should be 1/167 of viewing distance (8.2.5.6.5)
  (c) Character height to width ratios should be (8.2.5.6.10)
    - o <80 char per line, 1 to 0.7 up to 0.9 (15 pix tall, 10.5 to 13.5 pix wide) for monotype fonts
    - o >80 char per line, at least 1 to 0.5 (15 pix tall, 7.5 pix wide)
    - o 1:1 for M and W in a proportional font
- ❑ Larger fonts should be used for text read in poor viewing conditions

2.4.12 Legibility of Text—Spacing for Readability

- ❑ Text should be spaced appropriately to facilitate reading

- ❑ Line lengths should be appropriate for text content
- ❑ To facilitate readability, HFDS recommends the following:
  - (a) Use horizontal spacing between characters that is at least 10% of character height (15 pix tall, 1.5 pix spacing) (8.2.5.6.1)
  - (b) Use spacing between words of at least one character for equally spaced characters, or width of "N" for proportional fonts (8.2.5.6.2)
  - (c) Use spacing between lines of at least two stroke widths or 0.15 of character height (15 pix tall, 2.25 pix leading), whichever is greater (8.2.5.6.3)
  - (d) Separate paragraphs with blank line (8.2.5.6.4)

Symbols

2.4.4 Graphical Icons

- ❑ Icons should be accompanied with text labels
- ❑ Design of icons should minimize training and maximize intuitiveness for cross-cultural use

2.4.13 Non-Text Display Elements

- ❑ Non-text display elements should be distinguishable based on shape alone, without relying on secondary cues such as color or labels
- ❑ Non-text display elements should be designed for legibility on minimum expected display resolution viewed from the maximal intended viewing distance

Interactions: Accessing Functions and Options

2.4.5 Multi-Tasking

- ❑ The user should be able to identify the active application easily
- ❑ The user should be able to:
  - Select which of the open applications is currently active
  - Switch between applications easily
- ❑ Applications, running in the background, should be in the same state when the user returns to it, other than the completion of any background processing
- ❑ Responsiveness of an individual application should not suffer when all applications are running simultaneously
- ❑ The user should be able to exit applications with pending activities by completing them or by acknowledging that they are incomplete
- ❑ The system should discourage use of non-flight-related applications and ask for an extra confirmation to launch

2.4.6 Responsiveness

- ❖ The system should provide feedback when a user input is processed
  - Alphanumeric inputs should be shown within 0.2 seconds (SAE ARP 4791)
- ❖ A "system busy" indicator should be displayed if user inputs can not be processed within 0.5 seconds (SAE ARP 4791)
- ❑ The EFB applications should have a "system busy" indicator
- ❑ The type of feedback should be appropriate for the type of user input
- ❑ If tasks take more than a few seconds to complete, indicators should show their progress
- ❑ User entries made while the system is busy should be stored for later processing

2.4.7 Anchor Locations

- ❑ If the EFB supports more than one application, there should be an anchor location from which the user moves between applications
- ❑ Each EFB application should have its own anchor page
- ❑ It should be easy to move from any location in the EFB to an anchor location, and vice versa

2.4.18 Links to Related Material

- ❑ A consistent philosophy should be used for accessing different types of information. Similar types of information should be accessed in the same way
- ❑ Users should be able to keep track of how to move between topics. Users should be able to return to the starting point easily

Error Handling and Prevention

2.1.7 Failure Modes

- ❑ EFB should alert the flight crew to probable application/system failures (AC 120-76A, Section 10.e (2))

2.4.9 Display of System Status

- ❑ Any full or partial application failure should be indicated with a positive indicator (AC 120-76A, Section 10.d (2))
- ❑ The immediacy of indicator should be appropriate to the function that is lost or disabled (AC 120-76A, Section 10.d (2))

2.4.15 Ensuring Integrity of EFB Data

- ❖ EFB data should be checked prior to installation to ensure that they are accurate, current, and uncorrupted
- ❑ The EFB should check that the current date is within the valid date range
- ❑ The EFB should allow data with an effective date in the future to be installed
- ❑ The system should conduct a self-test to ensure that the data is current and generate a message to the flight crew if any data is out of date. The message should indicate where to go for further information.

2.4.17 Crew Confirmation of EFB Software/Database Approval

- ❖ The latest revision information should be available upon request

2.4.19 User-Interface Customization

- ❖ There should be an easy means to return all settings to their default values
- ❑ For Part 121 and 135, the default settings should be customizable only by an administrator
- ❑ For Part 91, the default settings should be specified by the manufacturer and configurable by the user

3 Electronic Documents

General Principles

3.5.1 Printing

- ❖ Pages or sections selected for printing should be clearly indicated
- ❖ The user should be able to terminate printing immediately
- ❑ Users should be able to select document subsets for printing
- ❑ The printed document should have the same visual structure as the EFB electronic document

3.5.2 Animation

- ❖ Start/stop functionality should be provided. The user should be able to stop the animation at any time
- ❖ Text describing the animation should be available even if the animation is not running
- ❑ Animation should not be overused
- ❑ If supplemental audio is provided, control of the audio and video should be integrated

Information Structure/Formatting/Layout

3.2.1 Consistency of Information Structure

- ❖ The information structure of the electronic document should be consistent with that of the hard copy

3.3.1 Visual Layout and Structure

- ❑ Windows and frames should be placed and used consistently
- ❑ Sections of text should be separated with plenty of white space
- ❑ Data should be formatted into short segments, where possible

3.3.2 Minimum Display Area and Resolution

- ❑ The minimum document display area and resolution should be specified by the manufacturer
- ❑ Operators should meet the manufacturer-specified display area and resolution requirements for training and operational use

3.3.3 Off-Screen Text

- ❖ The existence of off-screen content should be indicated clearly and consistently (AC 120-76A, 10.b (7))
- ❑ Whether it is acceptable for parts of the document to be off-screen should be based on the application and intended function (AC 120-76A, 10.b (7))
- ❑ Information regarding the document length and the current place within the document should be constantly available

3.3.4 Active Regions

- ❖ Active regions should be clearly indicated (AC 120-76A, 10.b (8))

3.3.6 Figures

- ❖ The electronic version of a figure should show all the content in the paper version
- ❖ The entire figure should be viewable at once, even if all the details are not readable
- ❖ All the details should be readable, although the entire figure may not be visible when doing so
- ❑ Figures should be displayed in their entirety with all details readable whenever possible
- ❑ Text information should be provided for each figure, independent of whether the figure is shown in full, or marked by a placeholder
- ❑ The user should be able to configure the figure for optimal viewing

- ❑ If zooming is supported, discrete zoom levels should be available (e.g. view whole page) and the current zoom level should be displayed at all times

Interactions: Accessing Functions and Options

3.4.1 Moving to Specific Locations

- ❑ The cursor should be visible at all times (AC 120-76A, 10.b (7))
- ❑ If links are supported:
  - Entries in the table of contents should be linked to its location in the text
  - Cross-references should be linked to each other within a document
- ❑ Users should be able to return to the previous location in one step

3.4.2 Managing Multiple Open Documents

- ❖ The active document should be indicated continuously (AC 120-76A, 10.b (9))
- ❖ The user should be able to choose the active open document
- ❑ A master list of all open documents should be available

3.4.3 Searching

- ❑ Search functionality should be available
- ❑ Users should be able to select the document(s) to include in the search

4 Electronic Checklist Systems

General Principles

4.2.1 Checklists Supported by the ECL System

- ❖ If normal checklists are supported, then *all* normal checklists should be supported
- ❖ If non-normal emergency checklists are supported, then *all* non-normal checklists should be supported
- ❖ Similar requirements apply for other checklist categories
- ❑ The ECL system should indicate the location of unsupported checklists in the paper document
- ❑ Non-normal checklists should retain as much commonality with normal checklists as possible

4.5.3 Task Reminders

- ❑ Reminders for high priority, time-critical tasks should be displayed constantly once in progress and should attract attention when delayed actions should be performed
- ❑ If multiple task reminders can be shown, crews should be able to determine how many are in progress and to what tasks they refer

Information Structure/Formatting/Layout

4.2.2 Information and Visual Layout/Structure of Electronic Checklists

- ❖ The resulting crew actions called for in the checklist should be identical for paper and electronic versions
- ❑ Layout of items should be similar to the paper version. Headings, sub-headings, and titles should be consistent (CAP 807)
- ❑ The format of the electronic checklist should make it clear which challenge is associated with which response (CAP 708)

4.3.2 Managing Checklists

- ❖ The checklist title should be displayed above the items and be distinguished throughout the checklist
- ❑ Parent-child checklists should be integrated into a single checklist
- ❑ If more than one checklist can be open at once, a master list of checklists should be available

4.3.3 Managing Non-Normal Checklists

- ❑ All checklists associated with on-going non-normal conditions that are sensed should be listed on one master list
- ❑ A master list should indicate the status of each checklist

4.3.6 Closing All Checklists

- ❑ The ECL should allow a state where no checklists are open
- ❑ The system should give a positive indication that no checklists are open; a blank screen is not sufficient

4.4.2 Displaying Item Status

- ❖ Item status, if available, should be clearly indicated.

4.4.4 Specifying Completion of Item

- ❑ The completion status of each checklist should be indicated clearly

(see also Interactions: Accessing Functions and Options)

4.5.4 Checklist Branching

- ❑ The selected branch should be clearly indicated

(see also Interactions: Accessing Functions and Options)

Interactions: Accessing functions and options

4.3.1 Accessing Checklists

- ❖ All supported checklists should be accessible for reference/review at any time while the system is active
- ❑ Normal checklists should be accessible in accordance with the normal sequence of use
- ❑ Electronic checklists should be as quick and accurate to access as paper checklists
- ❑ The ECL system should open checklists only upon crew request

4.3.2 Managing Checklists

- ❖ The title of each open checklist should be visible continuously
- ❖ If more than one checklist can be open at once, other checklists should be accessible without closing the displayed checklist
- ❖ If more than one checklist can be open, the user should be able to select which one is active
- ❖ If a checklist is a "child" of another checklist, the user should be able to select whether the parent or child is active
- ❑ A placeholder should be used to indicate which item was active prior to leaving the checklist
- ❑ The crew should be able to reset the checklist with a simple input

4.3.4 Lengthy Checklists

- ❖ The user should be able to look ahead (e.g., page down) without changing the active item
- ❑ Information regarding the length of the checklist, the user's current position within the checklist, and how much of the checklist has been completed should be continuously available
- ❑ It should not be possible to change the status of off-screen items
- ❑ If the active item is off-screen and the user makes an "item completed" entry, an error message should appear or the active item should be called into view

4.3.5 Closing or Completing a Checklist

- ❖ If item status is tracked and the user attempts to close an incomplete checklist, the system should provide an indication that the checklist is incomplete and present any deferred/incomplete items for review
- ❖ The user should be able to close incomplete checklists after acknowledging this indication
- ❑ If item status is tracked, a positive indication should be presented when the entire checklist, as well as each item, is completed
- ❑ The action for closing/completing a checklist should be distinct from the action for marking an item as complete

4.4.1 Indicating the Active Item

- ❑ The ECL should track and indicate the active checklist item
- ❑ When returning to an incomplete checklist, the item active prior to the move should again be active

4.4.3 Moving Between Items Within a Checklist

- ❖ The active-item pointer should be moved to the next item with a simple action
- ❖ Returning to a previous item should not change the status of any item
- ❑ If the status of individual items are tracked, the user should be able to:
  - (a) Move from uncompleted items, changing their status to deferred
  - (b) Move to the next item automatically after completing an item

- The user should be able to quickly select one item after another; system processing should not induce delays

4.4.4 Specifying Completion of Item

- User actions to mark an item as complete should be simple
- Completed items should not be removed from the screen immediately. The crew should be able to review the item and undo their action, if necessary
- If the system indicates active items:
  a) The next item in the list should become active when an item has been completed, unless it is on the next page. A separate action should be required to move to the next page
  b) Moving to the next item without completing the current item should require an input distinct from that of specifying the item as complete
- An *undo* function should be available
- The completion status of each checklist should be indicated clearly

(see also Information Structure/Formatting/Layout)

4.5.1 Links Between Checklist Items and Related Information

- The navigation between links in the ECL and related information needs to be simple and clear
- Related information should appear in a single window or area of the screen. Hyperlinks from the related information should be shown in the same window or area

4.5.2 Links to Calculated Values

- If the EFB provides calculation worksheets and allows integration between the application hosting the ECL and the application hosting the calculation worksheets, then:
  (a) Direct access to the appropriate worksheet should be provided for all items that can be calculated. This should be available for initial calculations and subsequent review/modifications
  (b) The user should be able to return easily to the checklist item from which the worksheet was accessed
- Calculated ECL values should appear in the corresponding checklist location. These fields should be blank prior to inserting the calculated value

4.5.4 Checklist Branching

- The user should be able to backup and select another decision branch
- Items not on the selected branch should not be selectable
- The selected branch should be clearly indicated

(see also Information Structure/Formatting/Layout)

5 Flight Performance Calculations

General Principles

5.1.1 Default Values

- ❑ Blank data entry fields should be used to indicate that there is no system assigned default value

5.1.3 Support Information for Performance Data Entry

- ❖ The units of each variable should be clearly labeled
- ❑ Labels, formats, and units of variables should match that in other sources (e.g., paper reports, flight deck systems)
- ❑ Related information for cross-checking should be in view or easily accessible

Interactions: Accessing Functions and Options

5.1.5 Modifying Performance Calculations

- ❑ The user should be able to modify previously computed results quickly
- ❑ Output relevant to earlier calculations should be erased once the user begins modifying those calculations

Error Handling and Prevention

5.1.2 Data-entry Screening and Error Messages

- ❑ The EFB should not accept user-entered data that is of incorrect format or type. Error messages should point out suspect entries and specify the expected data type. (AC 120-76A, Section 10.d (3))
- ❑ The system should detect input errors as early as possible during data entry (AC 120-76A, Section 10.d (3))
- ❑ The system should *only* discard erroneous input errors and not the whole set of entries related to the task in progress
- ❑ The system should present an error message when required values are missing; this error message should contain the name of the required value, using the label from the input field

6 Electronic Charts

General Principles

6.2.1 Transition from Paper to Electronic Charts

- ❑ Information structure of electronic charts should match that of paper charts
- ❑ Visual structure of electronic charts should be compatible with paper charts

6.2.3 Hard Copy Backups of Electronic Charts

- ❑ If the hard copy is used as a backup, it should be of sufficient quality to be used as effectively as the original paper chart. In particular:
  (a) The hard copy should be legible; all chart details should be visible
  (b) The quality of the paper should be acceptable for normal use
  (c) Color information should be distinguishable in the monochrome hard copy
  (d) All the chart information should fit on one printed page
  (e) The hard copy should be at least as large as a standard paper chart
  (f) The user should be able to select the size of the hard copy

Information Structure/Formatting/Layout

6.2.7 Orientation of Electronic Charts

- ❖ Orientation of the charts should be indicated continuously
- ❖ When charts are oriented with respect to directionality (e.g., track/heading), and directionality information becomes unusable, it should be clear to the pilot that that information is not available
- ❑ When charts are oriented with respect to directionality (e.g., track/heading), and directionality information becomes unusable,
  (a) The crew should be notified of the unusable directionality and informed that the charts must revert to north-up orientation.
  (b) After crew acknowledgement of the failure, the charts should revert to the north-up orientation, the chart orientation indicator should be updated, and any cues that could imply directionality should be removed
- ❑ Text and symbols other than those designed to reflect compass orientation should remain upright at all times
- ❑ Crew input should be required to change the orientation of the charts

Interactions: Accessing functions and options

6.2.5 Basic Zooming and Panning

- ❑ If zooming is supported, then panning should also be supported, and vice versa
- ❑ The chart's visual edges should be clearly marked. Visual edges should be shown only when no more information is outside that area
- ❑ When panning, the user should know which way to move to bring more of the chart into view
- ❑ Panning to an area where no portion of the chart will be displayed should be prevented
- ❑ If the user can change zoom levels, the user should be able to return to a default view easily
- ❑ If the display can be panned, the user should be able to return to a default view easily
- ❑ Zooming and panning should not result in lengthy processing delays

6.2.9 Access to Individual Charts

- ❖ The currently selected chart's label should be displayed continuously
- ❑ The system should allow rapid access to pre-selected charts
- ❑ The chart application should help the crew ensure that the correct chart was selected and allow corrections to be made quickly when an error occurs
- ❑ Multiple search methods should be supported

- ❑ Search results should be ordered with its best guesses at the top of the list and least likely to be used charts at the bottom
- ❑ Selection of alternate runways should be facilitated during approach

6.2.11 De-cluttering and Display Configuration

- ❑ The pilot should not be able to declutter safety critical display elements without knowing they are suppressed
- ❑ Changing map scale, orientation, and other options and settings should not induce significant levels of workload
- ❑ The information prioritization scheme should be documented

Error Handling and Prevention

6.2.2 Updates to Electronic Charts

- ❑ Corrections/updates should be made directly within the electronic chart application, unless they are temporary
- ❑ Corrections/updates that are of high priority or time-sensitive should not be made via paper notifications

6.2.4 Scale Information

- ❖ Scale information should always be visible for charts drawn to scale
- ❖ Scale information should be accurate. Scale information should be updated when the display is zoomed
- ❖ Static scale information should be removed unless it is always accurate
- ❖ Charts drawn “not to scale” should have a label indicating that fact continuously

6.2.10 Knowledge and Display of Own-Aircraft Position

- ❖ Display of ownship should not be supported on non-georeferenced or not-to-scale terminal charts
- ❖ See TSO C-165 and DO-257A for other applicable requirements
- ❑ The range of display zoom levels should be compatible with the position accuracy of the ownship symbol.
- ❑ An indication of ownship position should be provided if the chart is zoomed or panned such that ownship is not in the current view

Appendix C:
Tools from December 2003

A high level tool is shown in Section C.1, and a detailed tool is shown in Section C.2. The high level tool is a one-page summary of topics to consider during an evaluation. Evaluators go through the list commenting on each item. The detailed tool consists of guidance from Appendix B of the EFB Version 2 document to the topics provided in the high level Tool. Items in the detailed tool are presented in the order of the topics listed in the high level tool. Note that the content of the detailed tool shown here has not changed from the version in Appendix B. However, the items have been re-ordered to correspond better with the item order in the high level tool.

C.1 High Level Tool

Topic	General EFB System	Electronic Documents	Electronic Checklist Systems	Flight Performance Calculations	Electronic Charts
Hardware Considerations • Physical Ease of Use — Input devices and display, accessibility of controls • Labels and Controls — Clarity and consistency of language, terms, and abbreviations • Lighting Issues (day vs. night use) • Amount of feedback, potential for errors					
Software Considerations Formatting/Layout • Fonts (size, style, case, spacing) • Arrangement of information on the display — Consistent with user expectations and internal logic					
Symbols and Graphical Icons • Clarity of intended meaning • Legibility and distinctiveness					
Interaction: Accessing functions and options • Anchor locations and ease of movement between pages • Number of inputs to complete a task • Ease of accessing functions and options • Feedback (system state, alerts, modes, etc) • Responsiveness • Intuitive logic					
Error handling and prevention • Susceptibility to error (mode errors, selection errors, data entry errors, reading errors, etc.) • Error recovery					
Multiple Applications • Consistency and compatibility across applications • Identifying current position within system					
Automation (if any) • Is it disruptive/supportive? Predictable? User control over automation? (e.g., manual override)					
General • Consistency of controls/elements; are they distinctive where appropriate? • Visual, audio, and tactile characteristics • Use of color (esp. red and amber) and color-coding • Amount of feedback (system state, alerts, modes, etc) • Clarity and consistency of language, terms, and abbreviations • End-user customization					
Workload • Problem areas					
Other					

C.2 Detailed Tool

2 General EFB System

HARDWARE CONSIDERATIONS

2.1.5 Legibility—Lighting Issues

- ❖ Automatic brightness adjustment should be independent for each EFB (See AC 25-11)
- ❑ Screen brightness should adjustable in fine increments or continuously
- ❑ Buttons and labels should be adequately illuminated for night use

2.2.4 Kneeboard EFBs

- ❖ Kneeboard EFB should be easily removable

2.4.1 User Interface—General Design

- ❑ User interface should have a consistent set of controls and graphical elements (see also General Principles)
- ❑ Controls used for different functions should be visually distinct
- ❑ Graphic elements and controls should follow personal computer conventions, except where clearly inappropriate for flight deck environment (see also General Principles)

2.5.1 Pointing and Cursor Control Devices

- ❑ Input devices should be selected and customized based on the type and complexity of the entries to be made and flight deck environmental factors that affect its usability
- ❑ Performance parameters should be tailored for the intended application and for the flight deck environment
- ❑ Users should be able to rest and/or stabilize their hand when using the pointer or cursor control device
- ❑ Active areas should be sized to permit accurate selection with the pointer/cursor device under all operating conditions

2.5.2 Hardware Controls

- ❖ All controls should be properly labeled (14 CFR 23.1555, 25.1555, and 27.1555)
- ❖ All soft function keys should be labeled
- ❖ Inactive soft function keys should not be labeled or should use a visual convention to indicate that the function is not available
- ❑ Physical function keys should provide tactile feedback when pushed
- ❑ Key repeats should be filtered by the software if they occur too closely together
- ❑ Soft function keys should be drawn in a reserved space outside the main content area
- ❑ The same function should appear on the same function key, whenever possible
- ❑ Labels should be consistent
- ❑ Labels should be clear and brief
- ❑ Labels may use standard abbreviations; ambiguous abbreviations should be avoided
- ❑ Labels should be located near the controls they identify and should not be confusingly close to other labels or other controls
- ❑ Labels should be drawn in horizontal text
- ❑ Physical controls should be collocated with the display
- ❑ The most frequently used controls should be placed at the most accessible locations
- ❑ Controls presented in a small space may need to be grouped according to function and/or order of use
- ❑ Controls should be designed to deter inadvertent activation

2.5.3 Display

- The physical nature of the display screen should minimize the likelihood that information will be obscured

2.5.5 Keyboards

- Keyboard type should be appropriate for the given task
  - QWERTY type keyboards should be used for text entry
  - Numeric keypads are best suited for significant numeric entries
- Keyboards should provide appropriate tactile feedback
- Users should be able to rest/stabilize their hand to use the keyboard, especially during turbulence

SOFTWARE

Platform: Consistency and Compatibility

2.4.2 Application Compatibility and Style Guides

- All applications should follow a common style guide, preferably specific to that aircraft
- Color and other formatting should be internally consistent across applications (AC 120-76A, Section 10.b (1))
- Help facility, if available, should be standardized across applications
- Soft key labels and menus should be consistent across applications
- Common actions allowed on multiple applications should be performed in the same manner (see also Interactions: Accessing Functions and Options)
- Manufacturers should prepare style guides for third party developers

Formatting/Layout

2.4.10 Legibility of Text—Characters

- Typeface should be highly legible. HFDS recommends:
  - Spare use of upper case text (8.2.5.8.2)
  - Mixed upper and lower case for continuous text (8.2.5.8.4)
  - Serif fonts for high resolution displays (8.2.5.7.5)
  - Sans serif fonts otherwise (8.2.5.7.6)
  - Character contrast between 6:1 and 10:1 (8.2.5.6.12)
  - Characters stroke width 10 to 12% of character height (8.2.5.6.14)
- Individual characters should not be easily confused with other characters
- Slanting or italic text should be avoided

2.4.11 Legibility of Text—Typeface Size and Width

- ❖ Typeface should be appropriate for viewing distance, lighting conditions, and text criticality
- The FAA HFDS recommends that:
  - (a) Minimum character height should be 1/200 of viewing distance, e.g., for 35” viewing distance, 0.175” tall (17.5 pixels at 100 pix/inch) (8.2.5.6.6)
  - (b) Preferred character height should be 1/167 of viewing distance (8.2.5.6.5)
  - (c) Character height to width ratios should be (8.2.5.6.10)
    - o <80 char per line, 1 to 0.7 up to 0.9 (15 pix tall, 10.5 to 13.5 pix wide) for monotype fonts
    - o >80 char per line, at least 1 to 0.5 (15 pix tall, 7.5 pix wide)
    - o 1:1 for M and W in a proportional font
- Larger fonts should be used for text read in poor viewing conditions

2.4.12 Legibility of Text—Spacing for Readability

- Text should be spaced appropriately to facilitate reading

- ❑ Line lengths should be appropriate for text content
- ❑ To facilitate readability, HFDS recommends the following:
  - (a) Use horizontal spacing between characters that is at least 10% of character height (15 pix tall, 1.5 pix spacing) (8.2.5.6.1)
  - (b) Use spacing between words of at least one character for equally spaced characters, or width of "N" for proportional fonts (8.2.5.6.2)
  - (c) Use spacing between lines of at least two stroke widths or 0.15 of character height (15 pix tall, 2.25 pix leading), whichever is greater (8.2.5.6.3)
  - (d) Separate paragraphs with blank line (8.2.5.6.4)

Symbols

2.4.4 Graphical Icons

- ❑ Icons should be accompanied with text labels
- ❑ Design of icons should minimize training and maximize intuitiveness for cross-cultural use

2.4.13 Non-Text Display Elements

- ❑ Non-text display elements should be distinguishable based on shape alone, without relying on secondary cues such as color or labels
- ❑ Non-text display elements should be designed for legibility on minimum expected display resolution viewed from the maximal intended viewing distance

Interactions: Accessing Functions and Options

2.4.5 Multi-Tasking

- ❑ The user should be able to identify the active application easily
- ❑ The user should be able to:
  - Select which of the open applications is currently active
  - Switch between applications easily
- ❑ Applications, running in the background, should be in the same state when the user returns to it, other than the completion of any background processing
- ❑ Responsiveness of an individual application should not suffer when all applications are running simultaneously
- ❑ The user should be able to exit applications with pending activities by completing them or by acknowledging that they are incomplete
- ❑ The system should discourage use of non-flight-related applications and ask for an extra confirmation to launch

2.4.6 Responsiveness

- ❖ The system should provide feedback when a user input is processed
  - Alphanumeric inputs should be shown within 0.2 seconds (SAE ARP 4791)
- ❖ A "system busy" indicator should be displayed if user inputs can not be processed within 0.5 seconds (SAE ARP 4791)
- ❑ The EFB applications should have a "system busy" indicator
- ❑ The type of feedback should be appropriate for the type of user input
- ❑ If tasks take more than a few seconds to complete, indicators should show their progress
- ❑ User entries made while the system is busy should be stored for later processing

2.4.7 Anchor Locations

- ❑ If the EFB supports more than one application, there should be an anchor location from which the user moves between applications
- ❑ Each EFB application should have its own anchor page
- ❑ It should be easy to move from any location in the EFB to an anchor location, and vice versa

2.4.18 Links to Related Material

- ❑ A consistent philosophy should be used for accessing different types of information. Similar types of information should be accessed in the same way
- ❑ Users should be able to keep track of how to move between topics. Users should be able to return to the starting point easily

Error Handling and Prevention

2.1.7 Failure Modes

- ❑ EFB should alert the flight crew to probable application/system failures (AC 120-76A, Section 10.e (2))

2.4.8 Display of System Status

- ❑ Any full or partial application failure should be indicated with a positive indicator (AC 120-76A, Section 10.d (2))
- ❑ The immediacy of indicator should be appropriate to the function that is lost or disabled (AC 120-76A, Section 10.d (2))

2.4.15 Ensuring Integrity of EFB Data

- ❖ EFB data should be checked prior to installation to ensure that they are accurate, current, and uncorrupted
- ❑ The EFB should check that the current date is within the valid date range
- ❑ The EFB should allow data with an effective date in the future to be installed
- ❑ The system should conduct a self-test to ensure that the data is current and generate a message to the flight crew if any data is out of date. The message should indicate where to go for further information.

2.4.17 Crew Confirmation of EFB Software/Database Approval

- ❖ The latest revision information should be available upon request

2.4.19 User-Interface Customization

- ❖ There should be an easy means to return all settings to their default values
- ❑ For Part 121 and 135, the default settings should be customizable only by an administrator
- ❑ For Part 91, the default settings should be specified by the manufacturer and configurable by the user

General Principles

2.4.1 User Interface—General Design

- ❑ User interface should have a consistent set of controls and graphical elements (see also Hardware)
- ❑ Graphic elements and controls should follow personal computer conventions, except where clearly inappropriate for flight deck environment (see also Hardware)
- ❑ Menu functions should be accessible in proportion to frequency of use and criticality to mission

2.4.3 General Use of Colors

- ❖ Red and amber should be reserved for highlighting *warning* and *caution* level conditions respectively (AC 120-76A, 10.d(1))
- ❖ Color should not be sole means of coding important differences in information; color should be used redundantly
- ❖ Color-coding scheme should be interpretable easily and accurately.
- ❑ Each color should be associated with only one meaning

- ❑ No more than six colors with assigned meanings should be used in a color-coding scheme
- ❑ EFB colors should not conflict with flight deck conventions
- ❑ For Part 121 and 135, default colors that represent different types of data should be customizable only by an appropriately authorized administrator
- ❑ If colors are customizable, there should be an easy way to return to default settings

2.4.8 Alerts and Reminders

- ❖ Alerts and reminders should meet 14 CFR Part 23.1322, 25.1322, 27.1322 or 29.1322 as appropriate. Their intent should be generalized to the use of colors on displays and controls (AC 120-76A, 10.d (1))
- ❖ Red should be used only for warnings (AC 120-76A, 10.d (1))
- ❖ Amber should be used only for cautions (AC 120-76A, 10.d (1))
- ❖ Other colors should be sufficiently distinct from red/amber for use (AC 120-76A, 10.d (1))
- ❑ Alerts and reminders should be consistent with AC 25-11, 14 CFR Part 23.1311a, AMJ 25-11
- ❑ Alerts should be integrated or compatible with other flight deck alerts (AC 120-76A, 10.d (1))
- ❑ Messages should be prioritized and prioritization scheme should be documented and evaluated (AC 120-76A, 10.d (1) and AC 120-76A, 10.d (2))
- ❑ Strong attention-getting techniques (e.g., flashing or bright text) should be avoided (AC 120-76A, 10.d (1))
- ❑ During high workload phases of flight:
  - (a) Required flight information should be continuously present and unobscured, except those that indicate failure or degradation of the EFB application (AC 120-76A, 10.d (1))
  - (b) Messages should be inhibited, except those that indicate failure or degradation of the EFB application (AC 120-76A, 10.d (1))

2.4.14 Supplemental Audio

- ❑ Supplemental audio should be avoided in flight
- ❑ Users should be able to control the volume
- ❑ Users should be able to turn off the supplemental audio
- ❑ Objects with supplemental audio should be coded so the user knows of the associated audio before activating it
- ❑ Supplemental audio that is solely audio should have text description available
- ❑ Users should be able to stop the supplemental audio at any time

WORKLOAD

2.1.1 Workload

- ❑ Flight crew workload and head-down time should be minimized (AC 120-76A, Section 10.c)

3 Electronic Documents

General Principles

3.5.1 Printing

- ❖ Pages or sections selected for printing should be clearly indicated
- ❖ The user should be able to terminate printing immediately
- ❑ Users should be able to select document subsets for printing
- ❑ The printed document should have the same visual structure as the EFB electronic document

3.5.2 Animation

- ❖ Start/stop functionality should be provided. The user should be able to stop the animation at any time
- ❖ Text describing the animation should be available even if the animation is not running
- ❑ Animation should not be overused
- ❑ If supplemental audio is provided, control of the audio and video should be integrated

Formatting/Layout

3.2.1 Consistency of Information Structure

- ❖ The information structure of the electronic document should be consistent with that of the hard copy

3.3.1 Visual Layout and Structure

- ❑ Windows and frames should be placed and used consistently
- ❑ Sections of text should be separated with plenty of white space
- ❑ Data should be formatted into short segments, where possible

3.3.2 Minimum Display Area and Resolution

- ❑ The minimum document display area and resolution should be specified by the manufacturer
- ❑ Operators should meet the manufacturer-specified display area and resolution requirements for training and operational use

3.3.3 Off-Screen Text

- ❖ The existence of off-screen content should be indicated clearly and consistently (AC 120-76A, 10.b (7))
- ❑ Whether it is acceptable for parts of the document to be off-screen should be based on the application and intended function (AC 120-76A, 10.b (7))
- ❑ Information regarding the document length and the current place within the document should be constantly available

3.3.4 Active Regions

- ❖ Active regions should be clearly indicated (AC 120-76A, 10.b (8))

3.3.6 Figures

- ❖ The electronic version of a figure should show all the content in the paper version
- ❖ The entire figure should be viewable at once, even if all the details are not readable
- ❖ All the details should be readable, although the entire figure may not be visible when doing so
- ❑ Figures should be displayed in their entirety with all details readable whenever possible
- ❑ Text information should be provided for each figure, independent of whether the figure is shown in full, or marked by a placeholder
- ❑ The user should be able to configure the figure for optimal viewing

- ❑ If zooming is supported, discrete zoom levels should be available (e.g. view whole page) and the current zoom level should be displayed at all times

Interactions: Accessing Functions and Options

3.4.1 Moving to Specific Locations

- ❑ The cursor should be visible at all times (AC 120-76A, 10.b (7))
- ❑ If links are supported:
  - Entries in the table of contents should be linked to its location in the text
  - Cross-references should be linked to each other within a document
- ❑ Users should be able to return to the previous location in one step

3.4.2 Managing Multiple Open Documents

- ❖ The active document should be indicated continuously (AC 120-76A, 10.b (9))
- ❖ The user should be able to choose the active open document
- ❑ A master list of all open documents should be available

3.4.3 Searching

- ❑ Search functionality should be available
- ❑ Users should be able to select the document(s) to include in the search

4 Electronic Checklist Systems

General Principles

4.2.1 Checklists Supported by the ECL System

- ❖ If normal checklists are supported, then *all* normal checklists should be supported
- ❖ If non-normal emergency checklists are supported, then *all* non-normal checklists should be supported
- ❖ Similar requirements apply for other checklist categories
- ❑ The ECL system should indicate the location of unsupported checklists in the paper document
- ❑ Non-normal checklists should retain as much commonality with normal checklists as possible

4.5.3 Task Reminders

- ❑ Reminders for high priority, time-critical tasks should be displayed constantly once in progress and should attract attention when delayed actions should be performed
- ❑ If multiple task reminders can be shown, crews should be able to determine how many are in progress and to what tasks they refer

Formatting/Layout

4.2.2 Information and Visual Layout/Structure of Electronic Checklists

- ❖ The resulting crew actions called for in the checklist should be identical for paper and electronic versions
- ❑ Layout of items should be similar to the paper version. Headings, sub-headings, and titles should be consistent (CAP 807)
- ❑ The format of the electronic checklist should make it clear which challenge is associated with which response (CAP 708)

4.3.2 Managing Checklists

- ❖ The checklist title should be displayed above the items and be distinguished throughout the checklist
- ❑ Parent-child checklists should be integrated into a single checklist
- ❑ If more than one checklist can be open at once, a master list of checklists should be available

4.3.3 Managing Non-Normal Checklists

- ❑ All checklists associated with on-going non-normal conditions that are sensed should be listed on one master list
- ❑ A master list should indicate the status of each checklist

4.3.6 Closing All Checklists

- ❑ The ECL should allow a state where no checklists are open
- ❑ The system should give a positive indication that no checklists are open; a blank screen is not sufficient

4.4.2 Displaying Item Status

- ❖ Item status, if available, should be clearly indicated.

4.4.4 Specifying Completion of Item

- ❑ The completion status of each checklist should be indicated clearly (see also Checklist Interactions: Accessing functions and options)

4.5.4 Checklist Branching

- ❑ The selected branch should be clearly indicated
  (see also Checklist Interactions: Accessing functions and options)

Interactions: Accessing functions and options

4.3.1 Accessing Checklists

- ❖ All supported checklists should be accessible for reference/review at any time while the system is active
- ❑ Normal checklists should be accessible in accordance with the normal sequence of use
- ❑ Electronic checklists should be as quick and accurate to access as paper checklists
- ❑ The ECL system should open checklists only upon crew request

4.3.2 Managing Checklists

- ❖ The title of each open checklist should be visible continuously
- ❖ If more than one checklist can be open at once, other checklists should be accessible without closing the displayed checklist
- ❖ If more than one checklist can be open, the user should be able to select which one is active
- ❖ If a checklist is a "child" of another checklist, the user should be able to select whether the parent or child is active
- ❑ A placeholder should be used to indicate which item was active prior to leaving the checklist
- ❑ The crew should be able to reset the checklist with a simple input

4.3.4 Lengthy Checklists

- ❖ The user should be able to look ahead (e.g., page down) without changing the active item
- ❑ Information regarding the length of the checklist, the user's current position within the checklist, and how much of the checklist has been completed should be continuously available
- ❑ It should not be possible to change the status of off-screen items
- ❑ If the active item is off-screen and the user makes an "item completed" entry, an error message should appear or the active item should be called into view

4.3.5 Closing or Completing a Checklist

- ❖ If item status is tracked and the user attempts to close an incomplete checklist, the system should provide an indication that the checklist is incomplete and present any deferred/incomplete items for review
- ❖ The user should be able to close incomplete checklists after acknowledging this indication
- ❑ If item status is tracked, a positive indication should be presented when the entire checklist, as well as each item, is completed
- ❑ The action for closing/completing a checklist should be distinct from the action for marking an item as complete

4.4.1 Indicating the Active Item

- ❑ The ECL should track and indicate the active checklist item
- ❑ When returning to an incomplete checklist, the item active prior to the move should again be active

4.4.3 Moving Between Items Within a Checklist

- ❖ The active-item pointer should be moved to the next item with a simple action
- ❖ Returning to a previous item should not change the status of any item
- ❑ If the status of individual items are tracked, the user should be able to:
  (b) Move from uncompleted items, changing their status to deferred
  (c) Move to the next item automatically after completing an item

- The user should be able to quickly select one item after another; system processing should not induce delays

4.4.4 Specifying Completion of Item

- User actions to mark an item as complete should be simple
- Completed items should not be removed from the screen immediately. The crew should be able to review the item and undo their action, if necessary
- If the system indicates active items:
    a) The next item in the list should become active when an item has been completed, unless it is on the next page. A separate action should be required to move to the next page
    b) Moving to the next item without completing the current item should require an input distinct from that of specifying the item as complete
- An *undo* function should be available
- The completion status of each checklist should be indicated clearly
  (see also Checklist Formatting/Layout)

4.5.1 Links Between Checklist Items and Related Information

- The navigation between links in the ECL and related information needs to be simple and clear
- Related information should appear in a single window or area of the screen. Hyperlinks from the related information should be shown in the same window or area

4.5.2 Links to Calculated Values

- If the EFB provides calculation worksheets and allows integration between the application hosting the ECL and the application hosting the calculation worksheets, then:
  (c) Direct access to the appropriate worksheet should be provided for all items that can be calculated. This should be available for initial calculations and subsequent review/modifications
  (d) The user should be able to return easily to the checklist item from which the worksheet was accessed
- Calculated ECL values should appear in the corresponding checklist location. These fields should be blank prior to inserting the calculated value

4.5.4 Checklist Branching

- The user should be able to backup and select another decision branch
- Items not on the selected branch should not be selectable
- The selected branch should be clearly indicated
  (see also Checklist Formatting/Layout)

5 Flight Performance Calculations

General Principles

5.1.1 Default Values

- ❑ Blank data entry fields should be used to indicate that there is no system assigned default value

5.1.3 Support Information for Performance Data Entry

- ❖ The units of each variable should be clearly labeled
- ❑ Labels, formats, and units of variables should match that in other sources (e.g., paper reports, flight deck systems)
- ❑ Related information for cross-checking should be in view or easily accessible

Interactions: Accessing Functions and Options

5.1.5 Modifying Performance Calculations

- ❑ The user should be able to modify previously computed results quickly
- ❑ Output relevant to earlier calculations should be erased once the user begins modifying those calculations

Error Handling and Prevention

5.1.2 Data-entry Screening and Error Messages

- ❑ The EFB should not accept user-entered data that is of incorrect format or type. Error messages should point out suspect entries and specify the expected data type. (AC 120-76A, Section 10.d (3))
- ❑ The system should detect input errors as early as possible during data entry (AC 120-76A, Section 10.d (3))
- ❑ The system should *only* discard erroneous input errors and not the whole set of entries related to the task in progress
- ❑ The system should present an error message when required values are missing; this error message should contain the name of the required value, using the label from the input field

6 Electronic Charts

General Principles

6.2.1 Transition from Paper to Electronic Charts

- ❑ Information structure of electronic charts should match that of paper charts
- ❑ Visual structure of electronic charts should be compatible with paper charts

6.2.3 Hard Copy Backups of Electronic Charts

- ❑ If the hard copy is used as a backup, it should be of sufficient quality to be used as effectively as the original paper chart. In particular:
  - (a) The hard copy should be legible; all chart details should be visible
  - (b) The quality of the paper should be acceptable for normal use
  - (c) Color information should be distinguishable in the monochrome hard copy
  - (d) All the chart information should fit on one printed page
  - (e) The hard copy should be at least as large as a standard paper chart
  - (f) The user should be able to select the size of the hard copy

Formatting/Layout

6.2.7 Orientation of Electronic Charts

- ❖ Orientation of the charts should be indicated continuously
- ❖ When charts are oriented with respect to directionality (e.g., track/heading), and directionality information becomes unusable, it should be clear to the pilot that that information is not available
- ❑ When charts are oriented with respect to directionality (e.g., track/heading), and directionality information becomes unusable,
  (a) The crew should be notified of the unusable directionality and informed that the charts must revert to north-up orientation.
  (b) After crew acknowledgement of the failure, the charts should revert to the north-up orientation, the chart orientation indicator should be updated, and any cues that could imply directionality should be removed
- ❑ Text and symbols other than those designed to reflect compass orientation should remain upright at all times
- ❑ Crew input should be required to change the orientation of the charts

Interactions: Accessing functions and options

6.2.5 Basic Zooming and Panning

- ❑ If zooming is supported, then panning should also be supported, and vice versa
- ❑ The chart's visual edges should be clearly marked. Visual edges should be shown only when no more information is outside that area
- ❑ When panning, the user should know which way to move to bring more of the chart into view
- ❑ Panning to an area where no portion of the chart will be displayed should be prevented
- ❑ If the user can change zoom levels, the user should be able to return to a default view easily
- ❑ If the display can be panned, the user should be able to return to a default view easily
- ❑ Zooming and panning should not result in lengthy processing delays

6.2.9 Access to Individual Charts

- ❖ The currently selected chart's label should be displayed continuously
- ❑ The system should allow rapid access to pre-selected charts
- ❑ The chart application should help the crew ensure that the correct chart was selected and allow corrections to be made quickly when an error occurs
- ❑ Multiple search methods should be supported

- ❑ Search results should be ordered with its best guesses at the top of the list and least likely to be used charts at the bottom
- ❑ Selection of alternate runways should be facilitated during approach

6.2.11 De-cluttering and Display Configuration

- ❑ The pilot should not be able to declutter safety critical display elements without knowing they are suppressed
- ❑ Changing map scale, orientation, and other options and settings should not induce significant levels of workload
- ❑ The information prioritization scheme should be documented

Error Handling and Prevention

6.2.2 Updates to Electronic Charts

- ❑ Corrections/updates should be made directly within the electronic chart application, unless they are temporary
- ❑ Corrections/updates that are of high priority or time-sensitive should not be made via paper notifications

6.2.4 Scale Information

- ❖ Scale information should always be visible for charts drawn to scale
- ❖ Scale information should be accurate. Scale information should be updated when the display is zoomed
- ❖ Static scale information should be removed unless it is always accurate
- ❖ Charts drawn "not to scale" should have a label indicating that fact continuously

6.2.10 Knowledge and Display of Own-Aircraft Position

- ❖ Display of ownship should not be supported on non-georeferenced or not-to-scale terminal charts
- ❖ See TSO C-165 and DO-257A for other applicable requirements
- ❑ The range of display zoom levels should be compatible with the position accuracy of the ownship symbol.
- ❑ An indication of ownship position should be provided if the chart is zoomed or panned such that ownship is not in the current view

Appendix D:
Tools from April 2004

A full-length high level tool is presented in Section D.1and a detailed tool is presented in Section D.2. The high level tool contains a list of topics to consider during an EFB usability evaluation. Specific topics relevant to electronic documents, electronic checklists, flight performance calculations, and electronic chart applications have been integrated into the tool. Evaluators go through the list commenting on each item. The detailed tool consists of guidance from Appendix B of the EFB Version 2 document that has been reordered to correspond to the latest item order of the high level tool. Because the item order changed slightly from the Appendix C version, the entire detailed tool is provided again.

D.1 Full-Length High Level Tool

EFB Usability Assessment Tool

HARDWARE CONSIDERATIONS

- Physical Ease of Use
  — Input devices and display, accessibility of controls
- Labels and Controls
- Lighting Issues (day vs. night use)
  — Brightness adjustment, illumination of labels
- Amount of feedback, potential for errors

SOFTWARE CONSIDERATIONS

Symbols and Graphical Icons

- Clarity of intended meaning, confusability
- Legibility and distinctiveness

Formatting/Layout

- Fonts (size, style, case, spacing)
- Arrangement of information on the display
  — Consistency with user expectations and internal logic

Electronic Documents

- Indication of active regions and off-screen material
- Figures/tables
- Page format
- Structure and organization, consistency with hard copy

Electronic Checklists

- Display of item status, e.g., open, deferred, completed
- Indication of checklist status, e.g., open, closed, completed, active
- Formatting (e.g., associating challenges with responses)
- Consistency with hard copy

Electronic Charts

- Formatting
- Structure and organization, consistency with hard copy

Interaction (Accessing functions and options)

- Home pages and ease of movement between pages
- Number of inputs to complete a task
- Ease of accessing functions and options
- Feedback (system state, alerts, modes, etc)
- Responsiveness
- Intuitive logic

Electronic Documents

- Moving within a document, moving between documents

- Identifying open documents, identifying current document
- Zooming
- Search functionality

Electronic Checklists
- Accessing checklists and moving between checklists
- Managing checklists, e.g., parent-child relationships, master list
- Identifying open checklists, identifying current checklist
- Moving between items
- Linking between items, calculated values, other related information

Flight Performance Calculations
- Modifying performance calculations

Electronic Charts
- Access to charts
- Identifying open charts, identifying current charts
- Zooming and panning
- De-cluttering and display configuration (e.g., scale, orientation)
- Search functionality

Error handling and prevention
- Susceptibility to error (mode errors, selection errors, data entry errors, reading errors, etc.)
- Correcting errors (e.g., cancel, clear, undo)
- Error messages

Electronic Charts
- Updating chart information
- Scale information

Flight Performance Calculations
- Data entry

Multiple Applications
- Consistency and compatibility across applications
- Identifying current position within system
- Ease of switching between applications

Automation (if any)
- Is there enough? Too much?
- Is it disruptive/supportive? Predictable? User control over automation? (e.g., manual override)

General
- Consistency of controls/elements; are they distinctive where appropriate?
- Visual, audio, and tactile characteristics
- Use of color (esp. red and amber) and color-coding
- Amount of feedback (system state, alerts, modes, etc)
- Clarity and consistency of language, terms, and abbreviations
- End-user customization (if any)

Electronic Documents

- Printing (if available), printouts
- Animation (if any)

Electronic Checklists

- Set of checklists that are supported
- Presentation of task reminders (if any)

Flight Performance Calculations

- Unit labels
- Default values

Electronic Charts

- If own-aircraft/ownship display, see TSO C-165
- Printing (if available), printouts

WORKLOAD

- Problem areas

OTHER

D.2 Detailed Tool

2 General EFB System

HARDWARE CONSIDERATIONS

2.1.5 Legibility—Lighting Issues

- ❖ Automatic brightness adjustment should be independent for each EFB (See AC 25-11)
- ❑ Screen brightness should adjustable in fine increments or continuously
- ❑ Buttons and labels should be adequately illuminated for night use

2.2.4 Kneeboard EFBs

- ❖ Kneeboard EFB should be easily removable

2.4.1 User Interface—General Design

- ❑ User interface should have a consistent set of controls and graphical elements (see also General Principles)
- ❑ Controls used for different functions should be visually distinct
- ❑ Graphic elements and controls should follow personal computer conventions, except where clearly inappropriate for flight deck environment (see also General Principles)

2.5.1 Pointing and Cursor Control Devices

- ❑ Input devices should be selected and customized based on the type and complexity of the entries to be made and flight deck environmental factors that affect its usability
- ❑ Performance parameters should be tailored for the intended application and for the flight deck environment
- ❑ Users should be able to rest and/or stabilize their hand when using the pointer or cursor control device
- ❑ Active areas should be sized to permit accurate selection with the pointer/cursor device under all operating conditions

2.5.2 Hardware Controls

- ❖ All controls should be properly labeled (14 CFR 23.1555, 25.1555, and 27.1555)
- ❖ All soft function keys should be labeled
- ❖ Inactive soft function keys should not be labeled or should use a visual convention to indicate that the function is not available
- ❑ Physical function keys should provide tactile feedback when pushed
- ❑ Key repeats should be filtered by the software if they occur too closely together
- ❑ Soft function keys should be drawn in a reserved space outside the main content area
- ❑ The same function should appear on the same function key, whenever possible
- ❑ Labels should be consistent
- ❑ Labels should be clear and brief
- ❑ Labels may use standard abbreviations; ambiguous abbreviations should be avoided
- ❑ Labels should be located near the controls they identify and should not be confusingly close to other labels or other controls
- ❑ Labels should be drawn in horizontal text
- ❑ Physical controls should be collocated with the display
- ❑ The most frequently used controls should be placed at the most accessible locations
- ❑ Controls presented in a small space may need to be grouped according to function and/or order of use
- ❑ Controls should be designed to deter inadvertent activation

2.5.3 Display

- ❑ The physical nature of the display screen should minimize the likelihood that information will be obscured

2.5.5 Keyboards

- ❑ Keyboard type should be appropriate for the given task
  - ▪ QWERTY type keyboards should be used for text entry
  - ▪ Numeric keypads are best suited for significant numeric entries
- ❑ Keyboards should provide appropriate tactile feedback
- ❑ Users should be able to rest/stabilize their hand to use the keyboard, especially during turbulence

SOFTWARE

Symbols and Graphical Icons

2.4.4 Graphical Icons

- ❑ Icons should be accompanied with text labels
- ❑ Design of icons should minimize training and maximize intuitiveness for cross-cultural use

2.4.13 Non-Text Display Elements

- ❑ Non-text display elements should be distinguishable based on shape alone, without relying on secondary cues such as color or labels
- ❑ Non-text display elements should be designed for legibility on minimum expected display resolution viewed from the maximal intended viewing distance

Formatting/Layout

2.4.10 Legibility of Text—Characters

- ❑ Typeface should be highly legible. HFDS recommends:
  - Spare use of upper case text (8.2.5.8.2)
  - Mixed upper and lower case for continuous text (8.2.5.8.4)
  - Serif fonts for high resolution displays (8.2.5.7.5)
  - Sans serif fonts otherwise (8.2.5.7.6)
  - Character contrast between 6:1 and 10:1 (8.2.5.6.12)
  - Characters stroke width 10 to 12% of character height (8.2.5.6.14)
- ❑ Individual characters should not be easily confused with other characters
- ❑ Slanting or italic text should be avoided

2.4.11 Legibility of Text—Typeface Size and Width

- ❖ Typeface should be appropriate for viewing distance, lighting conditions, and text criticality
- ❑ The FAA HFDS recommends that:
  - i. Minimum character height should be 1/200 of viewing distance, e.g., for 35” viewing distance, 0.175” tall (17.5 pixels at 100 pix/inch) (8.2.5.6.6)
  - ii. Preferred character height should be 1/167 of viewing distance (8.2.5.6.5)
  - iii. Character height to width ratios should be (8.2.5.6.10)
    - o <80 char per line, 1 to 0.7 up to 0.9 (15 pix tall, 10.5 to 13.5 pix wide) for monotype fonts
    - o >80 char per line, at least 1 to 0.5 (15 pix tall, 7.5 pix wide)
    - o 1:1 for M and W in a proportional font
- ❑ Larger fonts should be used for text read in poor viewing conditions

2.4.12 Legibility of Text—Spacing for Readability

- ❑ Text should be spaced appropriately to facilitate reading

- ❑ Line lengths should be appropriate for text content
- ❑ To facilitate readability, HFDS recommends the following:
  - (a) Use horizontal spacing between characters that is at least 10% of character height (15 pix tall, 1.5 pix spacing) (8.2.5.6.1)
  - (b) Use spacing between words of at least one character for equally spaced characters, or width of "N" for proportional fonts (8.2.5.6.2)
  - (c) Use spacing between lines of at least two stroke widths or 0.15 of character height (15 pix tall, 2.25 pix leading), whichever is greater (8.2.5.6.3)
  - (d) Separate paragraphs with blank line (8.2.5.6.4)

Interactions: Accessing Functions and Options

2.4.5 Multi-Tasking

- ❑ The user should be able to identify the active application easily
- ❑ The user should be able to:
  - Select which of the open applications is currently active
  - Switch between applications easily
- ❑ Applications, running in the background, should be in the same state when the user returns to it, other than the completion of any background processing
- ❑ Responsiveness of an individual application should not suffer when all applications are running simultaneously
- ❑ The user should be able to exit applications with pending activities by completing them or by acknowledging that they are incomplete
- ❑ The system should discourage use of non-flight-related applications and ask for an extra confirmation to launch

2.4.6 Responsiveness

- ❖ The system should provide feedback when a user input is processed
  - Alphanumeric inputs should be shown within 0.2 seconds (SAE ARP 4791)
- ❖ A "system busy" indicator should be displayed if user inputs can not be processed within 0.5 seconds (SAE ARP 4791)
- ❑ The EFB applications should have a "system busy" indicator
- ❑ The type of feedback should be appropriate for the type of user input
- ❑ If tasks take more than a few seconds to complete, indicators should show their progress
- ❑ User entries made while the system is busy should be stored for later processing

2.4.7 Anchor Locations

- ❑ If the EFB supports more than one application, there should be an anchor location from which the user moves between applications
- ❑ Each EFB application should have its own anchor page
- ❑ It should be easy to move from any location in the EFB to an anchor location, and vice versa

2.4.18 Links to Related Material

- ❑ A consistent philosophy should be used for accessing different types of information. Similar types of information should be accessed in the same way
- ❑ Users should be able to keep track of how to move between topics. Users should be able to return to the starting point easily

Error Handling and Prevention

2.1.7 Failure Modes

- ❑ EFB should alert the flight crew to probable application/system failures (AC 120-76A, Section 10.e (2))

2.4.8 Display of System Status

- ❑ Any full or partial application failure should be indicated with a positive indicator (AC 120-76A, Section 10.d (2))
- ❑ The immediacy of indicator should be appropriate to the function that is lost or disabled (AC 120-76A, Section 10.d (2))

2.4.15 Ensuring Integrity of EFB Data

- ❖ EFB data should be checked prior to installation to ensure that they are accurate, current, and uncorrupted
- ❑ The EFB should check that the current date is within the valid date range
- ❑ The EFB should allow data with an effective date in the future to be installed
- ❑ The system should conduct a self-test to ensure that the data is current and generate a message to the flight crew if any data is out of date. The message should indicate where to go for further information.

2.4.17 Crew Confirmation of EFB Software/Database Approval

- ❖ The latest revision information should be available upon request

2.4.19 User-Interface Customization

- ❖ There should be an easy means to return all settings to their default values
- ❑ For Part 121 and 135, the default settings should be customizable only by an administrator
- ❑ For Part 91, the default settings should be specified by the manufacturer and configurable by the user

Multiple Applications

2.4.2 Application Compatibility and Style Guides

- ❑ All applications should follow a common style guide, preferably specific to that aircraft
- ❑ Color and other formatting should be internally consistent across applications (AC 120-76A, Section 10.b (1))
- ❑ Help facility, if available, should be standardized across applications
- ❑ Soft key labels and menus should be consistent across applications
- ❑ Common actions allowed on multiple applications should be performed in the same manner (see also Interactions: Accessing functions and options)
- ❑ Manufacturers should prepare style guides for third party developers

General Principles

2.4.1 User Interface—General Design

- ❑ User interface should have a consistent set of controls and graphical elements (see also Hardware)
- ❑ Graphic elements and controls should follow personal computer conventions, except where clearly inappropriate for flight deck environment (see also Hardware)
- ❑ Menu functions should be accessible in proportion to frequency of use and criticality to mission

2.4.3 General Use of Colors

- ❖ Red and amber should be reserved for highlighting *warning* and *caution* level conditions respectively (AC 120-76A, 10.d(1))

- ❖ Color should not be sole means of coding important differences in information; color should be used redundantly
- ❖ Color-coding scheme should be interpretable easily and accurately.
- ❑ Each color should be associated with only one meaning
- ❑ No more than six colors with assigned meanings should be used in a color-coding scheme
- ❑ EFB colors should not conflict with flight deck conventions
- ❑ For Part 121 and 135, default colors that represent different types of data should be customizable only by an appropriately authorized administrator
- ❑ If colors are customizable, there should be an easy way to return to default settings

2.4.8 Alerts and Reminders

- ❖ Alerts and reminders should meet 14 CFR Part 23.1322, 25.1322, 27.1322 or 29.1322 as appropriate. Their intent should be generalized to the use of colors on displays and controls (AC 120-76A, 10.d (1))
- ❖ Red should be used only for warnings (AC 120-76A, 10.d (1))
- ❖ Amber should be used only for cautions (AC 120-76A, 10.d (1))
- ❖ Other colors should be sufficiently distinct from red/amber for use (AC 120-76A, 10.d (1))
- ❑ Alerts and reminders should be consistent with AC 25-11, 14 CFR Part 23.1311a, AMJ 25-11
- ❑ Alerts should be integrated or compatible with other flight deck alerts (AC 120-76A, 10.d (1))
- ❑ Messages should be prioritized and prioritization scheme should be documented and evaluated (AC 120-76A, 10.d (1) and AC 120-76A, 10.d (2))
- ❑ Strong attention-getting techniques (e.g., flashing or bright text) should be avoided (AC 120-76A, 10.d (1))
- ❑ During high workload phases of flight:
  - (a) Required flight information should be continuously present and unobscured, except those that indicate failure or degradation of the EFB application (AC 120-76A, 10.d (1))
  - (b) Messages should be inhibited, except those that indicate failure or degradation of the EFB application (AC 120-76A, 10.d (1))

2.4.14 Supplemental Audio

- ❑ Supplemental audio should be avoided in flight
- ❑ Users should be able to control the volume
- ❑ Users should be able to turn off the supplemental audio
- ❑ Objects with supplemental audio should be coded so the user knows of the associated audio before activating it
- ❑ Supplemental audio that is solely audio should have text description available
- ❑ Users should be able to stop the supplemental audio at any time

WORKLOAD

2.1.1 Workload

- ❑ Flight crew workload and head-down time should be minimized (AC 120-76A, Section 10.c)

3 Electronic Documents

Formatting/Layout

3.2.1 Consistency of Information Structure

- ❖ The information structure of the electronic document should be consistent with that of the hard copy

3.3.1 Visual Layout and Structure

- ❑ Windows and frames should be placed and used consistently
- ❑ Sections of text should be separated with plenty of white space
- ❑ Data should be formatted into short segments, where possible

3.3.2 Minimum Display Area and Resolution

- ❑ The minimum document display area and resolution should be specified by the manufacturer
- ❑ Operators should meet the manufacturer-specified display area and resolution requirements for training and operational use

3.3.3 Off-Screen Text

- ❖ The existence of off-screen content should be indicated clearly and consistently (AC 120-76A, 10.b (7))
- ❑ Whether it is acceptable for parts of the document to be off-screen should be based on the application and intended function (AC 120-76A, 10.b (7))
- ❑ Information regarding the document length and the current place within the document should be constantly available

3.3.4 Active Regions

- ❖ Active regions should be clearly indicated (AC 120-76A, 10.b (8))

3.3.6 Figures

- ❖ The electronic version of a figure should show all the content in the paper version
- ❖ The entire figure should be viewable at once, even if all the details are not readable
- ❖ All the details should be readable, although the entire figure may not be visible when doing so
- ❑ Figures should be displayed in their entirety with all details readable whenever possible
- ❑ Text information should be provided for each figure, independent of whether the figure is shown in full, or marked by a placeholder
- ❑ The user should be able to configure the figure for optimal viewing
- ❑ If zooming is supported, discrete zoom levels should be available (e.g. view whole page) and the current zoom level should be displayed at all times

Interactions: Accessing Functions and Options

3.4.1 Moving to Specific Locations

- ❑ The cursor should be visible at all times (AC 120-76A, 10.b (7))
- ❑ If links are supported:
  - Entries in the table of contents should be linked to its location in the text
  - Cross-references should be linked to each other within a document
- ❑ Users should be able to return to the previous location in one step

3.4.2 Managing Multiple Open Documents

- ❖ The active document should be indicated continuously (AC 120-76A, 10.b (9))
- ❖ The user should be able to choose the active open document
- ❑ A master list of all open documents should be available

3.4.3 Searching

- ❑ Search functionality should be available
- ❑ Users should be able to select the document(s) to include in the search

General Principles

3.5.1 Printing

- ❖ Pages or sections selected for printing should be clearly indicated
- ❖ The user should be able to terminate printing immediately
- ❑ Users should be able to select document subsets for printing
- ❑ The printed document should have the same visual structure as the EFB electronic document

3.5.2 Animation

- ❖ Start/stop functionality should be provided. The user should be able to stop the animation at any time
- ❖ Text describing the animation should be available even if the animation is not running
- ❑ Animation should not be overused
- ❑ If supplemental audio is provided, control of the audio and video should be integrated

4 Electronic Checklist Systems

Formatting/Layout

4.2.2 Information and Visual Layout/Structure of Electronic Checklists

- ❖ The resulting crew actions called for in the checklist should be identical for paper and electronic versions
- ❑ Layout of items should be similar to the paper version. Headings, sub-headings, and titles should be consistent (CAP 807)
- ❑ The format of the electronic checklist should make it clear which challenge is associated with which response (CAP 708)

4.3.2 Managing Checklists

- ❖ The checklist title should be displayed above the items and be distinguished throughout the checklist
- ❑ Parent-child checklists should be integrated into a single checklist
- ❑ If more than one checklist can be open at once, a master list of checklists should be available

4.3.3 Managing Non-Normal Checklists

- ❑ All checklists associated with on-going non-normal conditions that are sensed should be listed on one master list
- ❑ A master list should indicate the status of each checklist

4.3.6 Closing All Checklists

- ❑ The ECL should allow a state where no checklists are open
- ❑ The system should give a positive indication that no checklists are open; a blank screen is not sufficient

4.4.2 Displaying Item Status

- ❖ Item status, if available, should be clearly indicated.

4.4.4 Specifying Completion of Item

- ❑ The completion status of each checklist should be indicated clearly
  (see also Interactions: Accessing functions and options)

4.5.4 Checklist Branching

- ❑ The selected branch should be clearly indicated
  (see also Interactions: Accessing functions and options)

Interactions: Accessing functions and options

4.3.1 Accessing Checklists

- ❖ All supported checklists should be accessible for reference/review at any time while the system is active
- ❑ Normal checklists should be accessible in accordance with the normal sequence of use
- ❑ Electronic checklists should be as quick and accurate to access as paper checklists
- ❑ The ECL system should open checklists only upon crew request

4.3.2 Managing Checklists

- ❖ The title of each open checklist should be visible continuously
- ❖ If more than one checklist can be open at once, other checklists should be accessible without closing the displayed checklist

- ❖ If more than one checklist can be open, the user should be able to select which one is active
- ❖ If a checklist is a "child" of another checklist, the user should be able to select whether the parent or child is active
- ❑ A placeholder should be used to indicate which item was active prior to leaving the checklist
- ❑ The crew should be able to reset the checklist with a simple input

4.3.4 Lengthy Checklists

- ❖ The user should be able to look ahead (e.g., page down) without changing the active item
- ❑ Information regarding the length of the checklist, the user's current position within the checklist, and how much of the checklist has been completed should be continuously available
- ❑ It should not be possible to change the status of off-screen items
- ❑ If the active item is off-screen and the user makes an "item completed" entry, an error message should appear or the active item should be called into view

4.3.5 Closing or Completing a Checklist

- ❖ If item status is tracked and the user attempts to close an incomplete checklist, the system should provide an indication that the checklist is incomplete and present any deferred/incomplete items for review
- ❖ The user should be able to close incomplete checklists after acknowledging this indication
- ❑ If item status is tracked, a positive indication should be presented when the entire checklist, as well as each item, is completed
- ❑ The action for closing/completing a checklist should be distinct from the action for marking an item as complete

4.4.1 Indicating the Active Item

- ❑ The ECL should track and indicate the active checklist item
- ❑ When returning to an incomplete checklist, the item active prior to the move should again be active

4.4.3 Moving Between Items Within a Checklist

- ❖ The active-item pointer should be moved to the next item with a simple action
- ❖ Returning to a previous item should not change the status of any item
- ❑ If the status of individual items are tracked, the user should be able to:
  - (d) Move from uncompleted items, changing their status to deferred
  - (e) Move to the next item automatically after completing an item
- ❑ The user should be able to quickly select one item after another; system processing should not induce delays

4.4.4 Specifying Completion of Item

- ❖ User actions to mark an item as complete should be simple
- ❖ Completed items should not be removed from the screen immediately. The crew should be able to review the item and undo their action, if necessary
- ❑ If the system indicates active items:
  - a) The next item in the list should become active when an item has been completed, unless it is on the next page. A separate action should be required to move to the next page
  - b) Moving to the next item without completing the current item should require an input distinct from that of specifying the item as complete
- ❑ An *undo* function should be available
- ❑ The completion status of each checklist should be indicated clearly
  (see also Formatting/Layout)

4.5.1 Links Between Checklist Items and Related Information

- ❑ The navigation between links in the ECL and related information needs to be simple and clear
- ❑ Related information should appear in a single window or area of the screen. Hyperlinks from the related information should be shown in the same window or area

4.5.2 Links to Calculated Values

- ❑ If the EFB provides calculation worksheets and allows integration between the application hosting the ECL and the application hosting the calculation worksheets, then:
  i. Direct access to the appropriate worksheet should be provided for all items that can be calculated. This should be available for initial calculations and subsequent review/modifications
  ii. The user should be able to return easily to the checklist item from which the worksheet was accessed
- ❑ Calculated ECL values should appear in the corresponding checklist location. These fields should be blank prior to inserting the calculated value

4.5.4 Checklist Branching

- ❑ The user should be able to backup and select another decision branch
- ❑ Items not on the selected branch should not be selectable
- ❑ The selected branch should be clearly indicated
  (see also Formatting/Layout)

General Principles

4.2.1 Checklists Supported by the ECL System

- ❖ If normal checklists are supported, then *all* normal checklists should be supported
- ❖ If non-normal emergency checklists are supported, then *all* non-normal checklists should be supported
- ❖ Similar requirements apply for other checklist categories
- ❑ The ECL system should indicate the location of unsupported checklists in the paper document
- ❑ Non-normal checklists should retain as much commonality with normal checklists as possible

4.5.3 Task Reminders

- ❑ Reminders for high priority, time-critical tasks should be displayed constantly once in progress and should attract attention when delayed actions should be performed
- ❑ If multiple task reminders can be shown, crews should be able to determine how many are in progress and to what tasks they refer

5 Flight Performance Calculations

Interactions: Accessing Functions and Options

5.1.5 Modifying Performance Calculations

- The user should be able to modify previously computed results quickly
- Output relevant to earlier calculations should be erased once the user begins modifying those calculations

Error Handling and Prevention

5.1.2 Data-entry Screening and Error Messages

- The EFB should not accept user-entered data that is of incorrect format or type. Error messages should point out suspect entries and specify the expected data type. (AC 120-76A, Section 10.d (3))
- The system should detect input errors as early as possible during data entry (AC 120-76A, Section 10.d (3))
- The system should *only* discard erroneous input errors and not the whole set of entries related to the task in progress
- The system should present an error message when required values are missing; this error message should contain the name of the required value, using the label from the input field

General Principles

5.1.1 Default Values

- Blank data entry fields should be used to indicate that there is no system assigned default value

5.1.3 Support Information for Performance Data Entry

- The units of each variable should be clearly labeled
- Labels, formats, and units of variables should match that in other sources (e.g., paper reports, flight deck systems)
- Related information for cross-checking should be in view or easily accessible

6 Electronic Charts

Formatting/Layout

6.2.7 Orientation of Electronic Charts

- ❖ Orientation of the charts should be indicated continuously
- ❖ When charts are oriented with respect to directionality (e.g., track/heading), and directionality information becomes unusable, it should be clear to the pilot that that information is not available
- ❑ When charts are oriented with respect to directionality (e.g., track/heading), and directionality information becomes unusable,
  (a) The crew should be notified of the unusable directionality and informed that the charts must revert to north-up orientation.
  (b) After crew acknowledgement of the failure, the charts should revert to the north-up orientation, the chart orientation indicator should be updated, and any cues that could imply directionality should be removed
- ❑ Text and symbols other than those designed to reflect compass orientation should remain upright at all times
- ❑ Crew input should be required to change the orientation of the charts

Interactions: Accessing functions and options

6.2.5 Basic Zooming and Panning

- ❑ If zooming is supported, then panning should also be supported, and vice versa
- ❑ The chart's visual edges should be clearly marked. Visual edges should be shown only when no more information is outside that area
- ❑ When panning, the user should know which way to move to bring more of the chart into view
- ❑ Panning to an area where no portion of the chart will be displayed should be prevented
- ❑ If the user can change zoom levels, the user should be able to return to a default view easily
- ❑ If the display can be panned, the user should be able to return to a default view easily
- ❑ Zooming and panning should not result in lengthy processing delays

6.2.9 Access to Individual Charts

- ❖ The currently selected chart's label should be displayed continuously
- ❑ The system should allow rapid access to pre-selected charts
- ❑ The chart application should help the crew ensure that the correct chart was selected and allow corrections to be made quickly when an error occurs
- ❑ Multiple search methods should be supported
- ❑ Search results should be ordered with its best guesses at the top of the list and least likely to be used charts at the bottom
- ❑ Selection of alternate runways should be facilitated during approach

6.2.11 De-cluttering and Display Configuration

- ❑ The pilot should not be able to declutter safety critical display elements without knowing they are suppressed
- ❑ Changing map scale, orientation, and other options and settings should not induce significant levels of workload
- ❑ The information prioritization scheme should be documented

Error Handling and Prevention

6.2.2 Updates to Electronic Charts

- ❑ Corrections/updates should be made directly within the electronic chart application, unless they are temporary
- ❑ Corrections/updates that are of high priority or time-sensitive should not be made via paper notifications

6.2.4 Scale Information

- ❖ Scale information should always be visible for charts drawn to scale
- ❖ Scale information should be accurate. Scale information should be updated when the display is zoomed
- ❖ Static scale information should be removed unless it is always accurate
- ❖ Charts drawn "not to scale" should have a label indicating that fact continuously

6.2.10 Knowledge and Display of Own-Aircraft Position

- ❖ Display of ownship should not be supported on non-georeferenced or not-to-scale terminal charts
- ❖ See TSO C-165 and DO-257A for other applicable requirements
- ❑ The range of display zoom levels should be compatible with the position accuracy of the ownship symbol.
- ❑ An indication of ownship position should be provided if the chart is zoomed or panned such that ownship is not in the current view

General Principles

6.2.1 Transition from Paper to Electronic Charts

- ❑ Information structure of electronic charts should match that of paper charts
- ❑ Visual structure of electronic charts should be compatible with paper charts

6.2.3 Hard Copy Backups of Electronic Charts

- ❑ If the hard copy is used as a backup, it should be of sufficient quality to be used as effectively as the original paper chart. In particular:
  - (a) The hard copy should be legible; all chart details should be visible
  - (b) The quality of the paper should be acceptable for normal use
  - (c) Color information should be distinguishable in the monochrome hard copy
  - (d) All the chart information should fit on one printed page
  - (e) The hard copy should be at least as large as a standard paper chart
  - (f) The user should be able to select the size of the hard copy

Appendix E:
Industry Presentation from April 15, 2004

Update to Industry on Tools for Structured Evaluation of Electronic Flight Bag Usability

Divya Chandra
US DOT Volpe Center, Cambridge, Massachusetts
Chandra@volpe.dot.gov

Michelle Yeh
US DOT Volpe Center, Cambridge, Massachusetts

Vic Riley
User Interaction Research and Design, Point Roberts, WA

http://www.volpe.dot.gov/opsad/efb

April 15, 2004

*This work is sponsored by the FAA (AAR-100)

Purpose

- **Update industry on Volpe EFB activities since publication of the Version 2 document last Fall**
  - **Describe latest version of our EFB usability assessment tools**
- **Present proposals for how industry could use these tools and conduct in-house usability evaluations**
- **Obtain feedback**
  - **Questions, comments, and concerns regarding the content, design, and use of the tools**
  - **Discussion forum TBD**
    - **Telecon, net-meeting, or maybe next meeting of Air Transport Association Digital Displays Working Group**
    - **Date TBD, around mid-May**

Overview

- **Introduction**
  - **Volpe EFB Research Overview**
  - **Disclaimers**
- EFB Assessment Tools
  - Detailed Tool
  - High-level Tool
- Steps for Conducting a Usability Evaluation
  - Setting up the evaluation
  - Performing the evaluation
  - Synthesizing the data
- Conclusion
  - Benefits to the FAA and Manufacturers
  - References

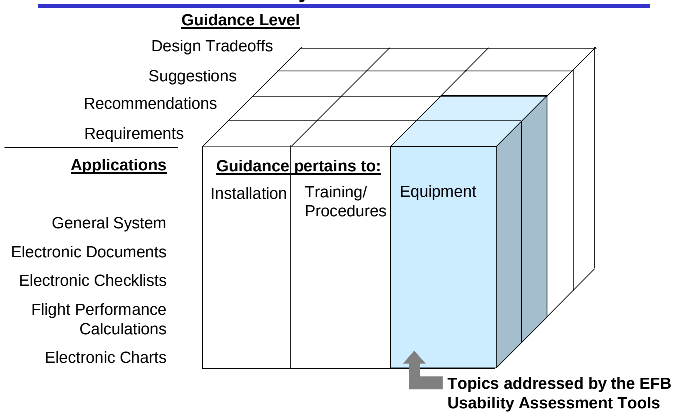
Volpe EFB Research Overview

- **EFB Human Factors Considerations Document (Version 2, September, 2003)**
  - Referenced in FAA EFB Advisory Circular (AC 120-76A)
  - HF Guidance for General EFB System, Electronic Documents, Electronic Checklist Systems, Flight Performance Calculations, and Electronic Charts
  - Appendix A, Industry Snapshot
  - Appendix B, Summary of EFB Equipment Requirements and Recommendations

- **AC 120-76A identifies need for evaluating EFBs from a human factors perspective and provides guidance on topics to consider, but does not specify a structure for the evaluation procedure**

- **EFB Usability Assessment Tools**
  - Detailed tool is a checklist based on Appendix B of Version 2 document (described above)
  - High-level tool contains a list of user interface topics for review
  - Both tools are designed for equipment evaluations only
  - Full report on tool due out in late-summer 2004

Relationship Between EFB Version 2 Document and Scope of the Usability Assessment Tools



Disclaimers

- No tool has been adopted by the FAA
- Research only
  - Content and use of the tools is not regulatory
- Intended end user of the tool is FAA
  - We think the tools may also be useful for manufacturers
  - Manufacturer feedback needed
- Tools may change based upon FAA feedback in next few months
  - To date, tested tools with (non-FAA) researchers as evaluators
    - Had aviation and/or human factors background
    - Provided FAA regulatory materials (e.g., AC120-76A and AC 25-11) in advance for review

April 15, 2004 6

Overview

- Introduction
  - Volpe EFB Research Overview
  - Disclaimers
- EFB Assessment Tools
  - Detailed Tool
  - High-level Tool
- Steps for Conducting a Usability Evaluation
  - Setting up the evaluation
  - Performing the evaluation
  - Synthesizing the data
- Conclusion
  - Benefits to the FAA and Manufacturers
  - References

April 15, 2004 7

Description of Detailed EFB Usability Assessment Tool

- Detailed tool contains the Version 2 Appendix B content (equipment requirements and recommendations) in a tighter format
  - Paraphrased versions of equipment requirements/recommendations
  - Non-regulatory "requirements" are indicated with a ❖
  - Recommendations are indicated with a ❑
  - Statements are "**heuristics**"; If system complies with heuristics, the result is a "more usable" system relative to systems that do not comply with the heuristics
  - References back to full EFB Version 2 document are provided
- Sample from detailed tool

2.4.10 Legibility of Text—Typeface Size and Width

❖ Typeface should be appropriate for viewing distance, lighting conditions, and text criticality

❑ The FAA HFDS recommends that:

(a) Minimum character height should be 1/200 of viewing distance, e.g., for 35" viewing distance, 0.175" tall (17.5 pixels at 100 pix/inch) (8.2.5.6.6)

(b) Preferred character height should be 1/167 of viewing distance (8.2.5.6.5)

April 15, 2004 8

Proposal for Use of Detailed Tool by Industry

- **Manufacturers evaluate the system using this tool and present results to FAA for review**
  - Why complete the detailed tool prior to FAA evaluations?
    - Helps manufacturers anticipate and prepare for FAA evaluation
    - Completing the detailed tool takes more time than FAA will have on-site
      - Some important items are very specific and require detailed analysis
      - Using the detailed tool requires familiarity with terms used in Version 2 document, which on-site FAA evaluators may not have
- **Detailed tool review could be performed as part of system design**
  - Consider whether the device meets the requirement or recommendation specified
  - Rating scale is flexible; FAA may want to know only the problem areas, but manufacturer may want to know both strengths and weaknesses
  - Open issues identified by the detailed tool could be examined further in the usability evaluation

Description of High Level EFB Usability Assessment Tool

- **List of topics to comment upon**
  - **Hardware**
  - **Software**
    - **Symbols and Graphical Icons**
    - **Formatting/Layout**
    - **Interaction (Accessing functions and options)**
    - **Error Handling and Prevention**
    - **Multiple Applications**
    - **Automation (if any)**
    - **General Principles**
  - **Workload**
  - **Other**
- **For each topic, more specific sub-topics are provided**
- **Items specific to electronic documents, electronic checklists, flight performance calculations, and electronic chart applications also provided**
- **Tool designed for use by FAA evaluators (in a brief evaluation) but could be used by manufacturers during system development**

Comparison of Detailed Tool and High-Level Tool

- **Similarities**
  - **Both tools help evaluators articulate positive and negative aspects of user interface**
  - **Both tools promote a comprehensive review of the interface**
  - **Many of the same issues are caught with both tools (but not all)**

- **Differences**
  - **Tools address issues at two different levels.**
    - **Detailed tool provides specific heuristics (see slide #8)**
    - **High-level tool provides list of topics only (e.g., "Fonts"), without heuristics**
  - **Because the detailed tool provides many heuristics, it takes more time to complete than the high level tool**
  - **Detailed tool contains objective and subjective items; compliance with heuristics can be "rated" (e.g., acceptable or not) to determine problem severity. Completion of high level tool and need to comply is more subjective.**

Overview

- Introduction
  - Volpe EFB Research Overview
  - Disclaimers
- EFB Assessment Tools
  - Detailed Tool
  - High-level Tool
- **Conducting a Usability Evaluation**
  - **Setting up the evaluation**
  - **Performing the evaluation**
  - **Synthesizing the data**
- Conclusion
  - Benefits to the FAA and Manufacturers
  - References

Conducting an In-House Usability Evaluation

- **Allow 2 to 3 weeks for the whole evaluation**
  - **Setting up the evaluation (3-5 days)**
  - **Performing the evaluation (3-5 days)**
  - **Synthesizing the data (3-5 days)**
- **"Costs"**
  - **"Setting up the evaluation" cost is fixed**
  - **"Performing the evaluation" is an incremental cost, e.g., adding participants and sessions**
    - **Note: Guidance on how to determine the number of participants is provided**
  - **"Synthesizing the data" has both a fixed and an incremental component**
    - **Increasing the number of participants or sessions increases the amount of data to be synthesized and adds some complexity over the fixed cost**

Steps for Conducting an In-House Usability Evaluation

Setting up the Evaluation

- Define goals for the user interface
- *Optional*: Complete the detailed tool
- Decide on the time/depth desired
- Select benchmark tasks
- Choose the participants
- Determine the number of evaluation sessions
- Determine the desired number of participants
- Schedule the participants
- Provide the tool to participants in advance and set expectations for the test
- Select a note-taker/observer
- Select a quiet location for the evaluation

Performing the Evaluation

- Four Stages
  - Describe the system capabilities briefly to the participants
  - Participants explore the interface with benchmark tasks while talking aloud
  - Participants evaluate the system using the high-level tool
  - Wrap up with a post-evaluation discussion
- Observer *transcribes* the session.
  - Should already be familiar with the system
  - Does not help participants to complete tasks
  - Does not analyze/interpret data in real-time

Synthesizing the Data

- Collate notes across all evaluation sessions
- Pre-process the data
- Identify specific difficulties encountered
- Look for relationships between difficulties
- Prioritize problem severity
- Write feedback report:

Iterate

NOTE: It may not be possible to uncover all usability issues in one session.

Setting up the Evaluation (1 of 5)

- ***Define goals for the user interface***
  - Understand what tasks can be performed, and how they are accessed
  - Examples
    - For a chart application, users should be able to bring up a specific chart
    - For a document application, users should be able to bring up a specific document
- ***Optional*: Complete the detailed tool**
- ***Decide on the time/depth desired***
  - Allow approximately 2 hours for each application
    - Less mature applications may have fewer functions and take less time to evaluate
  - Ideally, the evaluation should last no more than 4 hours
    - Participating in an evaluation is hard work
    - Allow time for breaks
    - Finding and scheduling participants for short evaluations will be easier than for long evaluations
  - FAA regulatory evaluations last approximately 2-4 hours

Setting up the Evaluation (2 of 5)

- ***Select benchmark tasks***
  - Tasks help participants learn the system
  - Tasks should be common functions that the users will want to accomplish
    - Tasks could include open issues identified by the detailed tool
  - Tasks should be specified at an appropriate level, with a beginning state and a desired goal.

    <u>Example</u>: For a chart application, an appropriate task might be "Find and display the airport diagram for Boston Logan (KBOS)"
  - Task descriptions can steer the users towards using particular functions (e.g., zoom to the upper left corner of the screen in only one step)
  - User deviations will help identify where the user interface structure is non-intuitive or inefficient.
  - A task is *over-specified* if it provides specific steps for accomplishing the goal, such as those in a user manual.

    <u>Example</u>: Select the airport from the Airport Menu and choose Boston Logan (KBOS).
  - A task is *under-specified* if it does not provide users with a clear goal.

    <u>Example</u>: Use the chart application as you would during an actual flight.

Setting up the Evaluation (3 of 5)

- ***Choose the participants***
  - Ideally, <u>participants should be representative of the target user population</u>
    - For manufacturer evaluations, the target user could the end user (i.e., air transport or general aviation pilots)
    - If the target user population is NOT available, then select participants who have knowledge that is representative of the target user population
      - e.g., If EFB is designed for air transport pilots, participants could be former air transport pilots
  - <u>Participants should not be too familiar with the device.</u> If participants know how to use the device well, user interface problems are not likely to be found.
    - Use of the EFB without training is important as it highlights how intuitive the device is to use. This is important in high workload conditions, when pilots may not remember a prescribed sequence of steps.
    - For example, icon meanings may not be intuitive to the user. The meaning may be taught in training, but could be forgotten in high workload situations.
  - Human factors background is helpful but not necessary

Setting up the Evaluation (4 of 5)

- ***Determine number of evaluation sessions***
  - Typically, 3 to 5 evaluation sessions provide comprehensive results
    - Diminishing returns for additional evaluation sessions
- ***Determine the number of participants***
  - Determine if you will use one or more participants per session
    - Consider how will the device be used, e.g., by one or two person crews?
    - Having more than one participant in a session changes the dynamics
      - Advantage for two or three participants
        - » Creates more opportunity for discussion and exploration
        - » Better for collecting large amounts of subjective data in a short time
      - Disadvantages
        - » Participants' interaction with each other as they complete a task may be unrealistic. Is collaboration the norm?
        - » Data collected are negotiated agreements that could hide underlying confusion or disagreements
    - Having more than 3 participants is not recommended
      - Too hard for all participants to get involved in the evaluation
    - Because an FAA regulatory evaluation includes multiple evaluators, we used teams of 2 participants each for each evaluation session.

April 15, 2004 18

Setting up the Evaluation (5 of 5)

- ***Provide the usability assessment tool to participants in advance and set expectations for the test***
  - Familiarity with the tool and expectations for the evaluation helps to limit digressions, helps participants shift mindsets between items, and reduces the amount of time spent reading during the evaluation
- ***Select a note-taker/observer***
  - A dedicated note-taker/observer allows participants to focus on the device without the burden of keeping their own notes
  - Observer notes capture participants' difficulties in completing tasks, solutions attempted, discussions, and comments
  - Ideally, observer should be more familiar with the system than participants, so that s/he can more easily follow the participants' discussion
  - Observer Role
    - Encourage participants to talk aloud
    - Write down what the participants say
    - Do *not* help participants complete the task, unless the participants can not complete it on their own after several unsuccessful attempts
  - Use the same observer across different evaluation sessions for consistency
- ***Select a quiet location for the evaluation***

April 15, 2004 19

Performing the Evaluation (1 of 2)

NOTE: Usability data can be collected using either a task-based exploration and/or tool-based review. We have participants do both in our research, but an evaluation could be performed using only one of these methods.

Four Evaluation Stages

- ***Provide a system overview***
  - System overview to familiarize participants with the purpose of the device and available functions
- ***Explore the interface with benchmark tasks***
  - Participants complete benchmark tasks to learn the system
  - Participants talk aloud, stating expectations and rationale for steps tried and difficulties encountered
  - Observer transcribes participants' comments. The observer can ask for clarification and examples, if necessary
  - Quantitative data could be collected, e.g., for each task, record number of steps, accuracy, time to complete the task
  - NOTE: Some training may be provided, but it is important that the user be allowed to explore the system on their own as well

April 15, 2004 20

Performing the Evaluation (2 of 2)

- ***Evaluate the system with the high-level tool***
  - Participants consider each topic and provide comments and examples
  - Comments can be positive or negative; these comments can be used to assess severity of a problem
  - NOTE: In our tests, evaluators tried using both the high-level tool and the detailed tool. The detailed tool could not be completed in the time available, so we do not recommended using it when under time constraints.
- ***Wrap-up/post-evaluation discussion***
  - Can be used to obtain feedback from the evaluator, e.g., through a post-evaluation questionnaire

Synthesizing the Data (1 of 3)

NOTE: Note-taker/observer should lead synthesis effort

- ***Collate notes across all evaluation sessions***
  - Notes from the task based exploration and tool based review can be analyzed together or independently, depending on the time available
    - Notes from the tool-based review may be an easier starting point than notes from the task-based exploration.
- ***Pre-process the data***
  - Delete incomplete thoughts and repeated comments
- ***Identify specific difficulties encountered***
  - Examples of where users became confused or inefficient are critical for demonstrating to system designers that the issue is real, not just a matter of opinion, and for illustrating more general problems that may be found elsewhere
- ***Look for relationships between difficulties***
  - Look across tool headings to consider problems by user interface components
  - Can reveal root causes for difficulties that initially seem unrelated

Synthesizing the Data (2 of 3)

- ***Prioritize problem severity***
  - Severity can be examined by addressing these questions:
    - Does it violate FAA regulations and/or guidance?
    - Is the problem systemic (found throughout the system) or local (self-contained)?
    - What is the performance impact?
    - How common is the problem between participants? Do only a few participants have trouble, or do most participants have trouble?
      - If many participants encounter the same problem, a better design may be needed.
    - Is the problem persistent within a participant? Do participants learn to ignore the problem, or do they have trouble throughout the session?
      - Some problems may exist only for first-time users, but the intuitiveness of the interface impacts training time.

Synthesizing the Data (3 of 3)

- ***Write feedback report***
  - Process
    - Writing the feedback report can be an iterative process
    - Relationships between problems and problem severity may become more clear as the feedback is drafted and revised
  - Content
    - Include an overview of the evaluation protocol and purpose
    - Feedback should be actionable and specific based on objective data
      - Group topics by user interface components
      - Provide general statement about the difficulty encountered
      - Provide specific examples and, where possible, quantitative results (e.g., 4 of 6 participants had this problem)
      - Where appropriate, include suggestions for design changes to address the issue

Overview

- Introduction
  - Volpe EFB Research Overview
  - Disclaimers
- EFB Assessment Tools
  - Detailed Tool
  - High-level Tool
- Steps for Conducting a Usability Evaluation
  - Setting up the evaluation
  - Performing the evaluation
  - Synthesizing the data
- **Conclusion**
  - **Benefits to the FAA and Manufacturers**
  - **References**

Benefits

- **Provides assistance in considering human factors topics**
- **For the FAA**
  - **Tool and procedure could be used to conduct structured and comprehensive EFB usability evaluations**
- **For manufacturers**
  - **Tools provide a relatively inexpensive way to identify problems early in the design process and track progress on addressing these problems**
  - **Tools and procedure could be used in advance of a regulatory inspection to anticipate results**

References

- Chandra, D.C., Yeh, M., Riley, V., (Draft of February 2004). Designing and testing a tool for evaluating Electronic Flight Bags. Submitted to HCI Aero Conference, to be held in September, 2004. Will be available this summer at Volpe EFB website (below).
- Chandra, D., Yeh, M., Riley, V., & Mangold, S.J. (2003). *Human factors considerations in the design and evaluation of Electronic Flight Bags (EFBs), Version 2*. DOT-VNTSC-FAA-03-07 and DOT/FAA/AR-03/67. USDOT Volpe Center: Cambridge, MA. Available at Volpe EFB website (below)..
- Federal Aviation Administration, Advisory Circular AC 120-76A, March 17, 2003. *Guidelines for the certification, airworthiness, and operational approval of electronic flight bag computing devices.*
- Nielsen, J. and Mack, R.L. (1994). *Usability Inspection Methods*. John Wiley and Sons, Inc.: New York, NY

- Volpe EFB research project website: http://www.volpe.dot.gov/opsad/efb

Appendix F:
Generic Task List for Evaluations

This generic list of tasks below is offered as a starting point for EFB usability evaluations.

General Tasks

1. Examine the hardware prior to use (physical form factor, display quality, arrangement of controls, etc.)
   - View the display at off-axis angles.
   - Can any of the controls be activated inadvertently?
   - Consider night-time operations. Try adjusting screen brightness.
2. Without clicking on anything, what looks selectable? Try this for the main page of each software function.
3. Review the icons. What do you think the functions are?
4. Provide general comments about the layout of the information.
   - Is any information you would expect missing or in a different place?
   - Are a common set of controls and graphical elements used?
5. If multiple applications are supported:
   - Switch back and forth between two applications.
6. If audio is available:
   - Turn the audio on/off.
   - Can you change the volume?
7. Is the data valid for the current date? What is the valid date rage for the EFB data?
8. Try shutting down and starting up the system. Provide comments on the process.

Electronic documents

1. View the table of contents. Provide comments on its design.
2. Move around to different sections of the documents. Can you get to a pre-selected section quickly, easily, and accurately? Is it easy to remember where you are within the larger document at all times?
3. Go to a long section of the document. Can you tell where you are within the section? Within the document?

4. If figures are available, go to a page that contains figures. Bring up a figure.
   - Try any figure-related functions (e.g., zoom, rotate, view thumbnail, open/close)
5. If tables are available, go to a page that contains a table. Is the information in the table as readable as it would be on paper?
6. Provide comments regarding the readability of the display and interaction with the document. Can you adjust the text to increase readability?
7. If search functionality is available,
   - Search for a word in all documents.
   - Search for a word within a subset of documents (if available).
   - Search for multiple words:
     - Find documents that contain *at least one* of the words.
     - Find documents that contain *all* the words.
8. If there is a list of pages that were viewed previously, comment on:
   - Organization of pages in the list
   - Ease of selecting pages from the list

Flight Performance Calculations

1. Is it clear what data needs to be entered and what data is computed by the system?
2. Try to enter inappropriate data. Are data entry errors caught, and are the error messages clear?

Electronic Charts

1. View all available charts for any airport.
2. Pick a chart to display. Display that chart.
3. Pick a section of the chart, or specific item on the chart. View that section/item.
4. Provide general comments on the legibility and accessibility of the information on the chart.
5. If zoom functionality is available:
   - Zoom in and out of the chart.
   - If there is more than one way to zoom, try the different methods.
6. If panning functionality is available:
   - Pan around the display to see different segments of the chart.
   - If there is more than one way to pan, try the different methods (e.g., use a finger or stylus on a touch screen, or use a touch pad, or keyboard equivalent)

7. If display elements can be added/removed as you zoom in/out of the chart:
    - do you have any comments about how and when display elements are presented?
8. If search functionality is available:
    - Search for an airport. View all the charts that are available. Display a chart.
9. If there is a way to pre-select a list of charts for easy access, comment on:
    - Ease of adding/deleting charts to the list
    - Intuitiveness of accessing and working with charts on the list versus accessing/working with charts that are not on the list

Appendix G:
Sample Observer Notes and Evaluation Feedback

One example of feedback we have provided to manufacturers is shown below. The feedback includes an overview of the evaluation protocol and purpose, and a table of contents, which provides an overview of the core problems to be discussed. The core problems were assigned high, medium, or low priorities. High priority problems were those that either (a) violated known FAA regulations and/or policy, or (b) were global and, in our opinion, had a potentially significant performance impact. Low priority problems were areas we felt could use improvement, but did not appear to have a significant performance impact. The bulk of issues were neither high nor low priority, and so were given a default label of " medium" priority.

Specific details and examples about the EFB and the application evaluated have been deleted.

G.1 Sample Observer Notes

Notes collected while three teams used the high level tool during an EFB evaluation are provided below as an example. Lines are used to separate comments from the different teams in the observer notes. Specific details and examples about the EFB that was evaluated have been deleted. This is raw data below. The important feature to note is how well the observer captures the conversational aspect of the evaluation; she does not insert any interpretation or analysis.

Workload

Pretty user friendly, in general. Workload will be fine. There will be a period of time for familiarization with the system. On submenu – some would give you details. Rather than take you to another screen, pop up a window instead so that you don't lose your place. If you select something, and it takes you to a different screen, then you might be off your path.

With workload – this is way too much work for the benefit of the system. It doesn't map intuitively to the mental map that the crew has. Didn't think it adds a lot more to the procedure than paper.

---

this is clearly head-down. Can't timeshare with this.

It's clear this isn't prioritized for any emergency or time-critical situation.

Hardware Considerations

When you select something, there seems to be very little space between the options. This could be a potential problem. Could select something you don't intend to select.

Hardware display – glare, lighting issues

Accessibility of controls – no issues.

Feedback – on poor side until we turned up audio.

What does it mean when one is highlighted white and the other blue? Could not use hands because the active areas are small

Resolution is ok, angle issues, particularly when trying to distinguish grays from whites. Screen is probably too bright for night use. Probably display issues for the cockpit

Need to get used to punching it well because buttons are small. If you're in a hurry, older guy with glass, or big hands, might take longer.

Nice size, reasonably lightweight. Wouldn't want it on his lap, but if it were on a docking station, …

It’s not clear what the best way to hold it is.

----

form factor is good for sitting in a room. Visibility is angularly sensitive; not clear how it would work if it was fixed. Clearly have glare issues.

Very few external controls. Did try the system on the right – primarily used touch screen.

Readability issues/glare. Would need to use in direct sunlight.

Feedback seems very weak – whether it took the number, whether the number you gave it is inconsistent.

Hadn’t anticipated all the different ways I wanted to go back and forth between sections. Fairly easy to ascertain what it thought was current though I couldn’t tell what had been saved.

Software Considerations

General

Consistency of left hand menu is very good. Use of color ties in well for navigation feature.

Consistency of layout and how information is separated is simple and effective and not distracting from its purpose.

Color coding – didn’t like the use of red. Used red for things that weren’t critical. Might have been important but not for the flight.

The layout is ok.

Thought that if she couldn’t select something it would beep at her. But it didn’t. Clicked on departure – why did it beep? Audio feedback is inconsistent

No audio feedback for errors → ties into workload.

----

inconsistencies with the use of gray

might be an inconsistency in the use of color – e.g., yellow. Not consistent with AC or other colors on the flight deck.

Audio – didn’t know what the beeps mean

Highlighting of knowing what page you’re on is done well.

The system doesn’t appear to be structured in a way I would have structured it, in that it has a lot of separate pages that don’t appear to need to be separate pages.

Audio not very powerful. In real airline operation, if someone is talking to me – wouldn’t want a beep. Would take a click but that’s a personal preference. Other things are beeping at me.

No tactile characteristics.

They have not taken advantage of the space that they have in making fonts larger and readable.

Feedback on system state – it's a little subtle. Use of red and amber not consistent with cautions and warning, in several places in which fields are blatantly red or amber independent of what is in them.

Biggest potential for errors is the erasing from pages.

Formatting/Layout

Information seemed really small. Liked the font. Sans serif so there's not a lot of clutter. Liked that category titles were bolded; showed hierarchy of information. One page where for labels, they did shadow text – didn't like that.

Grouping – (things looked like they were grouped when they weren't)

Font small but ok for his eyes. If supposed to be on lap, could read ok.

Use of rules good for grouping things but don't know what groups are.

Don't think titles of the boxes and title of pages should be in bold.

Numbers should be lined up consistently.

Inconsistencies with the use of gray and yellow – at the very least, the use of yellow for save violates some stereotypes of what yellow is. If important enough to be in red, then should be able to click on it and see what the problem is.

List down makes sense from logical, organizational point of view.

Readable and usable in room but concerns that print is small and fine. Would be concerned about turbulence.

Colors don't appear to be very contrast driven. Contrast not optimized for all conditions/max readability.

Spacing is a little tight but ok

Arrangement is fair – would link more information/group information as appropriate.

Symbols and Icons

Initial problems but as you go through, it made more sense.

Legibility – be careful with font size

Good that they don't use a whole lot of icons. Issue regarding red vs. yellow.

When correlated in green, is it redundant coding, or is it signaling that two different systems are ok.

Need some kind of indication to highlight required fields.

Not many icons. Mainly just buttons. There are some slight inconsistencies in the use of color.

Interaction

Thought number of steps was pretty good, with one exception. Otherwise, it was pretty easy to navigate within the different pages.

When moving between pages, previous page should be saved automatically.

Didn’t recognize abbreviation at first.

Could determine what all the abbreviations are.

----

moving between the fields is pretty straightforward and easy to do.

Number of inputs to complete the task – didn’t like the fact that you had to re-input information for the page after you put the information in (and left without saving)

Weakest feedback is when you’re inputting things. It’s unclear whether the input took. If you make a mistake, it’s not particularly easy to back up. Have a tendency to hit clear which brings you back to zero.

It’s not always good at telling you when it’s working or not. Am I just waiting for it to execute or am I waiting for it to take my input? For example, when trying to change pages. If I ask for something that it can’t do, it just doesn’t do anything. So it doesn’t tell me that I can’t do it.

Error handling and prevention

Didn’t get buttons.

Big error was when inputs were deleted. Error recovery painful.

Inadvertent selection of table headings.

Cancel really is a back button.

Error recovery – not a graceful recovery from errors in entering time. Why do you have to click ok for an error? Might not need a separate message in a separate window.

Easy to redo

----

lack of positive feedback, likelihood that you’ll erase a whole page – did not anticipate uninterrupted tasks very well.

Recommendation: enter for the field, automatic save.

Multiple Applications

unclear how much information is flowing across the pages. Haven't really thought about someone working multiple pages at same time

Automation

Could have more automation, e.g., figuring out what page is done.

----

few things where it was automatically going to lists. Not a lot of automation.

----

would have been nice to have autocomplete

G.2 Sample Feedback Introduction

Notes

1) Background

This research is being performed under the sponsorship of the Federal Aviation Administration (Human Factors Research and Engineering Division, AAR-100, Tom McCloy Program Manager). Our goal is to develop a tool for FAA evaluators to use when evaluation EFBs from a human factors perspective. Given the nature of the FAA evaluation, we focus on the weaknesses of the interface not the strengths.

2) Terms

User-interface (UI) **priorities** are assigned with respect to the usability of the device. We conducted a human factors evaluation only. Our feedback represents only the opinions of the project team; it does not represent the views of the DOT or FAA.

We use the term "**participant**" to refer to the people who participated in our study. The participants were not typical end users (line pilots), and they were not necessarily typical FAA evaluators either. However, they *used* the device to complete specific goals (just like an end user would) and they *evaluated* the device using our tools, which are being designed for the FAA.

Terms such as "**we**" and "**our**" below refer to the opinion of the human factors experts who led the tests.

Occasionally, we use the term "**you**" to refer to a generic system user.

Sometimes we use the term "**Observation**. In this case we are merely stating what we saw, without attaching any priority to it. These are items that we thought you would want to know about, but that did not have a clear human factors impact.

3) Platform

[A description of the hardware and software that were tested would be included.]

Overview of Protocol (3 to 4 hours total)

1) Intro (15 min)
   a. Project context and goals (i.e., develop a tool for FAA evaluations)
   b. Brief introduction to the application (not real "training," just context on the application and intended use)
2) Task-Based Exploration using Task List (next page) (1 to 1.5 hour)

   The tasks were designed to have a beginning state, and a desired goal. They were open-ended enough so that users could digress for a while, but specific enough so that participants knew when they had successfully completed the task.
3) Evaluation using High-Level Tool (up to 1 hour)

   This tool is a list of topics on which the evaluators provided comments. The list was just one page long and included items such as "Formatting/Layout," "Symbols and Graphical Icons," "Interaction: Accessing Functions and Options," and "Error Handling and Prevention." The tool was modified between test sessions. We will send you the latest version in a few weeks for comment.
4) Evaluation using Detailed Tool (up to 1 hour)

   The detailed tool is a reformatted version of the information in the full EFB human factors document (Version 2, Appendix B). We used only the information from the General EFB System chapter. This tool was much more specific to the EFB. It was also much longer (5 pages) and is more detailed than the other tool.
5) Wrap-up (survey and discussion comparing tools) (15 min)

Participants

Team 1: Two university graduate students. Neither is a pilot. They are working on air traffic control and rail human factors projects.

Team 2: One senior aviation-human-factors researcher (not a pilot), and one pilot/researcher (not in human factors). The pilot has air transport and military experience (15+ years).

Team 3: One senior aviation-human-factors researcher and faculty member at a university. He is an instrument-rated general aviation pilot with approximately 6000 hrs, and is highly familiarity with air transport operations. His partner cancelled at the last minute, so we ran him alone.

Team 4: One senior aviation systems engineer and one air-traffic specialist/researcher from a local university. Both are instrument-rated general aviation pilots. Neither specializes in human factors, but both have worked on project teams with human factors specialists.

Task List

[A copy of the customized task list for the test would be provided.]

G.3 Sample Feedback Overview

High Priority

1 Use of Red

Medium Priority

2 Audio

3 Check Boxes

4 Icons

5 Responsiveness

6 Text

7 Errors and Recovery

Lowest Priority

8 Abbreviations, Language, and Terminology

No assigned Priority

9 Use of Yellow

G.4 Sample Feedback

High Priority

1 Use of Red

- Red should be used only to indicate warning-level conditions. The use of red indicated to participants that there was a problem that needed to be addressed *immediately*. Inappropriate uses of red included:
  - *[Examples provided here]*

Suggestions

- Reconsider whether red is necessary. Use red only to denote situations where immediate pilot action is required.

Medium Priority

2 Audio

- Audio is used inconsistently. There was no clear pattern for when a beep was generated or not. The audio beep suggested that something was wrong and needed to be addressed, but the participants could not determine what they needed to do.

Suggestion

- Audio should be disabled unless clearly conveying an important meaning.

3 Check Boxes

- Check boxes should be used to toggle the state of a variable (on/off). We found cases where they were used differently, which caused confusion. For example:
  - *[Examples provided here]*
  - *[Observations]*

Suggestions

- *[Suggestions provided here]*

4 Icons

- Our participants had trouble understanding the meanings of the icons and distinguishing between them, creating a distraction.
  - *[Examples provided here]*

Suggestion

- *[Suggestions provided here]*

5 Responsiveness

- Our participants found the system was sometimes slow. System slowdowns were especially noticeable when changing pages.

Suggestion

- Provide a more obvious (and accurate) system busy indicator (e.g., clock-style cursor).

6 Text

- Participants had various difficulties in reading the text on the screen accurately. In some cases, it was because the font size was relatively small, or because the font it used a relatively thin stroke width.
    - Font size

      The font size was considered acceptable, but smaller than desirable. Also, participants observed that there often was space to use a larger font size.

Suggestion

- Consider using a larger font size where possible. In many cases, a larger font size could be accommodated without changing the screen layout.

7 Errors and Recovery

- We noted several points at which our participants tended to make errors. In some cases, recovery from the error involved multiple steps, such as reentering data that was lost. For example:
    - Size of input areas

      Participants noted that the size of the input fields were small and could result in inadvertent activation.
    - Lack of confirmation
        - The Cancel button does not ask users for confirmation. Participants were sometimes surprised when their entries were discarded and the previous page was presented.

Suggestion

- The size of the input field is somewhat related to the small font size chosen. A larger font would be more readable, and would create larger active regions.
- Confirmation of an action should be requested if data is going to be discarded.

<u>Low Priority Items</u>

8 Abbreviations, Language, and Terminology

- Some instances of unclear language were noted.
    - *[Examples provided here]*

Suggestion

- *[Suggestions provided here]*

<u>No Assigned Priority</u>

9 Use of Yellow

- Yellow should be used to indicate caution-level conditions only. In general, yellow is used in the flight deck to indicate the potential need for pilot action. The use of yellow software attracted participant's attention, but one participant commented that it was not always clear that the items presented in yellow were of that level of importance.
    - *[Examples provided here]*
- Participants noted some instances in which colors were used inconsistently.

Suggestion

- Reconsider use of yellow in general.