Development of a Usability Evaluation Framework for the flight deck

Victoria A. Banks, Katherine L. Plant & Neville A. Stanton

Transportation Research Group, University of Southampton, UK

ABSTRACT

Systems design is often criticised for bringing human factors (HF) expertise into the design process at the end of a product's development – often too late to have much impact on the design and usability of products and/or systems. This paper proposes a new Usability Evaluation Framework for the Flight Deck that can utilise HF expertise throughout the design lifecycle. It incorporates widely accepted design practices with a more user-centred approach that enables simultaneous usability testing and evaluation at every stage of the design lifecycle.

KEYWORDS

Usability, evaluation, end user, flight deck

Introduction

Human factors (HF) input can add the most value to a product early in design when it is still relatively inexpensive to modify mock-ups and prototypes (Stanton et al. 2013). The application of HF methodologies can offer researchers, engineers and designers with a structured framework in which to analyse and evaluate practical problems. However, HF is often bought in towards the end of the design lifecycle which runs the risk of considerable expense and redesign requirements. Figure 1 presents a schematic diagram of the basic design lifecycle. Within each of these phases, it is possible to involve HF methodologies and principles to varying degrees.

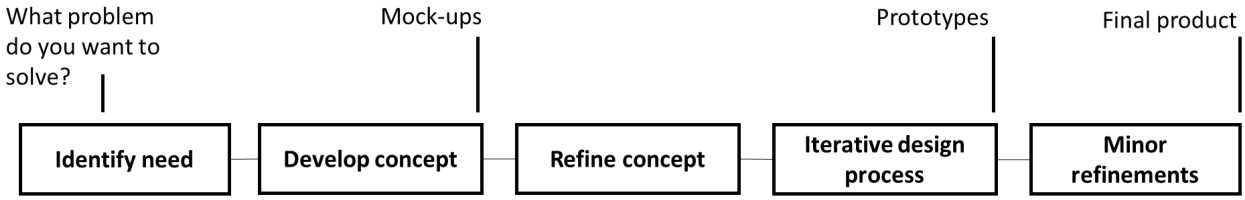

Figure 1: The basic application of HF to the design lifecycle (Adapted from Stanton et al. 2013)

In aviation, the Federal Aviation Administration (FAA) has published numerous design guidelines, considerations and standards relating to HF throughout the design process. These mainly result in the publication of technical documents and guidance materials that can be used by systems designers and engineers in pursuit of certification (e.g. Mejdal et al. 2001; Yeh et al. 2013; SAE ARP5056, ARP4754). These guidance materials typically align closely to usability principles rather than providing any specific criteria in which systems can be assessed. Consequently, the efficiency of the design process may be hindered by not detecting and recognising usability issues early on. This can lead to costly design modifications later on or the release of products that lack 'user

friendliness'; and in terms of usability assessment on the flight deck there is limited guidance actually available.

One notable exception is the National Aeronautics and Space Administration (NASA) who published a 'crew-centred flight deck design philosophy' in 1995 for high speed civil transport aircraft (Palmer et al. 1995). NASA sought to provide a set of guiding design principles that could help focus attention on the flight crew throughout the design process by developing a framework that could be used by engineers and researchers (see Figure 2). Whilst NASA's framework did not aim to represent the accepted design process within any particular organisation, it was intended to be descriptive of generally accepted design practice in flight deck design. However, despite this, it was not widely used. It also omits clear user requirements in its high-level description. This paper argues that user requirements are just as important as technical and functional requirements because failure to meet the needs and expectations of a target user group will impact upon the success of the product/system (Shackel, 1984, 1991; Nielsen, 1993). It is at this point that analysts should start to think about the development of context-specific usability criteria in which products/systems can be assessed. However, no formal criteria currently exist. So we propose a new Usability Evaluation Framework (UEF) that seeks to emphasise the role of HF within the design lifecycle.

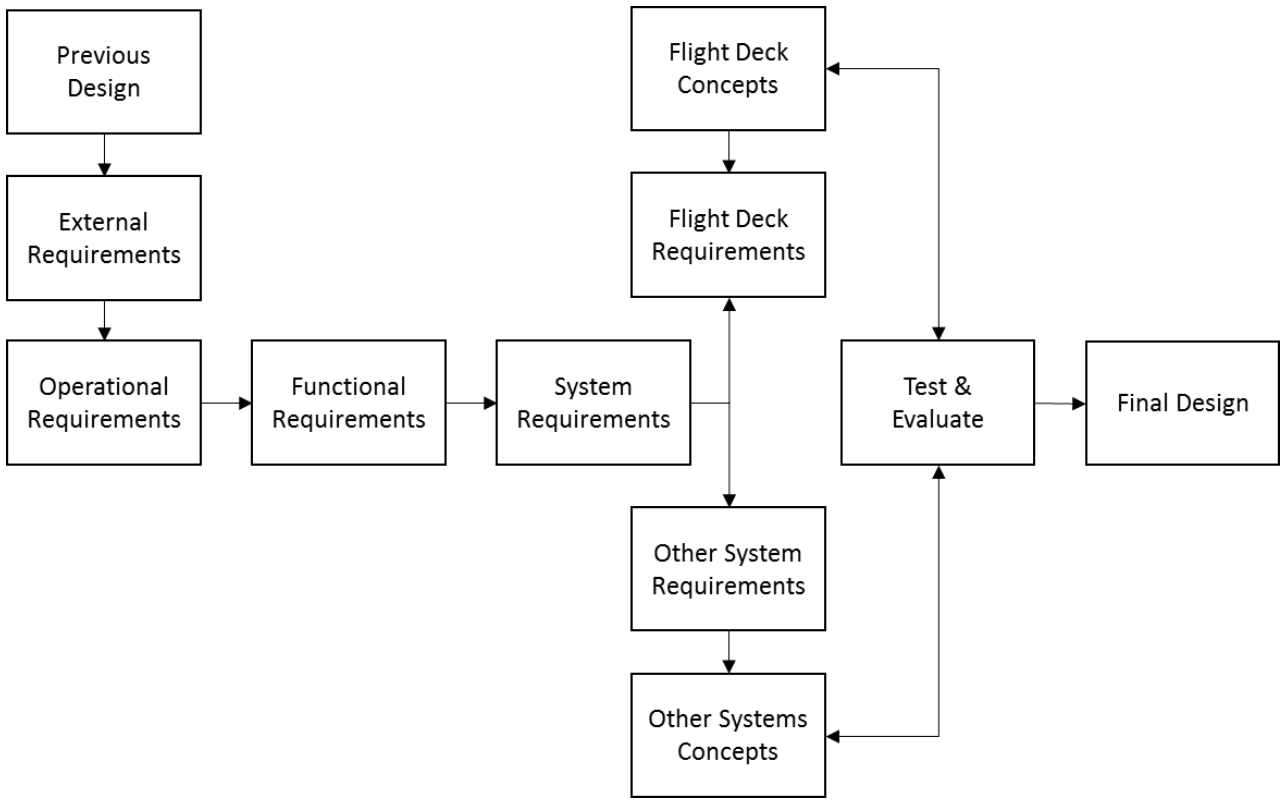

Figure 2:. Crew-centred design framework (Adapted from Palmer et al. 1995)

Development of a Usability Evaluation Framework

The need for user-centred design within aviation originates from the implementation of automated avionic systems. Automation has transformed the way in which pilots both operate and interact with the aircraft (Stanton & Marsden, 1996; Harris, 2011). The complexity of these systems can also make it difficult for flight crews to be fully aware of all the intricacies of operation that could impact upon the safety of the flight (Palmer et al. 1995). As such, the human-machine interface (HMI) becomes the bridge between the human operator and automated subsystems. Modern day

'glass cockpits' present flight instruments on electrical screens (HMI) that are interlinked with the Flight Management System (FMS) and Autopilot functions. Billings (1997) suggests that glass cockpits could include up to six interfaces, backup flight instruments and critical systems indicators on the main instrument panel. With so much information available to the flight crew at any one time, the principles of design for 'understandability and usability' for interface design should be utilised (Norman, 1988, 2002). Norman posits that if the user is to understand the system, the mapping and feedback of a system must be able to present an accurate representation of the product/system. Mapping refers to the relationship between the controls and how the user interacts with them whilst feedback is based upon the passage of information between a product/system and the user (Norman, 2002). Of course, there are many other notable contributions relating to usability design within the literature. Shneiderman (1992), for example, developed 'eight golden rules of dialog design' whilst Jordan presented 'ten principles of usable design'.

The proposed Usability Evaluation Framework (UEF) seeks to combine the widely accepted design practices involved in flight deck design (e.g. Palmer et al. 1995; Higgins, 1996; SAE-ARP-5056, 2006) with a more human-centred design approach, much like Harvey & Stanton (2013) who devised a usability framework for in-vehicle information systems. The proposed UEF for Flight Deck Design is presented in Figure 3. This represents the first step in developing a comprehensive usability toolkit for the design of the flight deck. It is envisaged that the final UEF will also incorporate context-specific usability criteria that will help guide the selection of appropriate research methodologies.

The UEF outlines that the initial step in designing a new system is to identify a need, highlight a current problem or create a new idea based upon existing products/needs or gaps in the market. Contextual factors must be considered as this will help identify target user populations, tasks and usability goals. Once a high-level description of a new product/system has been agreed, the envisaged system, and the task in which the system is designed for, needs to be explicitly defined (i.e. through the development of design requirements). The development of u*ser requirements* helps identify who the product/system is for and what they will need to know throughout the completion of the task. In addition to this, user requirements should also encompass physical, social and cognitive aspects of the task.

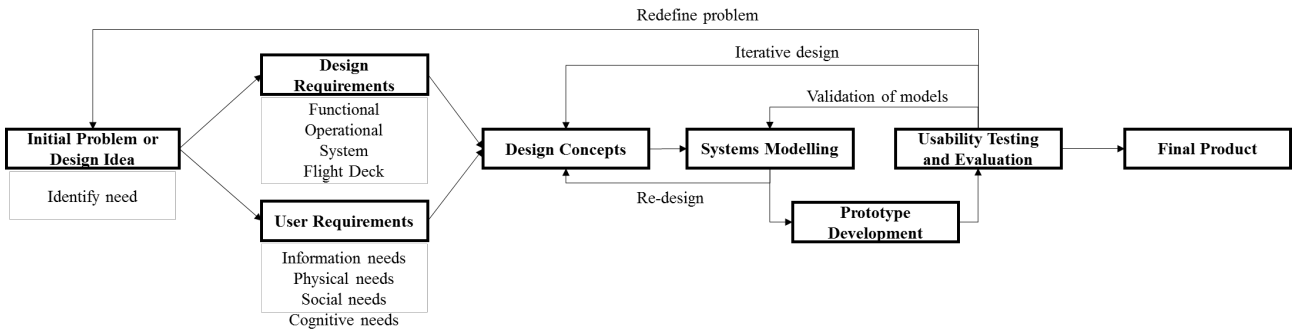

Figure 3: Proposed Usability Evaluation Framework for the Flight Deck

Developing usability criteria

Usability criteria differs from that of usability design principles as the former outlines specific expectations of a product/system whereas design principles are more aligned to best practice principles. Formally, the definition of usability has been outlined by the International Organisation for Standardisation (ISO) 9241 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) Part 11, Guidance on Usability as; "[the] extent to which a product can be used

by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (1998, p2). In this instance, context refers to the specific circumstances in which the product/system will be used. Heaton (1992) suggests that in the context of a particular product, an explicit definition of usability can be developed. It is this definition that can then be used to guide later evaluation. Context-specific usability criteria should be viewed as a means to mitigate risk and improve safety. It also represents a compatible approach to the certification process as they can be used to help demonstrate compliance to industrial standards and processes. For example, CS-25.1302 (European Aviation Safety Agency, 2017) specifically provides guidance and a regulatory basis in addressing design-related flight crew error. It outlines a number of design principles (similar to that of Norman, 1988, 2002; Shneiderman, 1992; Jordan, 1998) that seek to avoid and manage flight crew error (see Table 1). To demonstrate that a product/system complies to these requirements, we must have some way of assessing its performance. Context-specific usability criteria therefore offers a compatible approach. Usability metrics (or Key Performance Indicators) are measurable attributes that can be used to evaluate the success of a product/system. The Federal Aviation Administration identify a number of usability metrics in their online 'Human Factors Awareness Course'. These include training time, time to reach proficiency, error rate, time to complete a task and the number of positive versus negative statements during interview. There are of course likely to be many more product-specific criterion that are omitted from this list. However, to demonstrate how these may be used to show compliance to CS-25.1302, we have matched appropriate usability metrics to the requirements outlined in Table 1.

Table 1: Alignment between design requirements and usability metrics

CS-25.1302 Requirements	**Associated Usability Metrics (taken from FAA Human Factors Awareness Course)**
Controls and displays must be perceived correctly Controls and displays must be comprehended in the context of the task Support the crew in responding to, and performing, a task Be predictable and unambiguous Allow the crew to intervene when appropriate in the task	Training time Time to reach proficiency Number of times "help" is accessed Success vs. failure rate in task completion Time to complete a task Error rate Error recovery time Decision time / delay Positive versus negative statements/responses

The intention of the UEF is to further develop a set of context-specific usability criteria and associated key performance indicators related to flight deck design. This remains a work in progress.

Developing a methods toolkit

In addition to the development of context-specific usability criteria for the flight deck, the UEF will also provide guidance relating to the selection and implementation of various research methods. It is widely accepted within the literature that successful usability evaluations incorporate a mix of both analytical and empirical, quantitative and qualitative research methodologies. The UEF encourages the application of a wide range of research methodologies via its distinct modelling and test phases. For example, Hierarchical Task Analysis (HTA; Annett, 2000) and other process charting methods (e.g. Operator Sequence Diagram; Brooks, 1960) can show how the

product/system under development may impact upon the way in which the human will interact with it and the wider system during the completion of specific tasks (Kirwan & Ainsworth, 1992). HTA is particularly useful in understanding the task as it provides an exhaustive description of goals, sub-goals, operations and plans (Annett, 2004). It often acts as the input to many other research methods including the allocation of function, workload assessment, interface design and evaluation (Stanton et al. 2013). According to Harvey et al. (2011), HTA can be used to assess both the efficiency and effectiveness of systems design. In contrast, Operator Sequence Diagrams (OSD; Brooks, 1960) can provide insight into the allocation of function within a system. They can therefore give an indication of the *expected* pattern of task-user-system interaction. The benefit of systems modelling is that it provides the analyst with a snapshot into the inner workings of a system and offers a means to check the boundaries of the system are complete and appropriate. Depending upon the outcome of such activities, re-design activities may ensue. In the case of positive outcomes, prototype development will continue. This will allow for usability testing and evaluations at varying levels of complexity (i.e. mock-up evaluations, simulator evaluations, test-pilot evaluations) to be conducted during the cycle of iterative design using various methods. Examples of appropriate usability evaluation methods are presented in Table 2. Empirical testing using a representative sample of *actual* end users is strongly encouraged to ensure that a product/system meets the usability criteria that is established earlier on the design lifecycle along with their associated outcomes. Whilst this list is not intended to be exhaustive, it demonstrates the direction in which the UEF will be used. Importantly, whilst the use of test pilots represents a step closer to acknowledging end users within the design process, they represent a highly specialised cohort who complete flight test techniques (i.e. specific manoeuvres) enabling the design to be evaluated. They may not necessarily behave like an average pilot during day-to-day operations. It is therefore important to consider the needs and expectations of non-specialist pilots in the development and evaluation of new systems throughout the duration of the design lifecycle.

Table 2: Possible evaluation methods that can be used throughout the design lifecycle and their associated outcomes

Stage in Lifecycle	**Appropriate Human Factors Methods**	**Outcomes**
Initial problem or Design Idea	Focus Group	Evaluate design, establish user needs, generate new design concepts
	Interview	Real life insight into operational activities
	Observation	Real life insight into operational activities
Design and user requirements	Focus Group	Evaluate design, establish user needs, generate new design concepts
	Interview	Real life insight into operational activities
	Observation	Real life insight into operational activities
Design concepts	Focus Group	Evaluate design, establish user needs, generate new design concepts
Systems Modelling	HTA	Detailed breakdown of the structure of a task
	OSD	Graphical depiction of interaction / allocation of function between agents

	Event Analysis of Systemic Teamwork	Network representations of task, social and information characteristics
	Critical Path Analysis	Predict task times, identify conflicts, interference
	Systematic Human Error Reduction and Prediction Approach	Predict error types, rates, criticality, probability of error, mitigation strategies
Usability Testing and Evaluation utilising end user populations	Observation	Real life insight into operational activities
	Interview	Real life insight into operational activities
	User trial	Evaluate design, generate recommendations to improve design
	Questionnaires (e.g. Systems Usability Scale, Acceptance Scale, Cornell University Questionnaire for Ergonomic Comfort, NASA-TLX)	Subjective ratings of usability, satisfaction, usefulness, ease of use, learnability, workload, physical ergonomic assessment of product / system
	Performance measures	Information relating to user population e.g. number and type of errors made / time taken to complete a task
	Heuristic Evaluation	Estimate performance of system and identification of usability issues
	Layout Analysis	Redesigned menu structures

Discussion

Overall, the final UEF aims to provide a structured approach to flight deck design that may help reduce the risk of system failure from usability-related issues. Along with the UEF, cockpit-specific usability criteria and method selection criteria can be generated. This will result in a comprehensive toolkit that can help designers and engineers acknowledge HF throughout the design lifecycle and equip them with the tools necessary to deliver both usable and desirable products. This is still very much a work in progress and this paper represents the early stages of this process. Even so, the paper points to some very important research developments – specifically the development of cockpit-specific usability criteria that will be used to help guide the selection and application of different research methodologies (both analytical and empirical). This approach could significantly improve the efficiency of the design process because potential issues can be highlighted early (Stanton et al. 2013).

Acknowledgement

This project was co-funded by Innovate UK, the UK's Innovation Agency, with support from the UK Aerospace Technology Institute.

References

Annett, J. (2000). Theoretical and pragmatic influences on task analysis methods. *Cognitive task Analysis*, 25-37

Billings, C. (1997). Aviation automation: The search for a human-centered approach. New Jersey: Erlbaum

Brooks, F A. (1960). Operational sequence diagrams. IEE Transactions on Human Factors in Electronics, 33-34

European Aviation Safety Agency. "Certification Specifications and Acceptable Means of Compliance for Large Aeroplanes CS-25. Amendment 19, 12 May 2017", 2017

Federal Aviation Administration. 'Human Factors Awareness Course'. Available at www.hf.faa.gov/webtraining/Usability/u28_metrics.htm. Accessed 11.12.2017

Harris, D. (2011). Human performance on the flight deck. Ashgate

Harvey, C., & Stanton, N. A. (2013). Usability evaluation for in-vehicle systems. CRC Press

Harvey, C., Stanton, N. A., Pickering, C. A., McDonald, M., & Zheng, P. (2011). A usability evaluation toolkit for in-vehicle information systems (IVISs). *Applied Ergonomics*, 42(4), 563-574

Heaton, N. O. (1992). Defining usability. *Displays*, 13, 147-150

Higgins, J. K. (1996). Boeing's design approach for crew-centred flight decks. In The Future Flight Deck – Safe and User Friendly. London: Royal Aeronautical Society. ISBN 1-85768-037-5

International Organisation for Standardisation. (1998). ISO 9241. Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) Part 11, Guidance on Usability

Jordan, P. W. (1988). Human factors for pleasure in product use. *Applied Ergonomics*, 29, 25-33

Kirwan, B. & Ainsworth, L. K. (1992). A Guide to Task Analysis. London: Taylor & Francis

Mejdal, S., McCauley, M. E., & Beringer, D. B. (2001). Human factors design guidelines for multifunction displays. Available at: www.faa.gov. Accessed 23.08.2017

Nielsen, J. (1993). Usability Engineering. London: Academic Press

Norman D. A. (1988). The psychology of everyday things. Basic Books, New York

Norman, D. A. (2002). The design of everyday things. Basic Books, New York

SAE. (2006). Flight crew interface considerations in the flight deck design process for part 25 aircraft (ARP5056)

SAE. (2010). Guidelines for the Development of Civil Aircraft and Systems (ARP4754).

Shackel, B. (1984). The concept of usability. In *Proceedings of IBM Software and Information Usability Symposium*, Poughkeepsie, NY, 15-18 September 1981, pp1-30

Shackel, B. (1991). Usability-context, framework, definition, design and evaluation. In B. Shackel & S. J. Richardson, Eds., Human Factors for Informatics Usability, pp. 21-37. Cambridge: Cambridge University Press

Shneiderman, B. (1992). Designing the user interface: Strategies for effective human-computer interaction. Addison-Wesley, New York

Stanton, N. A., & Marsden, P. M. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science*, 24(1), 35-49

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). Human factors methods: A practical guide for engineering and design. 2nd Edition. Ashgate

Yeh M., Jo, Y. J., Donovan, C., & Gabree, S. (2013). Human factors consideration in the design and evaluation of flight deck displays and controls. FAA Final Report, Version 1.0. DOT/FAA/TC-13/44, DOT-VNTSC-FAA-13-09, 2013. Available at https://ntl.bts.gov/lib/50000/50700/50760/General_Guidance_Document_Nov_2013_v1.pdf. Accessed 11.09.2017