

Wright State University

CORE Scholar

International Symposium on Aviation
Psychology - 2019

International Symposium on Aviation
Psychology

5-7-2019

A New HMI Evaluation Method (Meria) Based on Pilot's Mental Representations

Théodore Letouze

Lisa Creno

Jamie Diaz

Sylvian Hourlier

Jean-Marc Andre

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2019



Part of the [Other Psychiatry and Psychology Commons](#)

Repository Citation

Letouze, T., Creno, L., Diaz, J., Hourlier, S., & Andre, J. (2019). A New HMI Evaluation Method (Meria) Based on Pilot's Mental Representations. *20th International Symposium on Aviation Psychology*, 355-360. https://corescholar.libraries.wright.edu/isap_2019/60

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2019 by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.

A NEW HMI EVALUATION METHOD (MERIA) BASED ON PILOT'S MENTAL REPRESENTATIONS

LETOUZE Théodore

Bordeaux University, BdxINP - ENSC, Lab IMS Cognitive UMR 5218
Talence, France

CRENO Lisa

Centre Aquitain des Technologies de l'Informations et Electroniques
Talence, France

DIAZ Jaime

Centre Aquitain des Technologies de l'Informations et Electroniques
Talence, France

HOURLIER Sylvain

Thales/ENSC, Human Engineering for Aerospace Laboratory (HEAL)
Talence, France

ANDRE Jean-Marc

Bordeaux University, BdxINP - ENSC, Lab IMS Cognitive UMR 5218
Talence, France

Continuous evolution of HMIs is necessary to keep operators in an optimal situation. In this context, we consider mental representations (MR) mobilized by operators as key elements for decisionmaking. Capturing and analysing these representations is not easy with existing tools. We propose a specific method (i.e. "MERIA" for Mental Representation Impact Analysis). Our case study focuses on a group of first officer (Airbus A320) in a dynamic situation with high time pressure. We are interested in cases where the HMI generates MRs that are inconsistent with the situation, resulting in a discrepancy between the prescribed activity and the actual activity. The goal is to identify the link between erroneous MR and the interface that created them. Our modelling structure allows us to create this link and place it in a proper temporal context. We observe that the constitution of the MR is different from one subject to another. However, invariants in the appearance of some erroneous MR make it possible to attribute the causality to an interface element well-defined in space and time. Thus, this analysis allows us to offer recommendations for HMI design to improve decision making. Our results show that the improvement does not lie in a drastic modification of the interfaces. Rather it allows a synchronization of the data coming from the cockpit with the pilot's MR of those data.

Introduction

We only observed the co-pilot activity. They intervene in a dynamic, uncertain, risky situation and they must make multiple decisions under the pressure of real time to achieve their performance objectives (Graziani et al., 2016). In the context of the cockpit of an Airbus A320, our objective is to determine which interfaces allow the co-pilot to build a good mental representation of the situation and which ones do not. In complex environments, HMIs and co-pilot cognitive activities can be evaluated in multiple ways. The methodologies we are interested in are those that aim to jointly evaluate the efficiency of the interface and its use by the operator. There are various categories of methods (Stanton, 2013): ETS (Annett, 2004), ACRC (Vicente, 1999), SAGAT questionnaire, SPAM method, Situational Awareness Requirements Analysis (Endsley, 1995, 2001, Selcon and Taylor, 1990), etc. None of these

methodologies creates a causal link between the user's cognitive process, his mental representation and the HMI used. This is the reason why we have proposed the MERIA (Mental Representation Impact Analysis) method, specifically adapted to this problem (see Letouzé & al., 2019).

Method

The methodological approach of MERIA is based on a triangulation of methods. It has been developed to design interfaces allowing operators to be more resilient in problem solving situations. It combines qualitative and quantitative methods that show, in a detailed and contrasted way, the observed activity (Altrichter, 2008). This method allows us to collect the general activity of the co-pilot (subject of this study) through three points of view: i) the experts describe the sequence of action expected in the scenario (prescribed task); ii), the experience of the co-pilot is collected by an interview (task performed); iii) we collect the general characteristics of the co-pilot (experience, personal data, etc.). The aim of this approach is to improve "the richness and sophistication of our analysis" (Guilbert & Lancry, 2007) and to get as close as possible to the "true value of the information collected" (De Battisti, Salini, & Crescentini, 2006), by crossing the three types of points of view.

Application of the method

The methodology is applied to a population of experts when performing a scripted and constrained activity in terms of progression and duration. The method is applied according to the following process: (1) A scenario representative of the co-pilot's activity is defined precisely. It is also verified that this scenario is reproducible under realistic conditions. The expected performances at each stage are defined by a collective of experts in pilot operations, aeronautics and cognitive sciences. Each performance element is associated with a mental representation. This is why cognitive science experts need to be involved in this phase. (2) A homogeneous panel of co-pilots is recruited. These co-pilots do not know the scenario. (3) The scenario is performed by the co-pilots of the panel in a cockpit of the current A320. During the scenario, cognitive science experts observe the activity and identify key events. A pilot expert comments on the co-pilot's actions to make the activity more explicit. The experts (cognition and aeronautics) are not in the copilot's environment. (4) Immediately after the end of the scenario, the cognition expert conducts a self-confrontation interview. During the interview, the co-pilot is "put back in the situation", the expert making him relive the scenario step by step. This expert identifies the RM associated with each step. (5) This interview shows the evolution of the RM over time using the MERIA grid (e.g. Figure 1). This grid is completed by observation of the experiment (3). We collect one grid per co-pilot. An inter-judge measurement method (Cohen kappa) makes the coding process of the grid more reliable. (6) From these grids, cognition experts identify the problematic interfaces and those that produce the expected effect.

The graphical representation: MERIA grid

The tool is constructed as follows: The prescribed scenarios are represented by white squares placed in the "NODES" column. In the "INPUT" column, squares indicate the different sources of information addressed to the co-pilot. In the "MENTAL REPRESENTATION" column, we coded the elements relating to the co-pilot's actual mental representation. It varies and evolves during the scenario depending on what the co-pilot

perceives, understands and anticipates. In the "IMPACT" column, we indicated the consequences of the actions implemented as soon as the mental load, the choice of the airport/runway (Bremen or alternatives) and/or the landing limitations were affected.

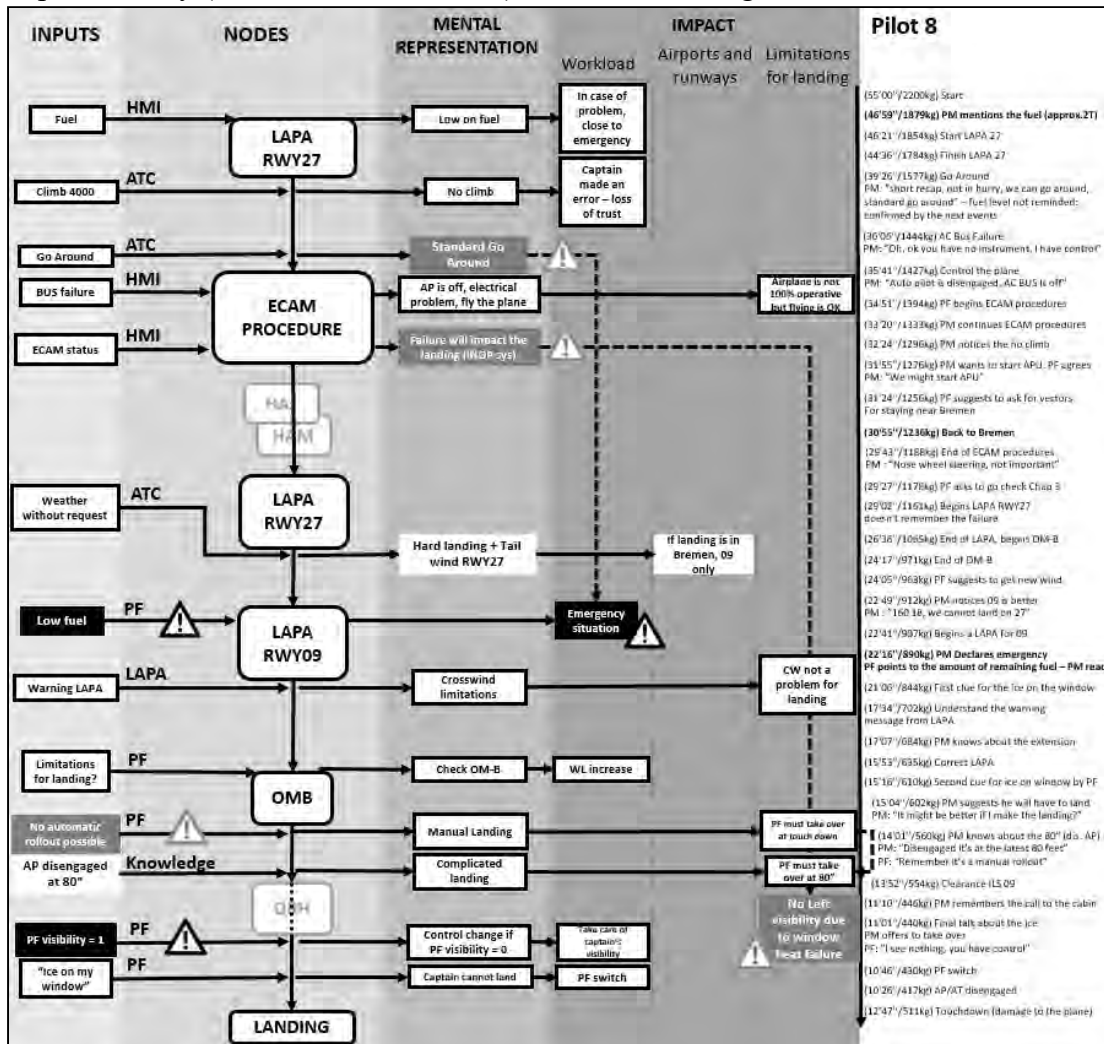


Figure 1. MERIA Model of Pilot # 8. (White without outline: actual performance > prescribed, White with black outline: real = prescribed, Gray: real < prescribed, without being critical, Black: real << prescribed, critical state, Triangles allow to quickly identify the elements that interfere with activity).

Scenario

The context of the critical scenario observed is unique since it begins in the middle of the flight (in a phase just before the approach and landing) so that the co-pilots are not aware of the amount of fuel remaining (fuel on board) to reach the end of the flight. The scenario has 4 phases of unequal duration (see Figure 2). Different "nodes" (or key elements of the scenario) structure these phases.

Description du panel

We have access to a panel of 10 co-pilots from Lufthansa Airlines, trained at Bremen Airport (Germany). They know all the particularities of this airport: runway length, nearby airports, unofficial runway extension, etc. These voluntary and paid participants are "experts"

of the task, which offers a particular interest for the analysis of decision-making, the knowledge of the situation and the mechanisms of mental representation studied.

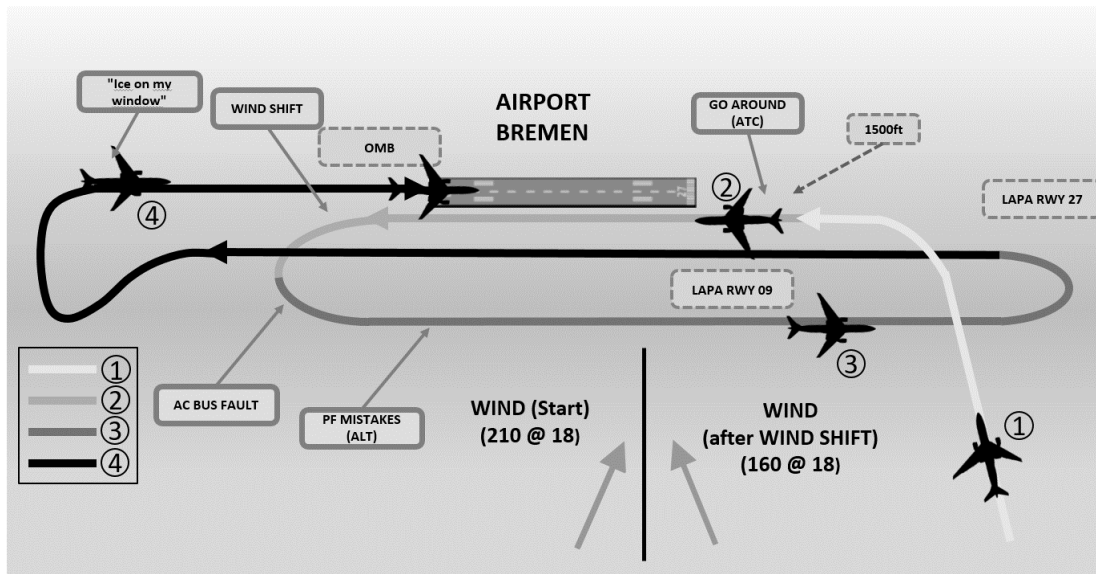


Figure 2. Schematic representation of the prescribed scenario (Phase 1 - Start of the go-around scenario, Phase 2 - Until failure, Phase 3 - Until the decision to land on runway 09, Phase 4 - Up to on landing).

As a professional, none of the pilots recruited are captain, that is to say, responsible for the plane or its passengers. Co-pilots averaged 30.9 years (min: 28, max: 36, standard deviation (SD): 3.28), a total of 4045 flying hours on average (min: 2250, max: 7000, SD: 1569), of which 3125 hours on average on Airbus A320 (min: 250, max: 6000, SD: 1557) and 667.78 hours on average over 12 months (min: 600, max: 750, SD: 42.78).

Results

From the MERIA grids we constructed from the 10 self-confrontation interviews, we were able to identify gaps between the expected mental representations and the actual mental representations. The different phases of the scenario (inputs and associated mental representations), allow us to identify needs for co-pilot. We can also identify services that the system could render to the co-pilot. Our analysis is among us to identify 10 services not rendered by the system that penalizes the mental representations of pilots. These services would be required for the completion of the requested task but the existing system is not designed to respond to it. In some cases, it is the training or the expertise / experience of the pilot that addresses this deficiency.

A total of 56 unreturned services are recorded for all 10 pilots as we can see on Table 1. These services can be categorized (some occurrences can belong to more than one category). It is observed that 9 occurrences can be identified as feedback defects from the system. That is to say a lack of visibility of the signal, or even its absence. 21 occurrences can be identified as a lack of data synthesis (cross-referencing of several types of information) and a lack of explanation of their consequences. 26 can be likened to a lack of spatial clarification of constraints and possibilities. Finally, 33 occurrences are similar to a failure to represent temporal constraints and the evolution of the system over time. The domains of spatial and temporal representations overlapping to a certain extent.

Table 1.

Number of services not provided by the system for all 10 pilots over the entire scenario. During the chosen scenario, the system must provide 10 services.

Service	Number of times the service has not been delivered (/10)	Services provided to the co-pilot 8
Warning about fuel level	9	Yes
Feedback : actions done	2	Yes
Projection of the flight action field compared to the needs to land	10	no
Combining failures to explain consequences and keep it in PM mind	7	no
Projection of the flight action field (fuel + wind + speed) compared to the needs to land	5	Yes
Combine the weather with the needs to land and explain options	7	no
Combine the aircraft state with the needs to land and explain limitations	4	Yes
Combining failures and information from documentation to explain consequences and keep it in PM mind	7	Yes
Explain what should be done to follow procedures	3	Yes
Combining failures and aircraft state to explain who should take control at each time	2	Yes
Total	56/100	3 services non delivered

Through this analysis of the results, and by combining this information with self-confrontation interviews, we observe that the representations provided by the system were not consistent with the representations expected by the co-pilots. The co-pilots evaluate the field of possible temporally (available flight time) and/or spatially (attainable distances). For the fuel on board the aircraft, for example, the system produces an indication in kilograms while the co-pilots convert it into minutes or nautical miles. This inconsistency is found for other information the aircraft provides: the inoperative systems, the co-pilot reflects the type of failure and consequences, the weather is transmitted in code form (METAR) and the data are relative to the ground, the co-pilots reflect in terms of cloud layers and relatively to the aircraft. Such difference between the information provided by the system and the cognitive functioning of the co-pilot creates a blow of information conversions that reduces performance and can potentially lead to conflicts in human-system collaboration or errors.

Conclusion

The use of the MERIA methodology highlights activities for which pilots are not properly assisted by the system. From this point of view, the results are similar to those obtained with SA measurement methodologies. The added value of our approach comes from the fact that the MERIA method highlights the discrepancies between the expected and actual mental representations of the co-pilots. It makes it possible to identify exactly where and when the source or sources of the offset on the manipulated HMI are. These results open the door to new studies on system design and evaluation. In this case of application, the

methodology allowed us to identify areas for improvement of the A320 cockpit system. This study highlights that improved co-pilot performance can be achieved through conceptual system changes and improved communication between operators and systems. In conclusion, the MERIA tool provides a solution to the evaluation and improvement of Man Machine Teaming.

Acknowledgement

The results presented here are based on the work carried out under the Future Sky Safety program, funded by the Horizon 2020 research and innovation program of the European Union under Grant Agreement No 640597.

References

- Altrichter, H. (2008). *Teachers investigate their work: an introduction to the methods of action research* (2nd ed). London; New York: Routledge.
- Annett, J. (2004). Hierarchical Task Analysis. In Neville Anthony Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. W. Hendrick (Eds.), *Handbook of human factors and ergonomics methods* (pp. 329–337). Boca Raton, FL: CRC press.
- De Battisti, F., Salini, S., & Crescentini, A. (2006). Statistical Calibration of Psychometric Test. *Statistica & Applicazioni*, 4(2), 1–25.
- Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors*, 37(1), 65–84.
- Endsley, M. R. (2001). Designing for situation awareness in complex systems. In *Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment* (pp. 1–14).
- Graziani, I., Berberian, B., Kirwan, B., Le Blaye, P., Napoletano, L., Rognin, L., & Silvagni, S. (2016). Development of the Human Performance Envelope Concept for Cockpit HMI Design. In *HCI-Aero 2016 International Conference on Human-Computer Interaction in Aerospace*. PARIS, France. Retrieved from <https://hal.archives-ouvertes.fr/hal-01409075>
- Guilbert, L., & Lancry, A. (2007). L'analyse des activités des cadres : l'intérêt de la triangulation des méthodes. *Le travail humain*, 70(4), 313–342. <https://doi.org/10/fm327w>
- Selcon, S. J., & Taylor, R. M. (1990). Evaluation of the Situational Awareness Rating Technique(SART) as a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 8 p (SEE N 90-28972 23-53)*.
- Stanton, Neville A. (2013). *Human factors methods: a practical guide for engineering and design*. Burlington, VT: Ashgate Publishing Company.
- Vicente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. CRC Press.