User-centred design and evaluation of future flight deck technologies

Katie J. Parnell, Rachael A. Wynne, Katherine L. Plant, Neville A. Stanton.

Human Factors Engineering, Transportation Research Group, University of Southampton, UK

ABSTRACT

This work presents an overview of the Human Factors methodologies, and applications thereof, that can be utilised across the design lifecycle of new technologies entering future commercial aircrafts. As advances are made to the architecture of commercial aircraft cockpits, it is vitally important that these new interfaces are safely incorporated and designed in a way that is usable to the pilot. Incorporation of Human Factors is essential to ensuring that engineering developments to avionic systems are integrated such that pilots can maintain safe interactions, while gaining information of value. Taking work from previous research studies, a case study example of technological advancements during its early conceptual stages is presented. This shows how different methods and processes can be applied and combined to ensure that the user is included within the design and evaluation of new flight deck technologies.

KEYWORDS

Aviation; Design lifecycle; Flight Deck

Introduction

Changes to cockpit design can take considerable time to enact due to the lengthy processes involved in ensuring that they are safe and meet certification. It is important that future aircraft can benefit from technological advancement without being limited by time and expense but also, critically, that they uphold safety. A movement towards an 'open flight deck' aims to facilitate innovation within the cockpit to enable a platform that can undergo regular updates of flight deck applications. This will allow the development of new applications that can more effectively present information to pilots in the cockpit, as well as bring in new sources of information that may have not been previously viable. The application of Human Factors (HF) to the design, modelling and evaluation of new applications to this platform is critical to ensure safety and usability (Parnell et al., 2020).

The Chartered Institute for Ergonomics and Human Factors (CIEHF) published a recent white paper on the future of aviation and the importance of HF to all facets of the domain going forwards. This white paper suggests that there will be many changes to the aviation domain across the next 30 years to 2050, with the implementation of Artificial Intelligence and augmentation, yet it emphasises that the human will still play a key role. Thus, rendering the importance of HF practices to advance human-machine relationships in aviation while maintaining safe and high-performing systems (CIEHF, 2020). Included within the white paper was the use of digital assistants within future flight displays to support pilots. There is an emphasis on a key difference between Artificial Intelligence (AI) and Intelligent Assistants (IA). With the implementation of AI into the cockpit predicted to incur complexity, there is a need to utilise assistive technology that can be succinctly understood by pilots. Conveying complexity within AI in a succinct and informative manner is as important as the development of the intelligence itself. This is particularly true in high workload, time critical events which can occur within commercial aviation and during which the pilot must make informed decisions regarding the safety of the flight. In the past, the adverse safety implications of poor cockpit design and a lack of HF awareness has been made evident with events such as the Kegworth Disaster. Analysis into the disaster by Plant and Stanton (2012) highlighted how a mismatch between the cockpit design and the pilots' expectations led to the fatal decision of shutting down the wrong engine of the aircraft.

Integrating HF throughout the design lifecycle of new technologies and interfaces that enter the flight deck stives to avoid further disasters. Following recent accidents with the Boeing 737 MAX (e.g. Wendel, 2019), new bipartisan legislation was recently introduced to reform the way the Federal Aviation Administration (FAA) certifies aircraft. This references an increased need for HF certification to assess the relation between humans and interfaces within the flight deck (Aircraft Safety and Certification Reform Act of 2020). One key approach to achieving this is by involving the end-user within the design process to ensure it matches their expectations and requirements (Norman, 1986; Gould & Lewis, 1985; Kaber et al., 2002, Parnell et al., 2019;2020). The work presented in this paper will demonstrate how HF practices and methods should be applied throughout the design lifecycle of new concepts that are to be implemented into a future open flight deck, intended for commercial aircraft. Utilising user-led practises, it will show how central usability is to effective design and the importance of including users from the beginning, and throughout, the design process. This builds on previous work by the authors, published as individual research studies, to show how they form a wider design lifecyle.

Design Lifecycle

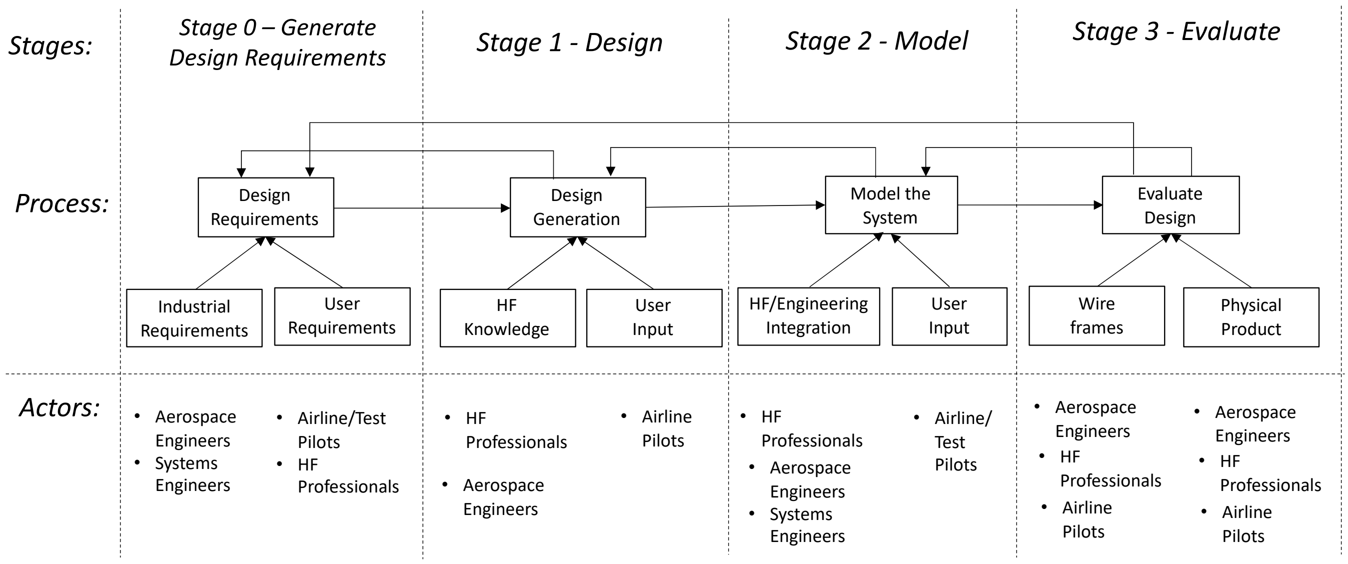


Figure 1. The design lifecycle for the development of flight deck technologies with the inclusion of HF practises and methodologies.

Despite encouragement from the FAA and the recent Whitepaper by CIEHF (2020) on the importance of including HF into the design of flight technologies, to enhance the relationship between human and machine, there are no specific guidelines on how this should occur. The 'Open Flight Deck' project is an aviation project comprised of aviation engineers, manufacturers and HF professionals and academics, seeking to explore the opportunity for an open architecture within flight deck technologies (see https://openflightdeck.co.uk for further details). The importance of designing and implementing new flight deck technologies that are both safe and usable sought to identify what the design lifecycle of new flight deck interfaces should involve when HF is considered throughout. An overview of this design process is shown in Figure 1, as well as the key actors that are required to facilitate and contribute to the work across each of the stages. The central key components of this are 'design', 'model' and 'evaluate'. While in reality there is likely to be much complexity surrounding this process within the wider system of systems that comprises aviation, these three components are considered invaluable within the design process.

Stage 0-Generate Design Requirements: The development of design requirements is stated to be stage 0, as it is important that the specific requirements are in place well before the designs and the design process itself are initiated. At this early stage it is important that a range of actors generate insight into their requirements of the technology. Any conflicting recommendations or visions of the intended design and use of the technology can then be overcome at this early stage which is both cheaper and more effective. Once the initial requirements have been established, they form the template to generate designs from, as well as acting as a reference to determine the success of the design.

Stage 1-Design: During the early stages of the design process multiple different design concepts can be generated in a cost-effect way through drawings and simple digital mock-ups. The Design with Intent (DwI) method (Lockton et al., 2010) has been found to be particularly useful in encouraging blue-sky thinking and encouraging creativity in the initial design stages (Allison & Stanton, 2020; Parnell et al., 2020). Especially when conducted with by aviation engineers and commercial pilots to facilitate a discussion surrounding the design requirements and their actualisations from different perspectives (Parnell et al., 2020). At this stage the designs can be innovative and creative before their technical feasibility and practicality is considered at a later stage.

Stage 2-Model: The designs must be modelled to ascertain their viability and the possibility for integration within the system. There are different modelling techniques available to review different facets of the design, such as the possibility for error using the System Human Error Reduction and Prediction Approach (SHERPA; Embrey, 1986) or predicted time of interactions using critical path analysis (Lockyer, 1984; Baber & Mellor, 2001). Alternative HF modelling techniques can model the wider system within which a technology will be placed, such as operator event sequence diagrams (OESD; Brooks, 1960; Kurke, 1961).

Stage 3-Evaluate: Once designs have been generated and then modelled to determine their feasibility, they require evaluation (Stanton et al., 2013; Stanton et al., 2014). Thiis occurs across two stages. The initial wire-frame evaluation involves drawings of the design it its basic form, which is time and cost saving. Successful designs from this evaluation process are then selected for physical reproduction and integration in a flight simulator with a replicated flight deck for physical testing. Reliable and robust evaluation occurring in this second stage, is required to ensure that the designs are usable and safe before they are considered for integration into future flight decks. This requires a realistic flight simulator and user testing with the intended user base, commercial airline pilots of ranging experiences and backgrounds. User testing must be designed to include a baseline condition and allow sufficient power to be drawn from the data to generate conclusions and enable statistical significance to be obtained.

It is unlikely that the process of designing a new technology will go through these three stages in a single linear fashion. The inclusion of multiple feedback loops across these stages (as indicated in

Figure 1) are intended to enhance the final design and allow for refinement in response to challenges that may arise.

Case Study

Having outlined the design lifecycle, an example is now presented to demonstrate how the Open Flight Deck project has applied and utilised it to facilitate the inclusion of the user and HF principles for the benefit of the interface and its use. This case study presents a novel technological feature that aims to provide enhanced information to the pilot on the status of the engine where problems may arise. Of specific interest was an engine oil starvation avoidance technology under development by an aerospace manufacturer that could provide advanced warning to pilots if oil was leaking from the engine. This section incorporates previous work conducted in Parnell et al., (2019) and Parnell et al., (2020), to suggest how they combine to inform the design process while also highlighting future intended work that has been delayed due to the Covid-19 pandemic.

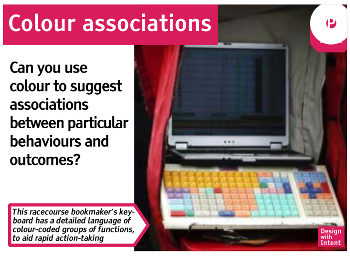
Oil leaks are a rare, but not unheard of, event on commercial aircrafts that do pose a significant threat to safety if they are not managed optimally (e.g. Australian Transport Safety Bureau [ATSB], 2012, 2017). Currently, pilots of commercial aircraft receive a low oil pressure warning only once the pressure levels have reached minimum limits. This reduces the options that the pilots have available to preserve and maintain the safety of the flight. The pilot must determine the validity of the message and then assess the severity of a suspected oil leak before taking the necessary action. Previous events have shown that pilots choose to either throttle the engine back to try and conserve the oil or shut the engine down, leading to a flight diversion (ATSB, 2012; 2017). This has considerable knock-on implications to the airline, passengers, and maintenance crews.

The current, late-stage alert does not allow pilots to be proactive and put in place effective mitigation strategies. Pilots have to do extra work to first ensure that the information is accurate and then calculate the options available to them at that moment in time. By presenting more up-to-date information, that can accurately and reliably inform the pilot of the engine status, it is thought that incidents and subsequent implications can be avoided. The subsequent section will now present the case study of the development of the oil starvation avoidance technology within the design lifecycle.

Stage 0-Generate Design Requirements: An aerospace manufacturer and their comprising aerospace and systems engineers had the requirement for what the system needed to achieve in its basic format (e.g. provide enhanced information to the pilot on the status of the engine when exposed to the oil leak). During our previous work conducting a series of structured interviews (see Parnell et al., 2019), pilots were asked to comment on their use of the current system and provide suggestions on how they may want to experience this information, in a more advanced future system. Furthermore, utilising HF practises, a SHERPA analysis was conducted to determine what the current opportunity for error was in the current system (Parnell et al, 2020). This information could then be fed back to aerospace manufacturers and engineers to allow them to elaborate on the utility of the future technology in mitigating opportunity for error.

Stage 1-Design: The first step in the development of the new technology was through design workshops conducted with airline pilots, aerospace engineers and HF professionals. This utilised the DwI method (Lockton et al., 2010) to encourage innovation while providing relevant prompts for creativity and consideration of different design principles. Table 1 provides some examples of how these principles were incorporated into the pilots' designs. For example, using colour which is already used in the cockpit to represent the oil levels, and building this into the dials and text given in relation to engine oil level.

Table 1. Example DwI cards and their design inspirations (taken from Parnell et al, 2019).

Lens	**Card**	**Design recommendation**	**Illustration**
Perceptual		Use green/amber/red for consistency. Code numbers on the oil gauge that show rate of oil loss.	
Interaction		Provide a count down timer to show how much time there is until oil starvation in the engine.	

A review of the concepts generated in comparison to the original outputs from the SHERPA showed where the recommendations can be applied. For example, real time oil parameter trending to gain up-to-date information on the oil status (Parnell et al, 2020). It also showed that involving pilots in the design process can facilitate rich detail into the concept designs with many more recommendations coming from the DwI than the initial SHERPA, that focused only on improving the original system. Interestingly, there were also cases where the detail from the workshops conflicted with the recommendations made in the SHERPA. For example, the SHERPA suggested that the potential to automate the detail regarding suitable diversion and landing options based on the available information in order to reduce pilots' workload and thus limit the opportunity for error. However, the workshops highlighted that they would not want a system to make the decisions for them or to lead them, as they are trained to make these decisions for themselves and like to have this level of control over the system. This illustrates the point made in the introduction reported in the CHIEF (2020) white paper with an important distinction between AI and IA. Pilots prefer a system that can assist them in making decisions rather than one that makes the decision on their behalf as they feel too removed from the system (Parnell et al., 2020). This aside, there were more similarities than dissimilarities between the design concepts generated in the workshops and the SHERPA recommendations. Four designs were taken forward to the next stage, evaluation. These designs bore some similarities on their placement and integration with the flight, yet there were differences in the imagery of the designs. For example, some involved graphs while others presented graphics or count down timers. Proving that there were multiple different design options that can be considered outside of those currently used on the flight deck.

Stage 2-Model: Taking forward the different design concepts that were generated from the DwI workshops involves the modelling of the wider system and its interaction with it. OESDs were chosen as a method to model the system surrounding the implementation of an oil starvation avoidance technology. This was because they allow a full systems overview with interactions demonstrated to capture the interactional nature of system (Harris et al., 2015). They also provide a joint understanding between systems engineer and HF professionals to convey system requirements. Two OESDs were developed, one that mapped the current system and one that conveyed what a future intended system could convey and how it would interact with other elements. From the OESDs, operational loading could be calculated to understand the frequency of operations in the system. The future system was found to have more operational loading, with more interactions between elements. This was representative of the increase information that was being provided in comparison to current practise. The diagrams highlighted the need for simplicity in the system between the interactions. Interactions that were defined in the OESDs were also feedback into the requirements and the design proposals.

Stage 3-Evaluate: The two stages of the evaluation were conducted; heuristic evaluation of the four wire frames by HF experts and then user testing in a flight simulator with airline pilots. The wire frames of the designs were developed by an aviation system architect. These were then presented to seven HF experts, along with the size dimensions of the displays and screens for ergonomic review. The heuristics used were those generated by a previous aviation project, deemed to be relevant to this technological feature. They included screen layout, navigation, alphanumeric, abbreviations, menus, undo/reversal, mimics, grouping, colour, message design and coding. As well as rating the concepts on the heuristics, participants also provided additional recommendations where they saw fit. There was one clear ‘winner’ from the heuristic evaluation, yet further recommendations were also made to enhance the design. For example, simplifying the information as much as possible to reduce clutter and not reproduce information unnecessarily. From this, the best design was generated taking into consideration the recommendations from each of the previous stages of the design. The design was developed for integration into a flight simulator for user testing with airline pilots.

A flight simulator is to be used that has dual-pilot configuration, requiring two participant commercial pilots per session, interacting with the simulation and each other. Unfortunately, due to the Covid-19 pandemic the full testing schedule has not been yet been completed, so the full results of the simulator testing cannot be reported. We are, however, hoping to resume testing in the next three months, with necessary provisions to ensure the safety of the participants. We hope to comment further on these results closer to CIEHF conference. A between-subjects methodology will be used, to avoid priming bias, with participants experiencing an oil-starvation event either in the current system or with the avoidance technology present. Participants follow standard take-off procedures and once in cruise, an oil leak will occur. They will then have to manage this and their actions and comments will be captured. The critical decision method (CDM; Klein et al., 1989) will be applied in a post-trial interview to capture decision making. Workload (Hart & Staveland, 1988), usability (Brooke, 1986) and the acceptance scale (van der Laan et al., 1997) will also be used to capture subjective ratings on the flight deck displays.

This process presents the starting point to facilitate design conception and review, where the integration of HF and the involvement of the user hold real value in shaping usable and safe design, as involvement at a later stage can lead to costly and risky issues (Stanton & Young, 2003; Stanton et al., 2014; Parnell et al., 2020). Yet, user testing in the simulator does not amount to full validation of the design, as further established testing in line with the FAA certification process is required before certification and integration can take place. The integration of such a feature into a wider ‘open’ cockpit system also needs to be carefully reviewed. As the cockpit functions within the wider system of systems in aviation, changes to one area will impact on others. There may be inconsistencies, redundancies and confusion that can arise from novel device integration. Training procedures are imperative to ensure effective uptake of such features.

Conclusion

There is increasing demand and importance placed on the appropriate application of HF practises in the development and implementation of future technological devices within the cockpit (Wendel, 2019; Aircraft Safety and Certification Reform Act of 2020; CIEHF, 2020). Collaboration between aerospace manufactures and engineers, systems engineers, pilots and HF practitioners has sought to develop a HF design process which can be utilised within the early stages of interface design. This paper has presented the proposed HF design lifecycle with support for its use, illustrated with a case study that brings together previous work by the authors to detail the wider design life cycle. Following the process, and its feedback loops, has been valuable in generating informative discussions between different actors who hold different perspectives on the design of aviation technologies. Sometimes these may be conflicting, yet these methods show these conflicts may be identified and challenged early on in the design process, before they become larger issues later on. It shows how HF practices can, and should, be integrated into the design process from the very start. Inclusion of the end-user can generate huge benefits to the development of usable designs that can be integrated alongside their current practises. Further work is still required to complete testing and review the next steps for further validation and certification.

References

Allison CK, Stanton NA (2020a) Ideation using the "Design with Intent" toolkit: a case study applying a design toolkit to support creativity in developing vehicle interfaces for fuel-efficient driving. *Applied Ergonomics*, *84*:103026

Australian Transport Safety Bureau . (2012). Engine oil leaks VH-OQG and VH-OQC en route Singapore to London, United Kingdom, 24 February and 3 November 2011. Retrieved from https://www.atsb.gov.au/media/3545749/ao-2011-034_final.pdf

Australian Transport Safety Bureau . (2017). Engine malfunction and in-flight shutdown involving Boeing 777 A6-EGA. Retrieved from https://www.atsb.gov.au/media/5772230/ao-2016-113-final.pdf

Baber, C. and Mellor, B., 2001. Using critical path analysis to model multimodal human-computer interaction. *International Journal of Human-Computer Studies*, 54 (4), 613–636.

Brooke, J. (1996). SUS: a "quick and dirty' usability. *Usability evaluation in industry*, 189.

Brooks, F. A. (1960). Operational sequence diagrams. IRE Transactions on Human Factors in Electronics, 1, 33-34.

Charter Institute for Ergonomics and Human Factors (2020) The Human Dimension in Tomorrows Aviation System [White Paper] https://www.ergonomics.org.uk/common/Uploaded%20files/Publications/CIEHF-Future-of-Aviation-White-Paper.pdf

Embrey, D. E. (1986). SHERPA: A systematic human error reduction and prediction approach. In *Proceedings of the international topical meeting on advances in human factors in nuclear power systems*. Knoxville, TE.

Gould, J. D., & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM*, *28*(3), 300-311.

Harris, D., Stanton, N. A., & Starr, A. (2015). Spot the difference: Operational event sequence diagrams as a formal method for work allocation in the development of single-pilot operations for commercial aircraft. *Ergonomics, 58*(11), 1773-1791.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology 52*, 139-183. North-Holland.

Kaber, D. B., Riley, J. M., & Tan, K. W. (2002). Improved usability of aviation automation through direct manipulation and graphical user interface design. *The International Journal of Aviation Psychology*, *12*(2), 153-178.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(3), 462-472.

Kurke, M. I. (1961). Operational sequence diagrams in system design. *Human Factors*, *3*, 66-73.

Lockton D, Harrison D, Stanton NA (2010) The design with intent method: a design tool for influencing user behaviour. *Applied Ergonomics 41*(3):382–392

Lockyer, K., (1984). Critical path analysis and other project network techniques. 4th ed. London, UK: Pitman.

Norman, D. A. (1986). Cognitive engineering In User Centered System Design, DA Norman and SW Draper, Eds. Erlbaum, Hillsdale, NJ, 31, 62.

Parnell, K. J., Banks, V. A., Plant, K. L., Griffin, T. G., Beecroft, P., & Stanton, N. A. (2019). Predicting design-induced error on the flight deck: an aircraft engine oil leak scenario. *Human factors*, 0018720819872900.

Parnell, K. J., Banks, V. A., Allison, C. K., Plant, K. L., Beecroft, P., & Stanton, N. A. (2020). Designing flight deck applications: combining insight from end-users and ergonomists. *Cognition, Technology & Work*, 1-13.

Plant, K. L., & Stanton, N. A. (2012). Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model. *Safety Science, 50*(2), 300-315.

Stanton, N., Salmon, P. M., & Rafferty, L. A. (2013). Human factors methods: a practical guide for engineering and design. Ashgate Publishing, Ltd..

Stanton, N. A., & Young, M. S. (2003). Giving ergonomics away? The application of ergonomics methods by novices. *Applied Ergonomics*, *34*(5), 479-490.

Stanton, N. A., Young, M. S. and Harvey, C. (2014) A Guide to Methodology in Ergonomics: Designing for Human Use (second edition). CRC Press: London, UK.

Van Der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies, 5*(1), 1-10.

Wendel, W. B. (2019). Technological Solutions to Human Error and How They Can Kill You: Understanding the Boeing 737 Max Products Liability Litigation. *Journal of Air Law. & Commerce, 84*, 379.