The Effect of Increasing Levels of Automation in Demand-Capacity Balancing processes on Human Actors and their Roles and Responsibilities

A Gaming Assessment

M.F.J. Koolloos, D. Escribano, R.A. Casar
Transport and ITC Department, R&D sector
Ingenieria de Sistemas para la Defensa de España –Isdefe
Madrid, Spain
racasar@isdefe.es

A. Groskreutz, R. Garcia
Centro de Referencia de Investigación, Desarrollo e Innovación ATM – CRIDA
Madrid, Spain
argroskreutz@crida.es

T. Bos
Training Simulation and Operator Performance Dep.
National Aerospace Laboratory – NLR
Amsterdam, The Netherlands
tanja.bos@nlr.nl

***Abstract* - This paper describes the first results of the assessments of the impact of increasing levels of automation (LoA) in on human factors aspects, like roles and responsibilities, interactions and trust in the system.**

A set of scenarios addressing different automation levels were defined for Demand-Capacity Balancing processes in an "Airspace Organization and Management" environment. These scenarios were simulated with three actors by means of Gaming Sessions. First a preparative "Paper Based" gaming session was organized, followed by a "Hardware Based" gaming session on the CHILL platform.

Both gaming sessions were performed with only three actors and a limited number of runs. Therefore, a thorough statistical analysis was not possible and the questionnaires were analysed mainly on a qualitative basis. Nevertheless, a graphical presentation of the ratings was provided and a trend analysis could be done.

The results showed that trust and acceptance of the system increased or was at least maintained with increasing LoA and that the situational awareness and the workload remained at sufficient levels. The teamwork decreased with increasing LoA which took away the feeling of doing it together and recommendations with respect to possible future support tools were especially related to how the actors communicate.

Keywords - Automation, ATM, evolution, human role, gaming

I. INTRODUCTION

A major issue in Air Traffic Management (ATM) is the impact of the automation and new technologies on the human operator. Automated systems must be compatible with human capabilities. Development of effective and usable automated ATM systems requires Human Factors input throughout the design life cycle, from concept formulation, through detailed design, to implementation and operation.

Because of the importance of the coming changes of the future ATM system and its effect on human behaviour, now and in the far future, many European projects have focused their efforts on the evolution of human roles [1, 2, 3, 4]. Not surprisingly, one of the key focuses of SESAR is the role of the human and their relation with advanced automation tools to work safely, with an appropriate workload and with a high level of situation awareness.

The necessary level of automation for 2020 in a SESAR environment will be studied in the SJU programme, but this level of automation will be higher in the ATM environment after SESAR. There is some knowledge regarding how automation impacts on the workload and on the performance of an actor, but there is little knowledge regarding how the same automation impacts on the interactions and situation awareness among several actors with different interests (e.g. airspace users and ANSPs). Experience (e.g. automation in cockpit) has shown that the transition to higher levels of automation has to be carefully addressed to promote the trust in the new system, (guarantying certain safety levels), and to ensure its acceptability by the human operator. The ADAHR project, from which the first results are presented in this paper, is focussed on the impact of these high levels of automation on the human roles, with a time horizon ranging from 2020 to approximately 2050.

II. PROJECT BACKGROUND

A. Objectives

The basic research questions that are addressed by ADAHR are:

- How do different levels of automation in ATM impact human interaction?
- How does automation impact the roles and responsibilities in different environments?
- Which automation functionalities can be brought forward in the SESAR master plan?
- What requirements are needed to promote the acceptance of high levels of automation in the ATM environment?

Based on these research questions, the main objective of ADAHR is to assess the behaviour of different ATM human actors in a highly automated environment. This objective will be achieved through the analysis of:

- the impact of automation in the interaction between human actors.
- the impact of automation on the new roles and responsibilities foreseen by the levels of automation addressed.
- the tools supporting the ATM human actors' duties.

These objectives are studied in two different environments: 1) Airspace Organization and Management; and 2) Airport Operation Centre. This paper presents the results of the study to the first environment.

B. Approach

Starting from the SESAR definitions of human roles in 2020 and the expected technical capabilities of automation, new levels of automation were assumed and new roles and responsibilities of human operators were defined, for both environments [5]. Next, the scenarios were defined for the two different environments and for three future automation levels (corresponding to a situation around approximately 2020, 2035 and 2050) [6]. Finally, the impact of the automation described in the scenarios on human roles is assessed. This is done using the Gaming Technique, which has proven to be very suitable for this kind of assessments [7, 8, 9], and which is briefly described in the following sub-section.

C. Gaming Technique

Human-In-the-Loop (HIL) Gaming technique are "serious games", designed for a specific purpose other than pure entertainment. These games are played with persons (mostly, experts) acting as actors and allow the exploration of concepts and definition of roles and processes in a structured way focusing the players' attention on the information flow and responsibilities associated to the processes. HIL Gaming technique has proven to be an excellent technique to explore the situation awareness and the human-human and human-machine interactions in automated environments [7, 8, 9] because it enables a research team to:

- learn about systems that have not been developed yet,
- study the behaviour of the people and/or machines' interactions with lower and controlled cost (compared to prototypes)
- stimulate actors to be open-minded and obtain results from different points of view.

Experience in ATM assessment has proved that the combination of role-based games using paper with role-based games using hardware-platforms provides a good quality assessment of the process involved in the concept under test [8, 9]. Paper-based games are performed using basic office material. They are basically board games where the rules are designed according to the processes and roles interactions to be studied/ clarified. Hardware-based games are basically performed in the same way as the paper-based ones, but the means/tool to play is a hardware platform. They use and complement the results obtained from paper-based games and have the following benefits:

- the platform contributes to execute the exercises in a more realistic context
- their results are more reliable and accurate.
- more complex game rules can be used
- performance analysis is easier,
- there is a closer link with different levels of automation.

Combination of both techniques will allow the definition and exploration of roles and their responsibilities and the interaction of these roles within an automated environment in two steps. Firstly, through paper-based games, obtaining the high-level results and next through hardware-based games, in which paper high-level and preliminary outcomes will support the platform/s configuration and also they will be the baseline used to produce the final results.

For the Gaming sessions described in this paper the CHILL (Collaborative Human-In-The-Loop Laboratory) platform was used. CHILL is a versatile collaborative ATM validation platform in which different categories of actors can work together to efficiently manage traffic demand and capacity, exchange ATM data and share information in support of a collaborative ATFM planning process. The CHILL modelling Platform has been designed as a suite of interoperable modelling services and components to support the evaluation, performance assessment and validation of existing and future ATM concepts of operation.

III. GAMING SESSIONS

A. Scope of the ExercisesAbbreviations and Acronyms

The environment of Airspace Organization and Management is focused on the processes of Demand and Capacity Balancing (DCB) at the planning phase. This scenario is based on timeframes beyond SESAR horizon, 2020, which is used as baseline for the exercises. Two more timeframes, 2035 and 2050, were defined with increasing LoAs for each. Such automation is assumed to be undertaken by systems that are supposed to progressively converge into a common one.

1) Objectives

The Paper-based gaming session had the following specific objectives [10]:

- Evaluate the appropriateness of the assumed levels of automation for 2020, 2035 and 2050;
- Assess the role definition and their responsibilities;
- Pre-selection of solutions to be evaluated;
- Definition of cost-index parameters;
- Preliminary analysis of the impact of various LoAs in DCB processes on human actors, on their (new) roles and responsibilities and on the interaction between human actors.

The Platform- based gaming sessions had the following objectives [11]:

- To assess the level of trust, acceptance and situational awareness that airspace traffic management related actors have in increasing levels of automation.
- To assess the impact on teamwork of increasing levels of automation.
- To assess the impact on workload of increasing traffic density in combination with increasing levels of automation
- To obtain needs of possible future support tools. Identify characteristics that should be in or out.

2) Metrics

For the last objective of the paper-based gaming sessions the indicators Trust and Confidence, Workload, Performance, and Situation Awareness were measured [10].

During the platform-based gaming sessions the indicators trust in automation, its acceptance, the participant's workload, situational awareness and teamwork were assessed [11].

3) Collection and Analysis methods

In order to assess the indicators, mainly qualitative data was collected by means of the following methods:

- Over the shoulder observations: during the different runs of the scenarios, experts on Human Factors and/or the subject matter performed direct observations on the behaviour of the controllers. These observations improved the debriefing sessions presenting issues to be discussed or points to be highlighted,
- Questionnaires: During the paper-based gaming sessions short tailor-made questionnaires were used for a first assessment. During the platform-based gaming sessions structured questionnaires related to the different indicators were to be completed by the participants post-run. The questionnaires were based on standard, validated questionnaires generally applied in the field of ATM and/or aviation (CARS – Controller Acceptance Rating Scale [12], Situational Awareness Rating Scale [13], NASA Task Load Index [14], SHAPE Automation Trust Index and SHAPE Teamwork Questionnaire [15])
- Debriefings: The debriefings were to obtain a variety of views from the experts performing the simulation. These debriefing were facilitated by the game management team. Given the set-up of the gaming exercises and the objectives of the research, the qualitative data revealed in the debriefing was expected to be particularly of interest. In particular to reveal mechanisms to enhance confidence in automation, and assessing the impact on the human roles.
- System data collection: data from the system was analysed such as the interaction between the different actors in the different LoAs. Performance data was considered as not relevant to this project.

Furthermore, the gaming method itself (both paper-based and platform-based) was also evaluated by means of pre- and post-questionnaires. This was done to assess the realism of the gaming aspects as perceived by the participants in order to reveal the value of the research and its outcomes in terms of Acceptability (of methodology) and Confidence (in results).

Both for the paper-based games and for the platform-based games, the feedback from the actors and the questionnaires was analysed mainly on a qualitative basis. The ratings provided through the post-run questionnaires are analysed making comparisons between the runs, using a within-subject analysis. A thorough statistical analysis was not possible given the limited number of runs and participants. Nevertheless, a graphical presentation of the ratings was provided as support to the qualitative feedback.

4) Scenario

The scenario features the impact of a non-severe capacity shortfall due to bad weather forecast. The bad weather is expected to be over Madrid TMA impacting the capacity of the airport and the TMA airspace. This happens during Madrid's high season, when unexpected events may disrupt operations and the agreed service level may fail to be met. The scenario encompasses the whole process from the detection of the imbalance until the approval of a solution to be implemented. More specifically, it consists of the following core processes:

- Build/Refine Reference Traffic Demand;

- Detect Airspace Demand Capacity Imbalance;
- Select/refine/Elaborate a DCB Solution at Network Level;
- Assess Network Impact of the DCB Solution;
- Start UDPP on SBTs.

The scenarios were defined three different time frames, the year 2020, 2035 and 2050 all with the expected traffic density for the year and a level of automation. Consequently, the roles and responsibilities of the actors change over these years. More details about the scenario can be found in [6, 10, 11]

5) Actors

The following actors take part in the scenario:

- Local Traffic Manager;
- Airspace Manager;
- Regional Network manager;
- Airport CDM Project Manager;
- Airspace User CDM Manager

However, given the number of experts available only three, the roles of Airspace Manager and Airport CDM Manager were played by the same actor. On the other hand, the UDPP process was not performed due to the lack of appropriate experts, so that the role of Airspace User CDM Manager was not played.

In total 3 participants, with their specific role of Local Traffic Manager, Regional Network Manager or CDM manager Madrid, participated both to the paper-based and to the platform-based gaming. The three male participants had a different nationality (British, Italian and Romanian). Two participants were between the age of 35-50 years old with approximately 25 years of experience. The other participant was between 20-35 years old and had approximately 10 years of experience.

B. Gaming Preparation

The main preparation activities performed were [10, 11]:

- Preparation of the simulation scenarios, including definition of the physical scenarios and events within each level of simulation, the definition of (catalogue of) solutions and cost model and the definition of the task, rules and gaming objectives of the actors (paper- and platform-based gaming);
- Preparation of the data gathering method including the questionnaires (paper- and platform-based gaming).
- Update of the simulation platform, CHILL. This activity started in parallel with the preparation of the paper-based phase. Information exchange was performed between both approaches, and outputs from the paper-based simulation were taken into account for the refinement of the platform (platform-based gaming).

C. Gaming Results and Discussion

1) Paper-based Gaming Results

Figure 1 shows the mean value of the responses of the participants to each question of the questionnaire and for the three scenarios. Since there were only 3 participants, it should be mentioned that the mean values have no statistical significance. However, the results were very useful for a preliminary trend analysis and for identification of the most interesting indicators and / or those who need further investigation. In other words, the results provided absolutely a support in the design of the platform-based gaming exercise.

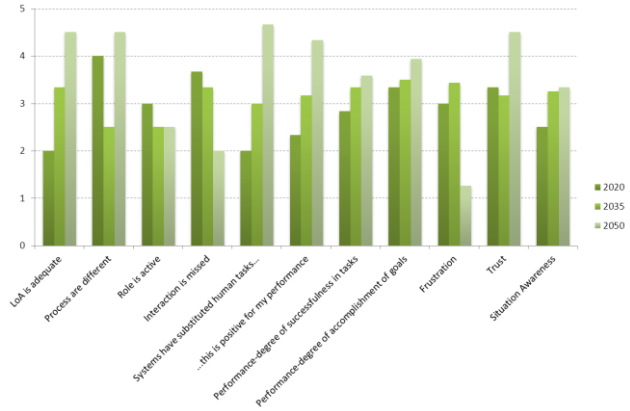

Figure 1. The effect of increasing LoA on the Human Actor (Paper-based games)

The main observations from Figure 1 are:

- The assumed LoA for 2020 was not adequate, and the one for 2050 was very adequate. The adequacy of the LoA for 2035 was in between that of 2020 and 2050;
- The difference in processes to be performed by the actors between 2020 and 2035 was considered small. Those between 2035 and 2050 were considered large;
- The obvious expectation that an increasing LoA results in less active roles is confirmed;
- Although the interaction between actors decreases with increasing LoA, it was decreasingly missed by the actors;
- The increasing substitution of human tasks by systems is considered increasingly positive for the actors' performance;
- The actors considered their performance to improve with increasing LoA;
- When going from a LoA of 2020 to that of 2035, the frustration increased and the trust in the system decreased slightly. However when subsequently going

to the LoA of 2050, both parameters the frustration decreased and the trust in the system increased significantly. This might be caused by the fact that some intermediate LoA (like that in 2035) was not well perceived: it is preferred that the human does (almost) everything (like in 2020 scenario) or that the system does everything (2050).

- The situational awareness (SA) increased with increasing LoA. However, the individual answers of the participants differed here. One of the actors had a high SA for those scenarios where he could perform tasks (2020, 2035) while for the fully automated scenario (2050) he indicated to have a very low SA. The other two participants did indicate to have an increasing SA with increasing LoA.

It should be mentioned that the questionnaire items were not independent but correlated and could therefore be considered as one construct, indicating a general measure of trust the actors feel working with the LoA under consideration. This construct, although also lacking statistical significance, indicated as well a positive trend with increasing LoA.

2) Platform-based Gaming Results

The questionnaire results of the platform-based gaming sessions have been grouped per validation objective so the change across the timeframes can be seen more easily.

Trust –It was seen that the level of trust is maintained as the level of automation increased. Feelings of system reliability, accuracy, usefulness and confidence increased over the periods of time. However, the NM did state that he would have liked to be able to perform a final "check" on the solution before accepting it.

Acceptance – The participants accepted the increasing levels of automation. The participants agreed that given the level of traffic in the 2035 and 2050 exercises, a high level of automation was required to be able to perform the work, but the NM felt more comfortable with the level of automation in 2035, where he had to review the solution and have several options available to him. The NM wanted to be able to review the solution, checking that everything is ok.

Situational Awareness – It was hard to determine a trend relating to situational awareness and its change as the level of automation is increased. However, the participants did state that their situational awareness was sufficient for the given automation and traffic density of that year. It is recommended that future CDB systems aiming at this LOA incorporate means for the actors to visualize the scenarios being proposed for implementation and the changes involved.

Teamwork –The participants found that the system did help them out more as the LOA increased, but that the interactions between participants decreased. The LTM did mention that the chat function worked well as a means of communication what he was doing to the other team members and was appropriate since the activities occurred during the planning phase of flight, not execution. It was also mentioned during the debrief that a key factor in the feeling of teamwork was that the participants were the same ones that performed the paper-based exercises and they assumed the same roles. This continuity led to a greater feeling of a team and gave a more realistic feeling of a group that works together often.

Workload – The workload was acceptable for the given automation and traffic density of that year, although there was a difference per participant in the amount of decrease in workload as the LOA increased. As there were only three participants, there can be no speculation as to the cause of this difference. During the final debrief, it was even suggested that since the automation is solving all the local DCB conflicts, the role of LTM approval of the solution could be incorporated into the duties of the NM.

Needs of possible future support tools – Many things were noted that should be taken into consideration when designing tools for this LOA environment. Most of them related to how the actors communicated (chat, telephone, etc.). As the populace in general becomes more and more accustomed to communicating textually (IM, SMS, etc.), the acceptance of this mode of communication in a CDB planning environment will increase and could be taken advantage of. Other needs revolve around the visualization of the solution and the actor's ability to see what has been implemented, if they so desire. Some proposals for improvements are:

- Ability to show multiple scenario data
- Ability to minimize the chat and collaboration window when needed.
- Predefined solutions should always be shown as reroutes to the NM.

3) The use of Gaming Technique

None of the three participants had previous Gaming Experience. However, no skepticism with respect to the technique was observed. Figure 2 shows the pre and post-gaming opinion about the suitability of the Gaming Technique for validation activities and the confidence in results.

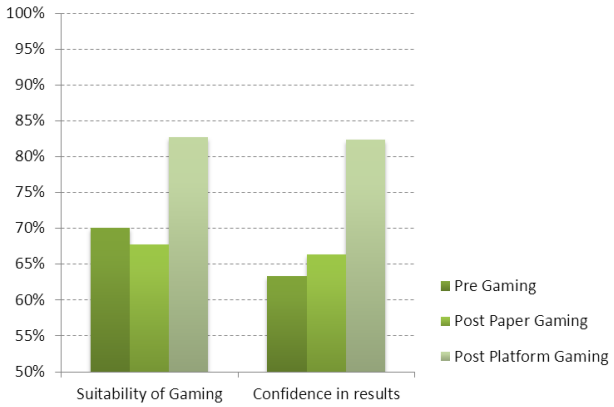

Figure 2. Usability of Gaming Technique

It can be seen that the differences between pre and post paper-based gaming opinions are relatively small. The belief in suitability of Gaming Technique for Validation decreased slightly. This can be attributed to the fact that the paper-based games do not "feel" realistic, as one of the participants remarked. Be that as it may, the confidence in the results increased slightly after having performed the games. After the platform-based games the opinions were – as expected – considerably more positive, mainly because of the higher degree of reality owing to using a platform. Another positive effect may come from the fact that the same actors as during the paper-based gaming session were participating in the platform-based gaming session, which was considered essential for the overall assessment by all three actors.

IV. CONCLUSIONS

The impact of increasing Level of Automation on the interaction between human actors and on their roles and responsibilities was assessed for the Airspace Organization and Management environment by consecutive paper-based and platform-based gaming sessions.

The paper-based games were used to refine the scenario, to provide the platform-based gaming with the right input and to do a preliminary assessment of the effect of higher levels of automation on the involved ATM actors. The platform-based gaming sessions were completely focussed on assessing that effect.

Both gaming sessions were performed with only three actors and a limited number of runs. Therefore, a thorough statistical analysis was not possible and the questionnaires were analysed mainly on a qualitative basis. Nevertheless, a graphical presentation of the ratings was provided and a trend analysis could be done..

The results showed that trust and acceptance of the system increased or was at least maintained with increasing LoA and that the situational awareness and the workload remained at sufficient levels. The teamwork decreased with increasing LoA which took away the feeling of doing it together and recommendations with respect to possible future support tools were especially related to how the actors communicate.

ACKNOWLEDGEMENT

The work presented here is carried out in the project ADAHR, which is a consortium of Ingenieria de Sistemas para la Defensa de España (Isdefe), Centro de Referencia de Investigación, Desarrollo e Innovación ATM in Spain (CRIDA), Deutschen Zentrums für Luft- und Raumfahrt (DLR) and National Aerospace Laboratory of the Netherlands (NLR). ADAHR is co-financed by EUROCONTROL on behalf of the SESAR Joint Undertaking in the context of SESAR Work Package E. This paper reflects only the authors' views and EUROCONTROL is not liable for any use that may be made of the information contained here.

REFERENCES

[1] T. Truman and A. de Graaf, "Out of the Box: Ideas about the future of air transport", ACARE - European Commission, November 2007

[2] ACARE Background Document "Aeronautics and Air Transport: Beyond Vision 2020 (towards 2050)", European Commission, January 2010

[3] High Level Group on Aviation Research, "Flightpath 2050 – Europe's Vision for Aviation", European Commission, EUR 098 EN, 2011

[4] The Federal Transportation Advisory Group FTAG "Vision 2050 – An Integrated National Transportation System", February 2001

[5] E.02.09-ADAHR Roles and Responsibilities D1.1, V00.01.00, 17/08/2011

[6] E.02.09-ADAHR Scenarios Description D1.2 Ed00.01.00, 27/04/2012

[7] J.P. Rafidison, "Episode 3 Final Report and Recommendations", D2.5-01, V3, March 2010. (website: http://www.episode3.aero/)

[8] P. Sanchez-Escalonilla, et al, "Collaborative Planning Results and Consolidation". Episode 3 project, 2009.

[9] R. Casar Rodríguez and R. García Lasheras, "Use of role games in automation assessment", 1st International Conference on Application and Theory of Automation in Command and Control Systems (ATACCS 2011), pp. 100-103, 26-27 May 2011, Barcelona, Spain

[10] E.02.09-ADAHR Exercise Plan Paper-Based Gaming ENV1 D2.1 Ed00.00.10, 20/06/2012

[11] E.02.09-ADAHR Exercise Plan Platform-Based Gaming ENV1 D2.2 Ed00.00.10, 27/07/2012

[12] Lee, K.K. Kerns, K., Bone, R. 2001 Development and Validation of the Controller Acceptance Rating Scale (CARS): Results of Empirical Research

[13] Waag, W.L., Houck, M. R., 1994. Tools for assessing situational awareness in an operational fighter environment. Aviation, Space and Environmental Medicine.

[14] Hart, S.G., Staveland, L.E. 1988 Development of the NASA-TLX, Results of empirical and theoretical research, in Human Mental Workload edited by P. A. Hancock & N. Meshkati Amsterdam: North Holland Press. (pp. 239-250)

[15] SHAPE (Solutions for Human-Automation Partnerships in European ATM), http://www.eurocontrol.int/humanfactors/gallery/content/public/docs/SHAPE_questionnaires/SHAPE%20User%20Guide%20v0.1.pdf