A novel Human Machine Interaction (HMI) design/evaluation approach supporting the advancement of improved automation concepts to enhance flight safety

Joan Cahill & Tiziana C. Callari
Centre for Innovative Human Systems (CIHS), School of Psychology, Trinity College Dublin, Ireland

Abstract

This paper presents a novel Human Machine Interaction (HMI) design/evaluation methodology, supporting the specification and evaluation of a new adaptive automation concept, both from a functional and an operational/safety perspective. This methodology has been advanced as part of the work requirements for the Applying Pilot Models for Safety Aircraft (A-PiMod) project, funded by the European Commission. Critically, this methodology integrates/combines formal HMI design/evaluation approaches (i.e. user interviews and simulator evaluation) with an integrated stakeholder approach to evaluation. The objective of this paper is to highlight (1) what is new in this overall approach (i.e. integration of formal HMI approaches such as simulator evaluation with stakeholder evaluation approaches, decomposition of project goals to project objectives, evaluation objectives and key performance indicators); (2) what is new in the specific stakeholder approach to evaluation (i.e. the set-up of a Community of Practice involving both internal and external stakeholders, and the integration of this methodology with wider HMI evaluation activities); and (3), what the methodology delivers in terms of ensuring improved levels of safety and reliability for the aviation sector. The evaluation of this methodology will be based on an analysis of project outcomes to date.

Introduction


Several human factors problems have been documented in relation to automation design. This includes: automation surprises (i.e. the crew does not understand what automation is [or is not] doing), workload concerns (i.e. whether or not automation actually increases workload in certain situations, given that the crew have to track the status/actions of automation, and/or lack of workload support in high workload situations), and issues pertaining to over-reliance on automation (i.e. potential that over reliance on automation might have a negative impact on pilot flight
management skills/competencies, overall impact on expertise etc.). In addition, certain aspects of automation design require more detailed consideration. Currently, the dynamic task allocation between the crew and automation is based on an assessment of the aircraft state (i.e. aircraft systems only). Indeed, automation is not really aware of the crew and at times, it acts forcefully. In this sense, new automation concepts must address the issues of teamwork (i.e. how to support cooperation/teamwork, what aspects of crew state to consider and how to distribute workload/tasks between the crew and automation). Also, some key questions concerning automation and the role of the pilot have not been fully addressed (i.e. level of authority in relation to key flight management tasks and vetoing automation). These questions can also be posited from an automation perspective (i.e. can or should automation veto the pilot’s decision?). In this regard, existing automation systems have built in ‘protections’ to ensure that the aircraft remains in a safe state. This mainly concerns abnormal ‘safety critical’ situations. Critically, the design of an improved automation system should support pilot task performance – and specifically, address the issues as outlined above.

This paper reports on a new Human Machine Interaction (HMI) design/evaluation methodology supporting the specification and evaluation of a new adaptive automation concept, both from a functional and an operational/safety perspective. This methodology has been advanced as part of the work requirements for the Applying Pilot Models for Safety Aircraft (A-PiMod) project, funded by the European Commission. First, a short introduction to the A-PiMod project and the Researcher’s role in the project (i.e. Human Factors research team from Trinity College Dublin) is provided. Existing HMI design/evaluation methods are then reviewed. Following this, the proposed novel HMI methodology is presented. An overview of the specific validation activities designed and implemented to date is then reported. Following this, the main outcomes and project achievements are reviewed. The benefits and application of this approach is then discussed. Finally, some conclusions are drawn.

The Applying Pilot Models for Safety Aircraft (A-PiMod) project

The A-PiMod project aims to address certain still-open automation problems, as outlined above. The high level goal of the project is to improve flight safety in a time of increasing levels of performance, automation and information provision to the flight deck. Specifically, the objective of the A-PiMod Project is to design a new adaptive automation concept based on a hybrid of three elements – (1) Multi-Modal Pilot Interaction, (2) Operator Modeling, and (3) Real-Time Risk Assessment. Three impact statements have been defined to assess the expected project outcomes: (1) to reduce accident rate by 80%; (2) to achieve a substantial improvement in the elimination of and recovery from human error; (3) to mitigate the consequences of survivable accidents.

The high level objective of our research in this project is to validate the A-PiMod concepts and technologies from a (1) functional and, (2) operational/safety perspective. This spans requirements specification/validation, prototype design and evaluation, and the final evaluation of safety/operational impact. To do so, a novel
a novel approach to operational/safety evaluation of automation technology

methodology has been proposed to support the specification and evaluation of the new adaptive automation concept. This is discussed in a later section.

Overview of existing HMI approaches to evaluation

The HMI literature defines a range of formal and informal methods for the design of human friendly technology adopting a ‘User-Centered Design’ methodology (Cooper, 2007; Preece, Rogers, & Sharp, 2007; Constantine & Lockwood, 1999; Hackos & Reddish, 1998). The specific approaches adopted reflect underlying theoretical assumptions about design practice. In particular, they represent diverse views concerning the role of end users, the specific process for envisioning new technology requirements, and the relationships between design and evaluation.

Formal HMI Design/Evaluation Methods

Typically, formal HMI methods start with analysing the existing task (Preece et al., 2007). To this end, a task analysis is first undertaken, involving the participation of end users. Structured or semi-structured interviews are used to understand and evaluate current work practices and supporting technology requirements (Hackos & Redish, 1998). Several analysis steps are then undertaken without the participation of end users. Analysis outputs include lists of end users, user and task matrices and task workflow diagrams. This is followed by different design activities such as storyboarding and prototyping. Once the prototype is developed, users are involved in different evaluation activities. In this way, design and evaluation are conceived as separate steps.

Informal HMI Design/Evaluation Methods

Formal HMI methods have been the subject of much debate in the HCI literature. Specific challenges have come from the fields of Ethnography and Participatory Design. Ethnographers argue that classical HCI methods do not take work practice seriously; failing to address the social aspects of work (Hutchins 1995; Vicente 1999). Participatory design theorists have questioned the separation between design and evaluation in formal methods (Bødker & Buur, 2002). Specifically, they have challenged the instructiveness of traditional user and task analysis outputs for design guidance. Central to Participatory Design theory is the idea that Usability Engineers design ‘with’ end users, as opposed to ‘for’ them. Accordingly, users are active participants in the design process (Bannon & Bødker, 1991, Bødker & Grønbæk, 1996). Several techniques are outlined in the literature. This includes concept generation, envisionment exercises, scenario role playing, story collecting and storytelling (through text, photography and drama), and the co-creation and evaluation of prototypes.

Operational Validation/Evaluation approaches

Arguably, existing HMI design/evaluation methods fail to address the broader operational issues underpinning the envisionment and specification of new technologies. Operational assessment involves more than the assessment of operator performance (i.e. in relation to task workflows, workload and situation awareness),
and the allied performance of the proposed system (i.e. usability of the proposed system/user interface). Crucially, wider ‘operational’ issues must be considered. This includes the fit between the technologies and the proposed operational scenarios, the specification of operational requirements (at a process as well as a task level), the assessment of operational benefits, the design of future operational processes/procedures, the specification of teamwork/co-ordination and information sharing requirements across relevant system actors, and the identification of potential implementation barriers.

**Stakeholder approaches to evaluation**

The involvement of stakeholders as part of programme/project evaluation has received increasing attention over the past three decades (Rodriguez-Campos, 2011). Overall these approaches follow from the idea that collaboration must tackle issues that matter and have impact/benefits for the stakeholder’s organization/domain of activity. Further, such collaboration requires a high level of interpersonal and organizational trust. Central to this, is the establishment of communication and discussion methods/sessions. The use of knowledge generation and tacit knowledge elicitation methods are favoured in these approaches. These methods promote ways to transfer users’ tacit knowledge as a source of sustainable competitive advantage. Stakeholder evaluation approaches do not necessarily involve technology design/evaluation. For example, such approaches have been applied to the evaluation of processes, the delivery of services, events, architecture, the layout of cities and relevant social spaces (i.e. parks/playgrounds), and so forth.

**The novel HMI Design/Evaluation Methodology adopted in A-PiMod**

**Introduction to Research**

The validation activities will address the following key issues pertaining to automation design:

- The design of the cockpit as a co-operative system (i.e. Pilot/automation co-ordination/teamwork, distribution of task activity between the crew and automation);
- Pilot comprehension of automation (i.e. status of automation, who is responsible for what task and what are they doing) and the avoidance of automation surprises
- How automation might be designed to enable workload management and reduce crew stress in high workload and potentially safety critical situations;
- How the A-PiMod concept enables/supports crew briefing/planning, situation assessment, information management and decision making (linking to Crew Resource Management concepts);
- How the A-PiMod concept enables/supports error identification and recovery.

Overall, the evaluation approach involves two strands of activity – (1) research with the A-PiMod Community of Practice, and (2) formal simulator evaluation. Collectively, this research can be characterized in relation to two key features - (1) early design/evaluation and (2) iterative design/evaluation.
Validation activities in A-PiMod are designed to be both early and iterative. Validation occurs after the initial specification of requirements elicitation and review (milestone 1), and then at two key milestones in project (milestone 2 and milestone 3). The first round of simulator evaluations (i.e. validation cycle 1/milestone 2) are designed to be explorative (i.e. using low fidelity prototypes), while the second round (i.e. validation cycle 2/milestone 3) will involve a full scenario run (i.e. using high fidelity prototypes). Also, there is on-going validation with internal and external stakeholders. Further, there will be a final evaluation of the overall system in relation to the overall safety/operational impact (i.e. milestone 4). For a graphical illustration of this, please see Figure 1 below.

Figure 1. Validation timeline/activities in A-PiMod.

The methodology proposed in A-PiMod integrates/combines formal and informal HMI design/evaluation approaches, along with an integrated stakeholder approach to evaluation. Overall this is an iterative process and links to the documentation of functional/technical requirements and associated prototyping activities.

The following sections will outline what is new and/or innovative in the proposed methodology, in relation to the following perspectives:

(1) What is new in the overall validation approach;
(2) What is new in the specific stakeholder approach to evaluation;
(3) What the methodology delivers in terms of ensuring improved levels of safety and reliability for the aviation sector.
New: The overall validation approach

The validation process in A-PiMod will support the assessment of how far the proposed technologies achieve the project goals and outcomes. It is underpinned by (1) User-Centred Design concepts and methods, and (2) the principle that safety is and operational concept. In determining the project evaluation objectives and questions, a hierarchical decomposition has been followed to ensure that validation activities are focussed on project outcomes and provides traceability. This process starts from the A-PiMod project goals/objectives, through to evaluation objectives (high level and detailed), evaluation questions (high level and detailed, and key performance indicators (KPI).

The proposed methodology integrates formal HMI approaches (such as simulator evaluation), with informal/participatory HMI methods (such as collaborative prototyping), along with tacit knowledge elicitation methods (such as semi-structured interviews following specific techniques – i.e. the Critical Incident Technique (Butterfield, Borgen, Amundson, & Maglio, 2005; Flanagan, 1954) and the Instructions to the Double technique (Oddone & Re, 1994; Oddone, Re, & Briante, 2008; Re & Oddone, 1991)).

New: The specific stakeholder approach to evaluation

Validation activities in A-PiMod have involved the application of a participatory/stakeholder approach to evaluation. The stakeholders involved in A-PiMod are referred to as the A-PiMod Community of Practice. Critically, these activities have developed a working collaboration with experts, which includes both ‘primary users’ (i.e. internal stakeholders representative of each project partner) and ‘all legitimate groups’ (i.e. external stakeholders representative of the aviation-related industry and Flight operational system). Both sets of stakeholders are involved in the specification and evaluation of the emerging adaptive automation concepts and technologies. This spans several activities pertaining to the specification and evaluation of user/technical requirements and user interface design prototypes. Internal stakeholders provide input based on their own domain knowledge. Further, they contribute in relation to assessing what is technically feasible and possible from a project perspective. On the other hand, external stakeholders provide feedback from direct experience and practice, to ensure that the emerging solution addresses real operational and safety requirements. Both internal and external stakeholders are conceived as active collaborators and contribute/engage in validation exercises on an on-going basis.

In the validation activities with the A-PiMod Community of Practice TCD’s role goes beyond that of a neutral facilitator. TCD’s role is to actively promote an interactive learning environment, where the stakeholders share their expertise and learn from the group collaboration. Indeed, TCD also act as a ‘key-broker role’ between the members of the Community of Practice to support (1) the review and specification of user requirements for the future system, (2) the production of relevant user interface design concepts/prototypes, and (3) the evaluation of prototypes.
New: What the methodology delivers in terms of ensuring improved levels of safety and reliability for the aviation sector.

A safety case has been advanced to support the specification of requirements and the assessment of safety/operational impact. The safety case comprises two parts – (1) the theoretical framework for the safety case and (2) the specific safety argument.

The safety framework provides a principled basis for conceptualizing/demonstrating how the A-PiMod adaptive automation concept and associated technologies will yield specific operational and safety benefits. This links to the demonstration of project impact, as discussed earlier. The framework is reported as a progression of ideas which form several phases. This includes: (1) background concepts which underpin the safety framework, (2) the starting point for conceptualizing the safety case, (3) the A-PiMod concept, and (4) the benefits of the A-PiMod adaptive automation concept and associated technologies from an operational and safety perspective. Each phase is associated with key points. The overall framework is depicted in Figure 2.
Figure 2. Safety Framework.

The safety argument articulates how specific operational/safety goals are achieved at the level of the A-PiMod technology (i.e. proposed architecture and technical components). Overall, the argument structure follows the theoretical approach and specific automation concept, as outlined in the safety framework. Specifically, the safety case/argument refers to specific steps in an overall use scenario – i.e. what technology does at different points in the scenario. See Figure 3 below.

![A-PiMod Safety Case/Argument](image)

**Figure 3. Safety argument.**

**What has been achieved so far in terms of validation approach**

**Research Undertaken**

The project started in September 2013. Overall, this research has involved two strands of activity – namely, (1) ongoing validation research with the Community of
Practice, and (2) the preparation of Validation Cycle 1 (i.e. comprising simulator evaluation, a parallel desktop evaluation and training evaluation).

In relation to (1), the on-going research activities with the A-PiMod Community of Practice, eight validation exercises sessions involving both internal and external stakeholders have been implemented. Overall, the objective of these sessions was to (1) define and integrate the A-PiMod adaptive automation concept, and (2) to harmonise the allied user/functional requirements. Both remote (by means of the WebEx functionality) and face-to-face workshops and/or interviews were undertaken. Prior to the validation exercise workshops, members of the Community of Practice were asked to complete tasks as defined by TCD. This served to facilitate the learning environment and promote the sharing of ideas and discussion in the specific workshops and/or interview sessions. Following every validation exercise workshop, TCD reported the minutes of the workshop and the consensus obtained on the topic. Further, TCD designed session specific templates to highlight the main results and integration of the Community of Practice members’ feedback.

In relation to (2), the first formal validation of the A-PiMod concept will take place in November 2014. The first Validation Cycle aims to evaluate and further specify the A-PiMod (1) adaptive automation concept, (2) the Multi-Modal Interaction concept and, (3) the training concept. In relation to the A-PiMod (1) adaptive automation concept, and (2) the Multi-Modal Interaction concept, this will involve an explorative user test with Pilots (i.e. four sets of crew), using a simulator. In addition, there will be some parallel evaluations (i.e. outside the simulator) with the same panel of Pilots (i.e. participatory review/design of concepts, semi-structured interviews to evaluate the concepts and so forth). In relation to (3), this will involve a parallel evaluation of the training concept, using semi-structured interviews.

The validation activities have produced a huge amount of qualitative data. Data recording and analysis has been undertaken with the assistance of a Computer-Assisted Qualitative Data Analysis Software (CAQDAS) tool - NVivo (© QSR International, V.8) (Bazeley, 2007). The use of difference sources of evidence during the data collection (i.e. interviews, observations, collaborative prototyping, etc.) allows for the assessment of convergence in relation to data evidence (data triangulation). This contributes to research validity. Further, the use of a concept-driven coding frame (based on the architecture and technology that A-PiMod intends to demonstrate) has supported the ongoing data analysis.

Emerging A-PiMod Adaptive Automation Concept

This research (i.e. use of innovate HCI design/evaluation methodologies) has resulted in the specification of (1) a new adaptive automation concept/approach and (2) the associated new technology concepts and requirements.

The problem of flying the aircraft is conceptualised as an ‘information processing decision’. This can be achieved in different ways (i.e. two/one person cockpit with different levels of automation, ground co-Pilot and/or ground support, or UAV/drone). In A-PiMod, these decisions will be undertaken by a two person crew with the support of automation. This is referred to as a ‘co-operative system’. The
underlying idea is that we can continuously monitor the operational situation and the
allied crew/automation state, to determine the best distribution of task activity
between the crew and automation. The basic philosophy is - if there is an increase in
workload, certain functions can be shifted to automation, to reduce the burden on the
flight crew. Automation is also used to support information management and
decision making tasks.

Critically, the A-PiMod system allows us to answer the following questions:

- Is the joint crew/automation system in a safe state (i.e. level of workload,
situation awareness)?
- Is there a potential for a safety critical aircraft state (i.e. now and/or the
near future)?
- Do we need to adjust the level of automation?

The crew obtain constant feedback via a new cockpit user interface as to status of (1)
the operational situation, and (2) the joint crew automation system. From an
operational/safety perspective this enables crew/automation teamwork, crew
workload management, and error identification and recovery. All of the above
ensures that the aircraft remains in a safe state. This in turn has consequences in
relation to the overall safety of the flight, and the achievement of process/operational
goals.

Discussion

The integration of formal and informal HMI methods, along with a stakeholder
approach to evaluation has proved effective in relation to the specification of the A-
PiMod concept. As outlined above, this has resulted in the preliminary advancement
of an innovative approach to automation, which addresses known problems.

Several points in relation to the stakeholder approach to evaluation should be noted.
First, the implementation of Community of Practice research is not straightforward.
This requires the advancement of a ‘working relationship’ with community members
(i.e. trust and teamwork), the set-up and acceptance of communication/information
sharing practices and the establishment of a decision making process. All of this
takes time. Further, the adoption of a participatory approach can make decision
making slow. However, on the positive side, this in turn fosters collaboration and
good co-ordination across project members.

In this regard, the TCD role has changed over the course of these validation
activities. Initially our role was one of a ‘facilitator’ and/or coordinator. We sought
to capture requirements and to advocate on behalf of the end user. Over time, we
have become more and more engaged in the current implementation of project
activities (i.e. in eliciting Human Factors requirements, suggesting user
requirements, designing user interface prototypes and so forth). In doing this, TCD
has adopted a ‘brokerage role’ between internal/external stakeholders. This is
underpinned by quality communication and the establishment of good working
relationships between TCD and internal/external stakeholders (i.e. trust and
teamwork).
The creation of an inclusive learning environment where members of the A-PiMod Community of Practice share ideas necessitates an appropriate setting (and potentially technology). In A-PiMod this has been mostly remotely telephone/web mediated (i.e. with WebEx), although some person to person interviews have been undertaken. Overall, person-to-person interaction has proved the most fruitful. As a result, the planning of the next validation exercises will consider more opportunities to meet in person. In time, technology may ‘catch up’, to provide a more natural/user-friendly environment for knowledge sharing.

Lastly, the importance of involving external stakeholders (i.e. pilots) cannot be understated. This involvement has been critical to the collection of user requirements and the emerging definition of the A-PiMod concept.

Conclusions

Safety is an operational concept and must be addressed at all levels: the air traffic management (ATM) system; the design of airline safety management system (SMS) processes and technologies; flight crew task activities and in particular, flight crew safety behaviour, and the design of cockpit systems/tools (including automation).

Overall, the evaluation/validation approach adopted has facilitated the preliminary specification and evaluation of a new adaptive automation concept. Specifically, the integration of a range of formal and informal HMI methods has proved effective in terms of enabling both operational and safety validation. The participation of stakeholders in the Community of Practice provides a strong link to the real world – in relation to (1) understanding automation issues, and (2) the capacity of technology to address these issues. Critically, the emerging adaptive automation concept is predicated on feedback in relation to flight crew experience with automation (and associated problems).

It is anticipated that these initial concepts will pave the way for an improved approach to automation. Preliminary evaluation feedback indicates that the concepts/technologies show promise in relation to solving pilot problems relating to teamwork (i.e. pilot/automation co-ordination) and workload management.

Acknowledgements

The research leading to these results/preliminary outcomes has received funding from the European Commission’s Seventh Framework Programme (FP7/2007-2013) under grant agreement N. 605141 - Applying Pilot Models for Safety Aircraft (A-PiMod) Project.

Thanks to the project coordinator – Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR) and to the other project partners – OFFIS (OFF), Honeywell (HON), National Aerospace Laboratory of the Netherlands (NLR), Symbio (SYM), Kite Solutions (KITE), Brno University of Technology (BUT). Also, thanks to the members of the A-PiMod Community of Practice, in particular the external stakeholders (i.e. pilots).
References


a novel approach to operational/safety evaluation of automation technology


