Feasibility study on operationability of a remote pilot station for multiple PATS control: a Human Factors simulation experiment

Henk van Dijk & Robbert Ouwerkerk
National Aerospace Laboratory - NLR
Amsterdam, the Netherlands

Abstract

The European Commission 7th Framework project called Personal Plane (PPlane) aims at developing a system to enable individual air transport. The objective of such a Personal Air Transport System (PATS) is to offer an alternative for the current transport system. The highly automated PPlane is characterized by having onboard only passengers, without piloting skills. A so-called Ground Pilot controls the flight from a Remote Pilot Station.The current paper presents a feasibility study on the operationability of the Remote Pilot Station for the control of multiple PPlanes. Relevant Human Factors issues - concerning the Ground Pilot - were identified and evaluated. The following Human Factors issues were recognized (through expert interviews / workshops) as relevant for the Ground Pilot: mental workload, situation awareness, and human error. The work proceeded with the evaluation of the identified Human Factors issues. A controlled experiment was conducted using a Remote Pilot Station simulation platform. The following variables were studied in the experiment: the user (or human-machine) interface for the Ground Pilot, the number of PPlanes under control of the Ground Pilot and the designated crew of one or two Ground Pilot(s) in the Remote Pilot Station.

Introduction

The PPlane (Personal Plane) project aims at developing operational concept ideas to enable individual air transport. The objective is to avoid the ever increasing congestion on European roads and to offer an alternative for the current transport system in the European Member states. The idea is that a PPlane offers inter-city personal air transport that will carry two to four passengers between cities. PPlane is designed to be a highly safe and secure mode of transportation with a low environmental impact. It flies for relatively short distances of a few hundred kilometers. Two separate operators are involved in the PPlane concept of operations. The passenger in the airplane itself forms the PPlane Passenger. The PPlane Passenger has control of the PPlane like a passenger in nowadays taxies. The PPlane passenger provides the desired destination or specific driving request. The driver is responsible for safely getting the passenger where he wants. In the PPlane concept this role of (taxi)driver is performed by a PPlane operator, the Ground Pilot that controls the plane from the ground. Compared to nowadays-private jet services this

In D. de Waard, K. Brookhuis, F. Dehais, C. Weikert, S. Röttger, D. Manzey, S. Biede, F. Reuzeau, and P. Terrier (Eds.) (2012). Human Factors: a view from an integrative perspective. Proceedings HFES Europe Chapter Conference Toulouse. ISBN 978-0-945289-44-9. Available from http://hfes-europe.org

means that the pilot controls the airplane from the ground instead of from on board the plane.

The current paper presents (a part of) the Human Factors work within the PPlane project, with specific focus on the Ground Pilot. A feasibility study on the operationability of the Remote Pilot Station for the control of multiple PPlanes is performed. The objective of this Human Factors study, using a simulated operational environment, is to evaluate whether or not the PPlane operational concept impacts performance of the Ground Pilot with regard to mental workload, situation awareness, and human error.

Methodology

The Eurocontrol Human Factors Case (Eurocontrol, 2007) was used as guideline to address and manage Human Factors issues throughout the PPlane project. The Human Factors Case supports pragmatic integration of Human Factors within the system concept's development life-cycle. It analyses and optimises the human interaction with a system.

Human Factors issues identification

A preliminary working PPlane operational concept was adopted to serve as a basis for the discussion with the experts and potential end users. This working operational concept of PPlane envisages the use of small, highly automated, ground controlled aircraft to carry passengers at their request from one destination to another. This operational concept involves two human actors as a minimum. Firstly, one passenger on board, whose role in the control of the vehicle may be either active (high level instructions / intentions, mostly related to navigation) or passive (fully automated system). Secondly, one pilot on the ground (in a Remote Pilot Station), who will be responsible for the safety and efficiency of the flight. The present paper has its focus on Human Factors issues concerning the Ground Pilot.

Workshops and interviews with air traffic controllers, (remote) pilots, and expert researchers were held to identify from different perspectives the Human Factors issues that need to be addressed. The analysis identified and prioritised a number of Human Factors related issues and their possible impact on performance of the Ground Pilot:

- Mental workload;
- Situation awareness;
- Human error.

Human Factors issues evaluation

In accordance with the Eurocontrol Human Factors Case methodology, actions were undertaken to resolve the identified Human Factors issues. The main action reflected the execution of an experimental study using a PPlane operational Remote Pilot Station simulation. The operational set-up for the simulation was fed by mitigation of the identified Human Factors issues and their possible impact on the Ground

Pilot. That is, adequately presenting information to the Ground Pilot in the Remote Pilot Station, designing an intuitive and user-friendly user interface for the Ground Pilot, and allocating tasks between the Ground Pilot and PPlane system.

Experiment set-up

In order to simulate the PPlane concept of operations two main applications were used: an application that simulated the PPlanes (simulator) and an application that controlled the PPlanes (user interface). For the simulation of the PPlanes an NLR tool called WinTMX was used. The main functions of the WinTMX are generating air traffic and simulating aircraft behavior, such as flight dynamics and flight performance. It can simulate up to 1000 aircraft simultaneously. Each aircraft is capable of flying a trajectory or route starting at an origin and going to a destination. During flight these aircraft can be controlled giving them commands to change speed, heading or altitude.

The design of the Ground Pilot user interface (Figure 1) included an overview map on which all PPlane under control and their position on the map were visualised. Requests / events from different PPlanes were illustrated in an action item list. The action items could be picked up, handled, and deleted by the Ground Pilot. This action list was inspired upon the lists with incoming calls that call centre managers have available. It also provided priority information for each individual task (i.e. colour coded: red for high priority actions, amber for medium priority actions and grey for low priority actions). Icons representing PPlanes with requests were distinguishable for other PPlanes by blinking. By clicking on an icon that represents a PPlane, its planned route and additional flight information became visible on the map. Further, a window was opened to communicate (via text message) with the passengers. Besides text messaging, communication was possible via voice (using a headset) to the different actors involved (e.g. passengers, air traffic controllers, and other Ground Pilots).

In the experiment, Ground Pilots were sometimes working alone and sometimes paired in teams of two. For the conditions, where two Ground Pilots were controlling a number of PPlanes together, a large wall display visualising all PPlanes under control and their requests was available. Ground Pilots were able to click on icons representing a particular PPlane after which they got control over that PPlane on their own desktop-computer. A shared action list was also visualised on the large wall display where both Ground Pilots could pick up and assign tasks.

The Ground Pilot had a regular mouse available for interaction with the overview map (i.e. selecting PPlanes and zooming in/out the map) and other onscreen elements such as the PPlane information panel and action list. A keyboard was used for text input for the messenger window. In addition, changing heading and speed, moving waypoints, and other PPlane specific interaction were done with a 3D mouse.

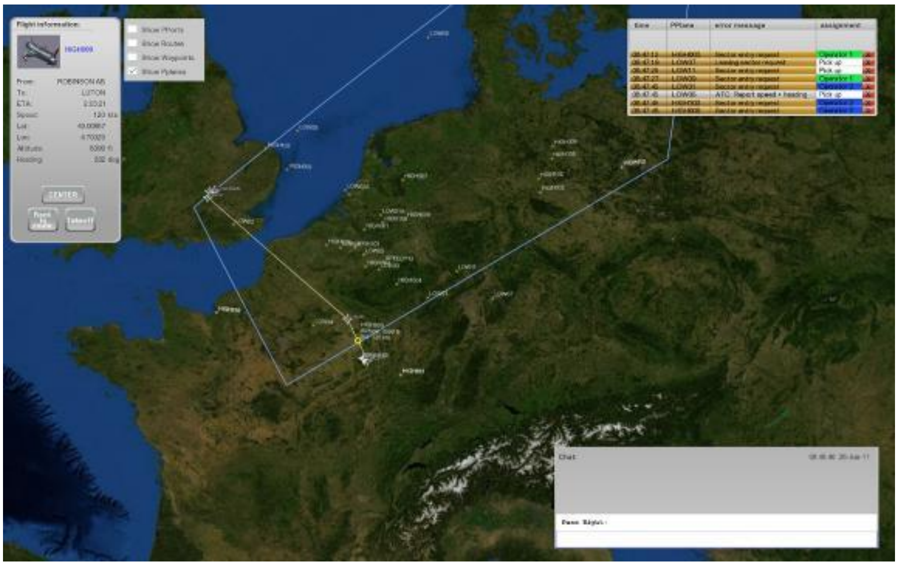

Figure1. Ground Pilot user interface

The Ground Pilot user interface was built with a display prototyping tool called Vincent (Verhoeven & De Reus, 2005). Vincent was upgraded with an Open Scene Graph extension: an open source high performance 3D graphics toolkit used by application developers in fields such as visual simulation, games, virtual reality, scientific visualization and modeling.

Experiment actors

Each PPlane would fly between flight level 50 - 200, have a maximum speed of 600 km/h (= 300 kts) and a minimum speed of 150 km/h (= 100 kts). They would have a range of 100 - 500 km, fly according to a 4D trajectory contract, and can automatically take off and land after receiving clearance from the Ground Pilot.

In the experiment the Ground Pilot was controlling the PPlanes in a particular area of operation in which a PPlane could fly in and out. For the experiment, the area of operation consisted of the square between London, Stockholm, Berlin and Paris. PPlane traffic in this area was simulated. In the experiment, the PPlane area was filled with PPlanes only. Ground Pilot tasks included communicating with air traffic control for take off and landing clearances; giving PPlanes a go for take offs and landings (after receiving clearance); checking handovers of PPlanes from and to another area; talking to PPlane passengers, setting them at ease; handling off requests for destination change; handling off emergency procedures (using checklists). The experiment management simulated both the passengers and the air traffic controllers.

Experiment scenario

Each scenario started with two or three PPlanes. After a short while, the number of PPlanes increased to a stable number. Two scenarios were foreseen here:

1. A low density traffic scenario that contained ten PPlanes that were controlled from take off to landing;
2. A high density traffic scenario with a total of 30 PPlanes.

Each scenario lasted about 30 minutes and included nominal and non-nominal events (emergencies) for the Ground Pilot to handle. Participants were briefed and trained on forehand, running two example scenarios and familiarizing themselves with the user interface and communication set-up.

Experiment design

The experimental design involved two independent variables, namely the number of PPlanes under control in the area (low vs. high density) and number of Ground Pilots in remote pilot station (single vs. team). In the team variant, only the high density scenario was run. The dependent variables were mental workload, situation awareness, and human error. This resulted in the following schedule:

1. Run 1 low density single Ground Pilot;
2. Run 2 high density single Ground Pilot;
3. Run 3 high density team of Ground Pilots.

Experiment metrics

Several metrics were used to measure the impact on the human performance during the experiment runs and post-run. Additionally, post-run surveys were used to indicate the estimated number of errors (i.e. number of actions with undesired outcome) made by the participant in performing the tasks at hand.

The Likert scale is a psychometric scale commonly used in questionnaires. When responding to a Likert questionnaire item, respondents specify their level of agreement or disagreement on a symmetric "agree - disagree" scale for a series of statements. The Likert-scale was used to address mental workload and situation awareness during run via the text messenger functionality.

The Crew Awareness Rating Scale (CARS; McGuinness & Foy, 2000) is a situation awareness assessment technique (post-run) that is based upon the three-level model of situation awareness (Endsley, 1995). The questionnaire consists of eight questions. Each question measures situation awareness on a four-point scale ranging from "very well" (1) to "very poor" (4). In the analysis, the numerical results from the questionnaire were inverted and normalized to generate a continuous scale ranging from 0 (low situation awareness) to 1 (high situation awareness).

The Rating Scale Mental Effort (RSME; Zijlstra, 1993) rates invested mental effort by a cross on a continuous line. The line runs from 0 to 150 mm. Along the line, at several anchor points, statements related to invested effort are given, e.g. “almost no effort” or “extreme effort”. The RSME was used post-run.

The System Usability Scale (SUS; Kirakowski & Corbett, 1988) is a simple, five-item attitude Likert scale giving a global view of subjective assessments of usability (post-run). SUS was used post-run. The sum of the scores was multiplied by 2.5 to obtain the overall value of usability within a range of 0 to 100. The active screens of the Ground Pilots were video recorded during each run. If necessary, these recordings were used in the debriefing that was conducted at the end of the experiment.

Experiment results

Participants

Table 1. Particpant information

Number of participants	14 students
Age	22 mean 1.7 standard deviation 20 - 26 range
Education	Aviation engineering (7x) Information science (6x) Journalism (1x)
Gaming experience	“I play games on a regular basis” (8x) “I sometimes play games” (5x) “I never play games”(1x)

Human performance

Table 2. Workload Likert scores during run

Scenario	Nominal	Non-nominal
Run 1 low density_single	1.9 mean 0.7 standard deviation 1 - 3 range	2.9 mean 0.8 standard deviation 2 - 4 range
Run 2 high density_single	2.6 mean 0.9 standard deviation 2 - 5 range	5.6 mean 0.9 standard deviation 4 - 7 range
Run 3 high density_team	1.8 mean 1.1 standard deviation 1 - 4 range	3.3 mean 1.8 standard deviation 1 - 6 range

Table 3. Workload RSME scores post-run

Scenario	
Run 1 low density_single	32.9 mean 15.6 standard deviation 10 - 70 range
Run 2 high density_single	70.7 mean 22.5 standard deviation 24 - 120 range
Run 3 high density_team	44.2 mean 14.4 standard deviation 25 - 65 range

Table 4. Situation awareness Likert scores during run

Scenario	Nominal	Non-nominal
Run 1 low density_single	6.2 mean 1.0 standard deviation 4 - 7 range	5.4 mean 0.9 standard deviation 4 - 7 range
Run 2 high density_single	5.6 mean 1.0 standard deviation 3 - 7 range	3.7 mean 1.0 standard deviation 2 - 5 range
Run 3 high density_team	5.9 mean 1.4 standard deviation 3 - 7 range	5.2 mean 1.6 standard deviation 2 - 7 range

Table 5. Situation awareness CARS scores post-run

Scenario	
Run 1 low density_single	0.9 mean 0.2 standard deviation 0.6 - 1.0 range
Run 2 high density_single	0.6 mean 0.3 standard deviation 0.3 - 1.0 range
Run 3 high density_team	0.8 mean 0.2 standard deviation 0.6 - 1.0 range

Table 6. Human error post-run

Scenario	
Run 1 low density_single	0.9 mean 1.4 standard deviation 0 - 5 range
Run 2 high density_single	2.4 mean 2.5 standard deviation 0 - 10 range
Run 3 high density_team	0.9 mean 1.6 standard deviation 0 - 5 range

The overall mean SUS score for the Ground Pilot's user interface was 76.3 (standard deviation 11.3 and range of 50 - 95).

Discussion and conclusion

Two different experimental variables (independent) were researched in the PPlane simulation experiment. Firstly, the number of PPlanes under control (low vs. high density); secondly, the number of Ground Pilots in the remote pilot station (single vs. team). The experiment looked into how these two variables affected the human performance in terms of mental workload, situation awareness, and human error.

Participating Ground Pilots were questioned, observed and measured throughout all experimental runs. Given the relatively low number of participants, data analysis was performed on a descriptive level. A clear trend between the low and high density traffic scenarios was illustrated. In the high density run (single) Ground Pilots showed higher mental workload, lower situation awareness, and more human errors compared to the low density and team (high density) run. This is in conformance to the air traffic control research done by Majumdar (2002). He concluded that traffic density is a major factor in mental workload for air traffic controllers.

The differences between team (high density) run and the low density (single) run were less great. The team run showed slightly higher workload, slightly lower situation awareness and slightly more human errors. That is, the team run divided the high density scenario into a lower density scenario that was handled by the two Ground Pilots. In some cases, however, working together resulted in some extra workload and stress. As Wickens et al. (1997) already stated in the ninetees, because air traffic control is a team activity, another possibility is that controllers may ask a colleague to take over a particular task. In general, controllers may use a variety of strategies to manage workload and regulate their performance: if they do not use any of these adaptive strategies, further increases in traffic load may result in errors.

In case of the non-nominal events that were simulated, too many of these events at the same time combined with the control over high density traffic resulted in an overload of the Ground Pilot. This was the case in the end of the high density (single) run. In the team run, this situation was prevented from happening because of the possibility to split the task load between the two Ground Pilots.

The participant's comments indicated that the Ground Pilot's user interface turned out to be highly intuitive. Of course, specific items could be improved, but the approach that was chosen to monitor / control the highly automated PPlanes (i.e. with the limited controls available for the Ground Pilot) was generally well received. The large wall display was hardly used in the team setting, as it presented no relevant additional information compared to the desktop display.

Emergency management seems to be very important aspect of the Ground Pilot's work. If all goes well (nominal), the work seems relatively low profile, working off all low / medium priority action items such as take off and landing, sector changes and talking to the passengers. Moreover, some of these tasks (e.g. destination changes and communication to the air traffic controller) are likely to be automated in the future as well. This leaves out the emergencies / non-nominal situations in which the Ground Pilot seems to have some sort of mediator role, making sure all parties involved are informed about the situation.

Acknowledgement

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement n° 233805.

References

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 32-64.
Eurocontrol (2007). *The Human Factors Case: Guidance for Human Factors Integration*: http://www.eurocontrol.int/humanfactors/.
Kirakowski, J., & Corbett, M. (1988). Measuring user satisfaction. In D.M. Jones and R. Winder (Eds.), *People and Computers IV*. Cambridge: Cambridge University Press.
Majumdar, A. (2002). Factors affecting air traffic controller workload: Multivariate analysis based on simulation modeling of controller workload. *Transportation Research Record Journal of the Transportation Research Board*, *1788,* 58-69.
McGuinness, B., & Foy, L. (2000). A subjective measure of SA: The Crew Awareness Rating Scale (CARS). In *Proceedings of the 1st Human Performance, Situation Awareness and Automation Conference*. SA Technologies, Inc.
Verhoeven, R.P.M., & Reus, A.J.C. de (2004). Prototyping interactive cockpit applications. In *Proceedings of 23rd Digital Avionics System Conference, 2004*.
Wickens, C.D., Mavor, A.S., & McGee, J.P. (1997). *Flight to the Future: Human Factors in Air Traffic Control*. Washington, DC: The National Academies Press.

Zijlstra, F.R.H. (1993). *Efficiency in work behavior. A design approach for modern tools*. PhD thesis, Delft University of Technology. Delft: Delft University Press.