

## **A Tool for Structured Evaluation of Electronic Flight Bag Usability**

Divya Chandra, U.S. DOT Volpe Center, Cambridge, Massachusetts

Thursday, March 27, 2014

**National Transportation Library preservation note:** The information in this document was present on the Bureau of Transportation Statistics (BTS) website from March 27, 2014, through approximately, December 16, 2016 at URL <http://www.volpe.dot.gov/safety-management-and-human-factors/aviation-human-factors/tool-structured-evaluation-electronic>. However, this URL is no longer functional

Please note that while the information gathered below may not be displayed in the same way as the original webpages, the intellectual content of those pages is captured and preserved for future use. Minor edits may have introduced to correct errors, or improve navigation or readability.

These preservation actions were taken on May 11, 2021. Information and tables were harvested from the Internet Archive's Wayback Machine, at

<https://web.archive.org/web/20140424170040/https://www.volpe.dot.gov/safety-management-and-human-factors/aviation-human-factors/tool-structured-evaluation-electronic>.

NTL thanks the Internet Archive for helping us preserve this important content.

# A Tool for Structured Evaluation of Electronic Flight Bag Usability

*Divya Chandra, U.S. DOT Volpe Center, Cambridge, Massachusetts*

## **Abstract**

Electronic Flight Bags (EFBs) are coming into the flight deck, bringing with them a host of human factors challenges. The first step in addressing these challenges was to identify and prioritize them. Good progress has been made on that front by Chandra and Mangold, whose comprehensive document is in active use by industry and the FAA today [1]. Unfortunately, using this document is a daunting task because of its breadth and depth. Our next goal is to develop and test a tool based on the full document that can be used for periodic structured assessments of EFB usability. We expect that this assessment tool will benefit designers, operators, and regulators by providing a structure for EFB human-factors evaluations. Both EFB-specific issues and general user interface topics are covered. The purpose of this report is to document the progress to date on constructing this usability-assessment tool for EFBs. We cover how the tool was developed and tested, what it looks like to date, and how it could be used to help assess and track EFB usability. Further testing is planned to ensure that the tool is usable and to ensure that it adds value to the evaluation process.

## **Introduction and Background**

Electronic Flight Bags (EFBs) are becoming a reality. Several airlines are either actively flying with EFBs, or looking into their options [2-5]. Some of the functions envisioned for EFBs include electronic documents, flight performance calculations, cabin surveillance, surface moving map displays, and display of weather information. Many other proposed functions for EFBs are mentioned in appendices of the recently issued EFB Advisory Circular (AC) 120-76A [6]. This AC provides guidance on a streamlined field approval process for certain types of EFBs, an attractive option for operators looking into deploying EFBs.

There are now a range of EFB solutions on the market, from laptop computers to fully installed devices such as the Astronautics Pilot Information Display offered by Boeing as a forward-fit option on some aircraft [5, 7]. In between are a range of devices to suit various flight decks and budgets. Some units are portable and can be used inside and outside the flight deck (e.g., ADR FG3600™ or the CMC CT-1000G). These portable devices, which tend to be favored by high-end general aviation operators, are essentially personal computers that run both flight-related software (e.g., JeppView FliteDeck®) and standard desktop software, such as Internet browsers and word processors. Other EFBs are more specifically designed for use in the flight deck (e.g., the mounted Teledyne Controls/Spirent AaVantage® and the tethered Universal Avionics UCD®); these units are designed with the air transport market in mind.

Some time ago, industry and the FAA recognized that human factors concerns would play a key role in the design of EFBs. While EFBs may look like familiar equipment, from a flight deck perspective, they are new and sophisticated devices. For example, they have graphical user interfaces and they can support multiple new functions, such as electronic charts and documents, that could impact established flight deck procedures. With all these capabilities, EFBs could well

play a central role in the future of flight deck information management [8]. In the future, EFBs may develop uses that we cannot even foresee today.

In 1999, the Volpe Center was tasked with identifying and prioritizing EFB human factors considerations in support of the draft EFB AC. Working with the FAA and the Air Transport Association Digital Data Working Group, Volpe produced a lengthy document in September 2000 on EFB human factors considerations [1]. This report is referenced in the EFB AC, along with other documents that show the FAA's commitment to ensuring that human factors issues are addressed in the evaluation of new flight deck devices [9, 10]. In addition, topics from the Volpe document that were considered especially important are brought into the main text of the EFB AC [6].

Chandra and Mangold [1] contains an extensive list of human factors topics that EFB designers and evaluators need to consider. Its format was crafted for use by system developers, operators, and evaluators who are not necessarily human-factors specialists. The document has been distributed widely and is actively in use by EFB developers and customers. It contains guidance on system considerations that could apply to any EFB as well as three specific functions: electronic documents, electronic checklists, and flight performance calculations. The purpose of this document is not to tell designers how to build an EFB, but rather to help them make informed choices. Established user interface design principles are described, recommendations are made, tradeoffs are described, and sources for more information are referenced. The document is discussed in more detail elsewhere [11].

While Chandra & Mangold [1] is valued as a comprehensive and readable reference, using it to keep track of human factors issues that arise during EFB development and regulatory evaluations is a daunting task because of its breadth and depth. The original document contains nearly 80 topics, and an updated version, due out soon, will contain approximately 100 topics. In order to ensure that it can be used effectively by evaluators and designers, our next goal is to develop and test a tool based on the full document that can be used for periodic structured assessments of EFB usability.

While FAA field evaluators are the primary intended audience for the tool, we expect that it will benefit designers and operators as well. For example, designers and operators could use the tool in internal reviews to anticipate and resolve human factors issues even before going through a formal regulatory evaluation. In addition, the tool could help evaluators conduct more structured, thorough, and predictable regulatory evaluations. Feedback to the manufacturer from these evaluations would be more specific as well. Using the tool may also help designers and evaluators to focus their discussions, ideally leading to quicker resolution of human factors issues. Finally, when used by experienced staff, the tool can help both designers and evaluators see general patterns in user interface difficulties, not just isolated problems. This deeper understanding should help everyone to design more usable EFBs, the ultimate goal.

The purpose of this report is to document the progress to date on developing a tool for structuring EFB usability evaluations. Initial plans for this project were described by Chandra in an earlier paper [12]. Since then, we have made significant progress in refining both the process for developing the tool, the tool itself, and a methodology for testing the tool. In the sections

below, we cover recent progress on developing and testing the tool, what it looks like to date, and how it could be used by manufacturers to assess and track EFB human factors considerations. We conclude with our plans for further testing of the tool.

## **Developing and Testing the Tool**

In developing the EFB usability assessment tool, we actually started with three sources of information. One, of course, was Chandra and Mangold [1]. Another source was our knowledge of the general principles of user interface design, which are well documented [13, 14], and our expertise in human factors. In addition, we reviewed earlier checklist tools that were designed for assessing usability within an aviation context [15, 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 is designed for use by a variety of non-human factors experts such as FAA field evaluators (e.g., engineers or test pilots), EFB system developers (e.g., software engineers), and even EFB operators (e.g., airline personnel). Note, however, that some experience and familiarity with the tool and the original document [1] is necessary to get the full benefit. While there is some accommodation for new or infrequent users who may need more explanation of terms and topics, the intended user is already familiar with the topics. Also, note that human factors experts are not excluded from using the tool. They too may find it helpful for structuring their reviews, as long as it matches reasonably well with their internal model of what to look for in a user interface; if it does not match, the human factors expert may find the tool's structure helpful, but somewhat constraining.

The biggest problem we faced in developing the tool was the need for the evaluations to be brief (i.e., under four hours). There is a necessary tradeoff between the time spent on the evaluation and the depth of the review; more time will produce a more thorough review (until evaluators are fatigued or saturated, of course). For example, every item in Chandra and Mangold [1] could be assessed, page-by-page-but that is not a viable option in general. Instead, we tried to develop a tool that could be used in different ways depending on the available time. To minimize the evaluation time, only compliance with core issues could be examined, or, if time is available, a detailed review could be performed to identify both compliance issues and areas for optimization or improvement. Results from the detailed evaluation could help the manufacturer understand usability issues in order to reduce certification risk and optimize training time, workload, likelihood of errors, etc.

Creating the tool was just the first step in this effort, however. Testing the tool with representative users, typical systems, and realistic methods was also an important step in ensuring that the tool is usable and adds value to the process. In order to do this step, we sought volunteer aviation/human-factors experts to participate in evaluations of EFBs that vendors volunteered for the test. The tool described in this report is the product of evaluations of two different EFB systems. The evaluations were conducted in an office setting, similar to a bench test that could be conducted by a manufacturer or regulator. We expect to do at least one more test prior to releasing the full tool.

Chandra and Mangold [1] is a good source for information on general principles of user interface design, but it does not cover methods for evaluating EFB usability. Fortunately, we were able to

draw upon knowledge from the usability engineering community, whose focus is to assess and improve the usability of more common devices and applications, such as cell phones and websites [17].

While some standard usability evaluation methods were not suitable for the EFB evaluations we envisioned, we were able to modify and customize other methods for our purpose. Because of the time limit, we did not pursue formal usability testing in which representative system users are observed and their performance is recorded and analyzed (e.g., errors are counted, times to complete tasks are measured, or the number of input steps are recorded and compared against optimal scenarios). Instead, we focused on the co-discovery technique and expert reviews (also known as heuristic evaluations). Co-discovery is an informal usability inspection method in which two individuals examine a user interface collaboratively. We selected this protocol because it is simple, meets our time constraints, and fits well with the process already in use by the FAA. Through discussion, the evaluators stimulate each other to form observations and insights regarding user interface issues that they might not have arrived at working independently. Our expert reviews were just that: usability, aviation, and human factors experts reviewing the system independently. After the individual expert reviews were completed, we compared notes and synthesized our collective observations.

A team of two evaluators worked together to complete the evaluation. The protocol consisted of three phases. First, evaluators were given a brief demonstration of the unit's capabilities and functions by an experimenter who played the role of the manufacturer. Next, participants explored the system as a team while a facilitator guided them through a set of benchmark tasks and took notes about any areas of difficulties they encountered. Finally, the participants used the tools to complete their evaluation and provided feedback to the experimenters about the tools.

### **Tool Description**

In a generic sense, assessment tools need just two components: items and a rating scale. Often, these tools also suggest where to go for further information and provide an area for written notes. When there are many items to review (the usual situation), items must be formatted and grouped with care. In addition, to provide real value, the wording of each individual item must be brief but clear and the rating scheme must be both clear and appropriate for the task and the user's purpose.

Because we were interested in having a tool that could be used for different levels of assessments, we designed the tool in layers. One layer can be completed relatively quickly because there are few items and the ratings are coarse. To complete a detailed analysis, we include items that are more specific to EFBs and ratings scales of finer resolution. Evaluators can choose to mix the high level and detailed items at their discretion, customizing the tool for their purposes. The items and ratings scales are discussed further below, and samples from the tool are presented. We also present our thoughts on acquiring evaluator comments, and how these can be gathered in a general sense, and in more structured ways.

## Items

There are two sets of items in our tool, one brief and one lengthy. The brief set is designed for a high-level assessment of the system, and the long set is designed for a detailed review. The topics in the brief set represent general dimensions along which the quality of the user interface can be assessed. Topics in the long set are gathered from Chandra and Mangold [1].

Items from the first set are shown in Table 1. A variety of topics is included in order to ensure a systematic and thorough review. Note that some topics overlap with others. For example, movement between pages and number of inputs to complete a task could be related to each other if completing the task requires the user to move between pages. The advantage of overlapping topics is that difficulties with the user interface are less likely to be missed. Plus, there is little cost to having an issue appear under multiple headings; redundancies are easily reconciled. The exact titles for the items in Table 1 are still evolving as we get feedback from potential users, but the basic list is mature. Note that this list is generic and it could potentially be applied to other systems beyond EFBs.

- Physical ease of use
- Visual, audio, and tactile characteristics
- Movement between pages
- Number of inputs to complete a task
- Ease of accessing functions and options
- Manipulating data/content (e.g. panning)
- Susceptibility to error
- Error recovery
- Graphical icons/symbols
- Formatting and layout
- Use of color
- Language, terms, abbreviations
- Feedback (system state, alerts, modes, etc.)
- Labels and controls (hardware and on-screen)
- Responsiveness
- Automation (if any), amount and integration

**Table 1:** General usability assessment items.

Updated: Thursday, March 27, 2014