Twelfth USA/Europe Air Traffic Management Research and Development Seminar (ATM2017)

Effectiveness of the application of the Human Performance Assessment Process in SESAR 1

Sharing Lessons Learnt

Sonja Biede
Human Factors and Ergonomics
Airbus
Toulouse, France
sonja.straussberger@airbus.com

Renee Pelchen-Medwed
Directorate ATM, Performance & Methods
EUROCONTROL HQ
Brussels, Belgium
renee.pelchen-medwed@eurocontrol.int

***Abstract*—This paper introduces a comprehensive and systematic process to analyze the effectiveness of the application of the Human Performance Assessment Process (HPAP), which had been developed during the SESAR 1 program. An effectiveness evaluation was conducted over a selection of indicators addressing the process as well as the final outcome. These indicators considered the coverage of the HPAP, the involvement of human factors (HF) specialists, the number of identified issues and recommendations, as well as the status of closing these assessments at a defined level of maturity. The analysis shows that the process was widely applied, but variations in methods and tools of application were observed. It can hence be concluded that there is still some need for harmonization. In order to prepare its effective transition towards SESAR 2020, this analysis allows providing recommendations for the next steps.**

Keywords-Human performance; SESAR; Air Traffic management; Assessments; Human Factors; Air/Ground; Shared process;

I. INTRODUCTION

From 2009 until 2016, the so-called SESAR (Single European Sky Aviation Research) Development phase (renamed into SESAR 1) took place. The scope of this phase was to define in detail operational concepts and develop their technical enablers. Those enablers had initially been agreed in the SESAR definition phase, with the objective to improve ATM performance. For this purpose, SESAR has adopted a top down performance based approach, and target performance criteria relating to safety, as well as capacity and efficiency, have been defined.

If human factors (HF) are not adequately considered, the proposed system performance benefits in terms of safety, capacity and efficiency may not be achieved. It was therefore essential that all operational and technical projects adopted a consistent approach to HF integration, in order to ensure that the relevant HF findings were aggregated and linked back to the relevant target performance criteria. [1]

New operations are impacting numerous actors that cooperate with each other, hence, coherence and aggregation needs to be ensured on a system-wide basis. In addition, numerous partners across air and ground industrial and operational domains are collaborating to produce the new ATM systems. To support such an activity, not only did a shared air/ground methodology between stakeholders not exist, also, the collaboration throughout a multinational and multi-organizational environment represents a very specific context, as each local environment has its own HF culture.

Hence the need for a new HF integration process for SESAR was identified [1], and a systematic human performance assessment process (HPAP) was developed from 2010 to 2014. Initial versions were applied from 2012 on, resulting in a good coverage of the application of this process by the end of SESAR 1. In order to evaluate the extent of the effectiveness of this process, a study was conducted and is presented in this paper. With the HPAP becoming an integral part for SESAR 2020, the study allows the establishment of recommendations in order to improve the assessment of the impact on Human Performance (HP) and share this experience with the wider ATM community.

So far, no studies are known that have systematically evaluated the effectiveness of the application of a HF process in the design of complex systems and operations. This paper is an attempt to fill this gap.

II. BACKGROUND

A. Inputs for SESAR HPAP

To date, no situation existed in the ATM domain where collaborators developed a shared process for the application across multiple industries, organizations, and nations. Ergonomic standards such as ISO 9241 [2] provide guidance on how to consider human factors during a design phase and are dedicated to industrial organizations without providing further detailed recommendations.

Within the ATM domain, different organizations had developed processes at local premises. EUROCONTROL made a first step with the development of the Human Factors case [3] for the European air traffic control (ATC) community.

The Human Factors Case has been developed to provide a comprehensive and integrated approach to ensure that the design of a technical, human, and/or procedural system can deliver desired performance improvements. The Human Factors Case (HF Case) was launched in August 2004, supported by the first edition of the HF case document. The primary focus of the original HF Case was for high level application in European Air Traffic Management (EATM) projects within EUROCONTROL. Lessons learned using the HF Case in a number of EUROCONTROL projects led to the publication of revised versions of the document at a later stage. Today the HF case is used widely within European air navigation service providers (ANSPs).

Certain organizations such as NATS, DFS, Skyguide, Frequentis, ENAV and Thales, have identified HF groups or focal points. Some of these organizations have also additionally a mature HF integration process. Accordingly the way how human factors specialists are involved in design, varies. The participation ranges from a deep involvement by accompanying a project throughout its entire life cycle till little involvement by assessing an advanced product. Some other organizations do not have identified HF specialists at all. For the airborne side, based on certification requirements for large aircraft, the human factors department of Airbus has a systemised contribution to aircraft design integrated in the company's process, whereas other aircraft manufacturers may be organized differently.

In the US, the HF Workbench of the FAA [4] provides recommendations for HF activities. Similarly, there are international norms in the military domain, the UK ministry of defense has defined standards on Human Factors integration [5], and processes also exist for the nuclear domain and railway. It is however known, that there is variability in the applications of the various references.

B. Developing HPAP[2]

This and the next section (B. and C.) are based on Chalon et al. 2012 who described the principles of the HP assessment process [1].

The HP assessment process [6] informs the design and development of an operational concept through the identification of recommendations and / or requirements necessary to prevent or mitigate any potential negative impacts of the proposed concept on human performance. The new HP assessment methodology for SESAR built on existing HF integration processes used in ATM, such as the EUROCONTROL HF Case [3] and inputs provided by the contributing partners across the industry, and adopted a systematic argument and evidence approach. A human performance argument can be understood in this context as 'a human performance claim that has to be proven'. The HP arguments are structured into four different HF areas, namely:

- Human Roles,
- Human and the System,
- Teams and Communication and
- Transition Factors (which includes general acceptability, training, skills and competencies, and staffing).

Each of the four main high level arguments is broken down into lower level, more detailed arguments, to form an argument tree. For instance, '*Argument 1-The role of the human is* consistent *with human capabilities and limitations'* is subdivided into **'***Argument 1.1-Roles and responsibilities of human actors are clear and exhaustive'*, '*Argument 1.2-Operating methods are clear and support human performance'*, and '*Argument 1.3-Human actors can achieve their tasks (under normal, abnormal and degraded modes of operations*'. In turn, Argument 1.1 is further broken down into ''*1.1.1 The description of roles & responsibilities cover all affected human actors*" and.''*1.1.2 The description of roles & responsibilities cover all tasks to be performed by a human operator*" and so forth.

Fig. 1 shows the different levels of the arguments for a selected example.

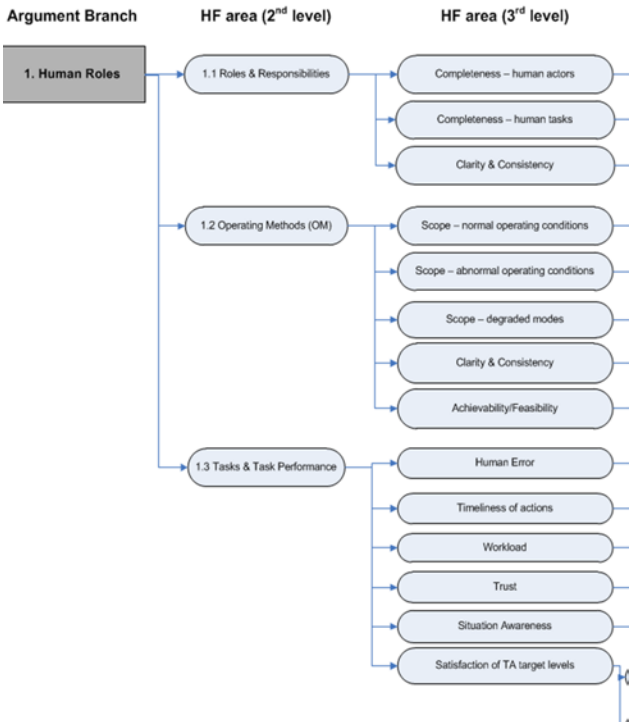
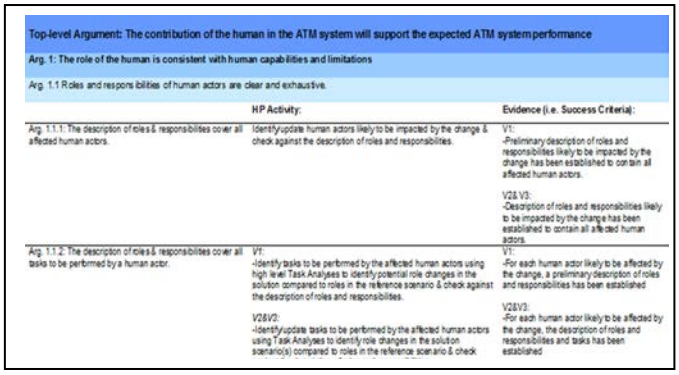


Figure 1 Graphical display of the argument structure: Argument tree

Identify applicable sponsor/s here. *(sponsors)*



Top-level Argument: The contribution of the human in the ATM system will support the expected ATM system performance		
Arg. 1: The role of the human is consistent with human capabilities and limitations		
Arg. 1.1 Roles and responsibilities of human actors are clear and exhaustive.		
	HP Activity:	Evidence (i.e. Success Criteria):
Arg. 1.1.1: The description of roles & responsibilities cover all affected human actors.	Identify/update human actors likely to be impacted by the change & check against the description of roles and responsibilities.	V1: -Preliminary description of roles and responsibilities likely to be impacted by the change has been established to contain all affected human actors. V2& V3: -Description of roles and responsibilities likely to be impacted by the change has been established to contain all affected human actors.
Arg. 1.1.2: The description of roles & responsibilities cover all tasks to be performed by a human actor.	V1: -Identify tasks to be performed by the affected human actors using high level Task Analyses to identify potential role changes in the solution compared to roles in the reference scenario & check against the description of roles and responsibilities. V2&V3: -Identify/update tasks to be performed by the affected human actors using Task Analyses to identify role changes in the solution scenario(s) compared to roles in the reference scenario & check	V1: -For each human actor likely to be affected by the change, a preliminary description of roles and responsibilities has been established V2&V3: -For each human actor likely to be affected by the change, the description of roles and responsibilities and tasks has been established

Figure 2 Arguments linked to HP activities and evidence

For each argument, the 'evidence' required to satisfy that specific argument is defined. Each argument is linked to a set of suggested HP activities. Fig. 2 illustrates an extract of the argument structure including the HP activity and the evidence.

The HP assessment process methodology includes the development of templates for the HP assessment plan, the HP assessment report and the HP log. While the HP log is a living document maintained throughout the project by updating issues and results, the HP assessment plan and report are handed over at a certain point to the project team and delivered with the project documentation. As support to the HP assessment process a toolbox - the eHP repository [7] - was developed to provide a catalogue of means to conduct the HP activities. The eHP Repository is a collection of standard human performance methods, tools, guidelines and techniques as well as tools specifically developed within SESAR 1 to support ATM design. These tools include guidance for automation best practice [8], competence and training assessment [9], and guidance for information presentation [10].

C. Application of HPAP

The HP assessment process consists of four steps:

- Step 1 Understand the ATM Concept
- Step 2 Understand the HP implications
- Step 3 Improve and validate the concept
- Step 4 Collate findings and produce the HP assessment report.

Fig. 3 illustrates these four steps and their link to the HP deliverables.

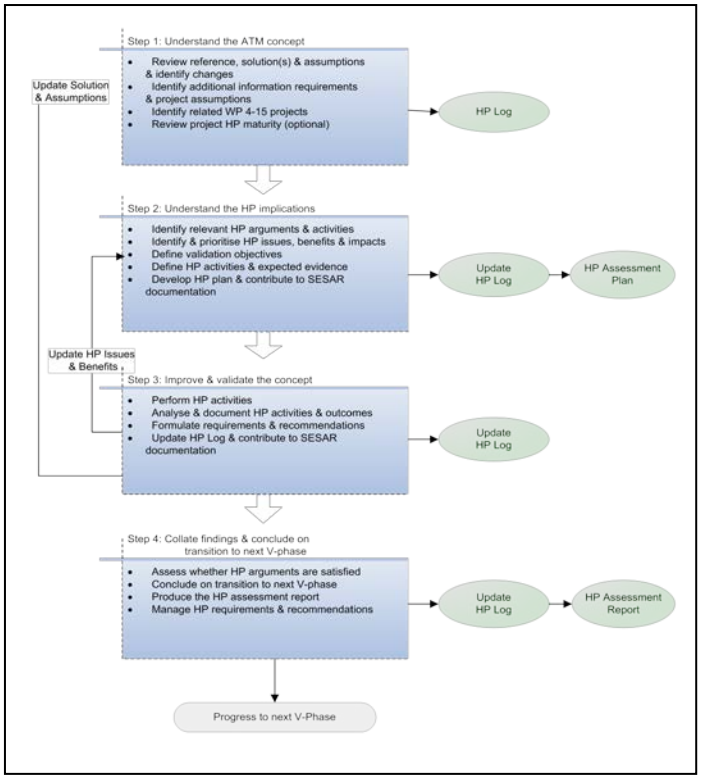


Figure 3 Steps of the HPAP

In Step1, an understanding of the ATM concept is acquired by reviewing relevant project documentations and in-depth discussions with the concept developers. HP related assumptions are defined and agreed and the HP level of maturity determined using the maturity criteria provided in the HP reference material [6]. In Step 2 once an understanding of the ATM concept is built, the high level questions relating to the argument structure are raised to identify the relevant high level HP arguments that need to be considered and addressed in the system design and validation process. When the relevant high level HP arguments are determined, the HP argument structure is then used to identify the lower level arguments that need to be considered. Based on the low level arguments the required HP activity(ies) that need to be performed can be identified (e.g. task analyses, real time simulations, workshops). For each argument and related HP activity, the evidence required to satisfy that specific argument is defined. Based on this information, an initial HP Assessment Plan can be developed. However, in order to select, refine and prioritize the HP activities, HP issues and benefits have to be reviewed. This can be done by interviewing relevant stakeholders as well as reviewing previous work conducted on related concepts.

The issues and benefits are described in terms of their impact on human and system performance and rated in terms of priority level by the stakeholders. Where possible, actions to address or mitigate the issue are identified. On the basis of the results of the HP issues / benefits analyses, the initial identification of relevant HP activities is refined and the proposed activities prioritised. Each activity recommended in the HP assessment plan is described in detail, in terms of: the

arguments and issues to be addressed and hence the HP objectives of the activity; the required evidence (or success criteria); and the general planning of the activity i.e. timeline, resource and approach, so that they can be understood and discussed with the project manager and project team members including the safety and validation teams. The HP assessment plan is a key element to be integrated in the global validation plan. After the conduct of the HP activity in step 3 the results are documented and transformed into recommendations and requirements. These recommendations and requirements are then discussed, consolidated and prioritised with the concept development and safety team and in a further step fed back into the project documentation supported by the HP assessment report in step 4.

III. EFFECTIVENESS OF HPAP IN SESAR PROJECTS

This paper describes a comprehensive and systematic process to analyze the effectiveness of the application of the above described HPAP.

A. Evaluation objective

The objective of this evaluation is to identify to what extent HPAP is used and if it is used in a consistent and effective manner across SESAR projects. This evaluation allows identifying recommendations for a more efficient HPAP application in the coming phase of the SESAR program.

B. Method

The study is focused on an effectiveness evaluation, which determines whether an initiative has had the intended effect on outcomes. As only few solutions continue to the deployment phase by the end of SESAR 1, indicators for the effectiveness on the application of HPAP cannot be retrieved from operational feedback. Due to SESAR 1 being a research and development initiative, outputs of the application of the process are expected to achieve a certain level of maturity. These maturity criteria stem from the European Operational Concept Validation Methodology (EOCVM) [11], as the E-OCVM was the validation process adopted within SESAR. The maturity criteria applicable for SESAR1 were V1 (feasibility) to V3 (pre-industrialization).

For this reason, a retrospective evaluation was conducted to assess if the application of the HPAP was effectively achieved related to the expected maturity level. Based on the data available at the end of SESAR 1, effectiveness was measured by analyzing how the HPAP was applied to ensure the final outcome, i.e. HF adequately integrated in the operational and technical design by having resolved HP issues& benefits.

C. Development of effectiveness measures

The first question is on the measure to be used to adequately assess effectiveness. Generally, effectiveness refers to the extent to which goals are accomplished, or in terms of Sproles [12], effectiveness refers to “a measurement of how well the problem has been solved”. The identification of these measures is inspired from research initiatives in the field of organizational effectiveness [13].

Within the domain of organizational effectiveness, different approaches may be used. Amongst the most structuring approaches, the goal-oriented approach, the system-based approach, the strategic constituency approach and the competing values approach are mentioned.

A goal-focused approach is looking at the achievements of targets to define essential operating objectives. Hence, effectiveness is assessed by first determining the task objectives of the system or organization, and then to develop criterion measures to assess how well the objectives are achieved. A systems approach to effectiveness considers the organization as an open system, whereby the organization acquires inputs, engages in transformation processes, and generates outputs. In that context a measure of effectiveness establishes the degree to which an organization realizes its goals under a given set of conditions. A so-called strategic constituency approach deals with the effect of the organization on the main stakeholders and their interests, and effectiveness refers to the minimal satisfaction of all of the strategic constituencies of the organization. Finally, the competing values approach assumes that diverse preferences can be consolidated and organized into a holistic approach.

Based on these inputs, a framework was developed to characterize effectiveness measures on both a goal-focused and a system-focused approach to ensure a complete view. A proper selection of the measurement attributes (e.g. be operational, understandable, non-redundant) is essential to this process. The strategic constituency approach can be considered as partly addressed by having the HPAP deliverable approved by all partners involved in the project as well as the SESAR Joint Undertaking as final customer. In addition, it became part of the Project Management Handbook for SESAR 2020.

Based on all possible indicators that could be determined to assess such completeness, indicators were selected in relation to four dimensions that represent both the goal-oriented and the system-oriented approach (see Table I).

The overall effectiveness is characterized by comparing the different indicators with regard to the final outcome at a certain level of maturity. The final task objective from a goal model perspective may be considered in assuring the human factors integration into design. This is expressed by a complete application of the HPAP which results in resolving all issues and recommendations by adequate activities.

TABLE I. INDICATORS FOR HPAP EFFECTIVENESS

Dimension	*Category*	*Indicator*
System view	Inputs	• Number of partners and HF specialists involved
	Process	• HPAP applied in operational focus area • OFA requiring combined air/ground assessments compared to separate air or ground assessments • Form of HF specialist involvement (continuous vs. punctual)
	Outputs	• Available Form of Evidence (HP reports/logs produced) • Arguments addressed (Roles, System, Teams, Transition)
Goal view	Outcome	• Number and Status (open vs. closed) of Issues& Benefits at Airside/Groundside/Air-Ground • Number and Status (Open vs. closed) of Recommendations Airside/Groundside/Air-Ground

D. Data collection

A document analysis was conducted based on the evidence produced by HF specialists, which is available in the following form:

- Contribution to validation documentation
- Production of HP plans and HP reports
- Production of HP logs

This document analysis was supported by information which had been gathered in a shared spreadsheet along the SESAR 1 program to monitor the application of the HPAP. The units of reviews in this spreadsheet were the so called operational focus areas (OFA). The operational focus areas, is a limited set of dependent operational and technical improvements, comprising specific interrelated operational improvements (OI) designed to meet specific performance expectations of the ATM Performance Partnership [14]. Within one OFA more than one SESAR 1 project was clustered and one project was very often conducting several validation exercises.

E. Results

*1) Coverage of HPAP:*For the application of the HPAP, a total of 32 OFAs were reviewed and descriptive measures provided to characterize the conducted assessments. Sixty-six percent of these OFAs were identified as requiring human performance assessments. OFAs were not selected when an evolution only concerned technical solutions without HF impact or in case human performance was already addressed within other scopes.

However, due to the evolution of the program, not all projects where an assessment would have been required, were effectively conducting one.

TABLE II. HPAP COVERAGE OF OPERATIONAL FOCUS AREAS

Application of HPAP in OFA	*Number*
Total number of OFA	**32**
HPAP applicable	**21**
Applied	13
Not applied	8
HPAP not applicable	**11**

This resulted in a total of 13 OFAs effectively applying the HPAP. Table II lists the number of OFAs that applied or did not apply the HPAP. Several reasons for not applying the HPAP while it would have been applicable can be listed:

- Unavailability of adequate human resources,
- Late communication on applicable reference material for assessments,
- Difficulty to integrate an external person to a project that is already going-on,
- Operational specialists or validation specialists performed the HF assessment and analyses, or

In few cases, the documentation containing the evidence was no longer accessible due to technical issues.

When starting to collect the available HP evidence in OFAs for the analysis, it was identified, that finally several independent HP assessments took place within many OFAs. These assessments were either conducted for a specific validation exercise, or across validation exercises for one project, but not often aggregated for the concept. The fact that there were reasons that no HP assessment was conducted while it would have been needed, led also to the case where one HP assessment was conducted within an OFA but not the entire OFA could be assessed and therefore certain areas/projects in the OFA did not apply the HPAP. In the following, these types of assessments are referred to as separate air or ground assessments. Hence, these assessments may either address an air or a ground solution, or an assessment of an air/ground concept only from part of the concept (either air or ground). A combined air/ground assessment refers to an assessment that has an explicit impact on actors both at air and at ground side and these impacts are addressed at the same time.

The table III lists the OFAs where HP assessments were conducted and if they were split into several assessments at a sub-unit level. The sub-unit refers to projects within these OFAs. Therefore it cannot be assumed that in those cases the performed HP assessments cover the entire OFA.

In order to be able to characterize the effectively undertaken HP assessments, the following data are analyzed at the level of the OFA sub-units (projects). This analysis leads to a total of 24 assessments conducted across the considered 13

OFAs. Some OFA or OFA sub-units were not considered for this analysis even if HF assessments were conducted as part of the validation activity, as they were not yet aligned with the HPAP or did not integrate this process.

TABLE III. OPERATIONAL FOCUS AREAS TYPE AND NUMBER OF ASSESSMENTS AT SUB-UNIT LEVEL

OFA	*Number of HP assessments within OFA sub-units*
AIM/MET	1
LVPs using GBAS	1
Airport safety nets	3
Enhanced Runway Throughput	5
Business and Mission Trajectory	1
ASAS Spacing	3
ASAS Separation	1
Enhanced ACAS Operations	1
Enhanced Arrival & Departure Management in TMA and En Route	3
Integrated Surface Managemen	1
Airport Operations Management	1
Remote Tower	2
Airspace Management and AFUA	1

*2) Effectiveness related to available inputs: T*he availability of HF specialists represents the major precondition for being able to perform Human Performance assessments. HF specialists followed the HPAP process after having been informed by the so-called SESAR HP Support and Coordination function. Based on a subjective estimation, an average number of partners involved in projects are six. Within 21 OFA sub-units, the identified number of HF specialists involved ranges from one to three with an average of two involved human factors specialists. Ten of these sub-units are considered as being either an airborne (for example Surface Alerts for pilots) or a ground solution (Safety nets for Airport Vehicle drivers) or were assessed on only a part of a solution instead of a complete air/ground solution. In these cases, usually only one HF specialist was involved. In the case of projects covering both air and ground aspects of solutions, at least two human factors specialists – one with ground one with airborne expertise - were involved.

3) Effectiveness related to process: The coverage of the OFA by HPAP had already been described in the previous section. Regarding the way how the HF contribution was organized, in 84% of the assessments HF specialists were continuously involved in the project.

4) Effectiveness related to output: Regarding the way to produce the evidence, the HPAP suggests the production of an HP log, an HP plan and an HP report. The HP log is an excel sheet that lists the HP issues, the (proposed) mitigation to the issue and the validation results with recommendations and requirements. The HP plan and the HP report are word documents expressing the same content like the HP log but in a different format and therefore on a different level of detail.

While the detail of the HP log is defined via the issue analysis, the HP plan and the HP report is very often used to detail the validation plan and the validation report and therefore relate often more to the validation exercises than to the overall concept assessment. The data show, 13 HP logs are available, and in the remaining cases HP plans and reports were produced. In some cases, HP logs were directly integrated also together with HP plans and reports in project validation documentation like the validation plan and the validation report.

TABLE IV. ARGUMENTS COVERED

Argument Categories	*Number of OFA sub-units**
Roles & System	2
Roles, System & Transition	2
Roles, System & Team&Com	3
Roles, System & Team&Com & Transition	15

*In 2 OFA sub-units, no argument allocation was done.

Table IV illustrates which type of arguments was addressed in the various OFA sub-units. From the 22 sub-units where the information is available, in 66% all argument types are addressed, whereas in a reduced number, either team-related arguments or transition-related arguments were not considered relevant.

TABLE V. DESCRIPTIVE STATISTICS* OF HP ISSUES&BENEFIS AND RECOMMENDATIONS PER ASSESSMENT CATEGORY

Segment	*Issue/Recommendation*	*Total*	*Air*	*Ground*	*Air/ Ground*
Air/ Ground	Issues	**62** *4-278* 11	**7** *0-22* 10	**43** *0-229* 9	**13** *0-34* 10
	Recommendations	**40** *8-51* 8	**8** *1-19* 6	**16** *2-43* 6	**14** *3-51* 7
Air	Issues	**16** *5-36* 4	**12** *5-19* 4	0	**19** *0-17* 3
	Recommendations	**13** *11-15* 3	**11** *8-15* 3	0	**2** *0-6* 3
Ground	Issues	**32** *8-86* 9	0	37 *8-86* 7	1 *0-4* 7
	Recommendations	**21** *2-65* 8	0	**22** *2-65* 7	1 *0-4* 6

***Average**, *min-max*, sub-units considered; indicated for all statuses

5) Effectiveness related to outcome: Finally, the above mentioned inputs, processes and intermediate outputs result in the production of a final outcome. Closed HP Issues & Benefits and HP Recommendations are considered as such major outcomes to characterize the effectiveness of HPAP, as they present an indication if pertinent issues have been studied to assess a concept, and they indicate if any discovered issues were adequately solved.

Table V presents an overview of the number and type of HP issues & benefits as well as recommendations separately for the air, the ground or the connected air/ground segment in order to illustrate the variation in the elements addressed. For each category, averages across the involved sub-units, as well as the minimum and maximum counts are shown. As the available data varies, also the number of analyzed OFA sub-units varies across these different elements. Consequently, averages presented per sub-cell may not be aligned with the total averages.

Finally, the scope of the HPAP is to ensure that all issues are adequately studied and recommendations addressed before an entry into industrialization and deployment.

TABLE VI. STATUS OF HP ISSUES&BENEFIS AND RECOMMENDATIONS FOR INITIAL AND ADVANCED MATURITY LEVELS

Table column subhead	***Number of OFA-Subunits***	***Percentage of Closed Elements***
Issues (maturity V3)	9	30% (0-64%)
Issues (maturity V1-V2)	6	20% (0-90%)
Recommendations (maturity V3)	4	24% (0-53%)
Recommendations (maturity V1-V2)	5	0%

To illustrate this, table VI shows the percentage of HP Issues & Benefits as well as recommendations for lower and higher levels of maturity that were closed by the end of SESAR 1 based on sufficient evidence. For each category, the average percentage as well as the minimum and maximum percentage observed over the various OFA sub-units is indicated.

IV. DISCUSSION

As the results show, it is possible to identify measures for the effectiveness of the HPAP along different dimensions to gather an overall picture of the situation. There are however some limitations with respect to the applied approach. The classification of some elements during the document analysis may be subject to inter-rater difference and impact the reliability. This was however as far as possible addressed by a clear definition of the addressed concepts.

Despite an OFA being expected as one central aggregation entity due to the linked concepts, it was noted that for some OFAs several independent HP assessments were developed. It can be explained by the difficulty of clearly allocating concept elements to OFAs, which also would allow setting clear borders for the HP assessments. This illustrates also the difficulty of identifying the adequate level of analysis. Consequently, in many cases HP assessments were conducted at validation exercise level, and only few at concept level.

Positively, the results showed that overall a high level of involvement of HF specialists occurred across projects. However, compared to the number of overall partners the number of involved HF specialists in the projects was still rather low. We have identified a reduced HF specialist involvement in projects that only address either the ground or the air side or sub-units of air/ground assessments despite an important number of partners involved. This bears the risk that issues that have - despite the nature of the project - an impact on the air/ground interaction are overlooked or if identified not addressed adequately. This risk was generally covered in explicitly defined air/ground projects, where a higher number of HF specialists was involved. However, the way to collaborate between HF specialists may have been very much driven by a focus on the production of documents rather than focusing on a shared conduct of the HP assessment throughout the design phase.

One efficient way to ensure that mitigations to air/ground issues are taken on board, further developed and deployed, are workshops involving technical and operational ground expertise as well as airborne expertise to gain a mutual understanding and the buy in on both sides. These workshops should be driven and facilitated by the HPAP and the specialists applying it.

Various factors impact the produced outputs. First of all, due to the distributed nature of the activity over many organizations, the information was not reaching all project partners at the same time. Due to the different significance of the HF roles in different organizations the importance of the input provided by the HF specialists to the projects was perceived differently. Secondly, there is a large variety in the creation of HP documentation. The HP log concept is a new concept as such, and it was observed in some cases, that HF specialists preferred to apply familiar methodology by focusing on documents which remained centered on exercise level rather than on concept level.

The difference in the number of defined issues can be explained in several ways. It can arise due to

- The nature of the OFA and the arguments addressed as a consequence
- Dependent on the specialist and resources available. Whereas some HF specialists conduct a profound analysis, other analyses are remaining on a rather generic level. Also, the background of HF specialist varies; consequently, HP assessments may be approached differently.
- The input available for the issue analysis. If the resources for a systematic approach like a task analysis or use cases are available a more detailed issue analysis can be conducted.
- The maturity level of the concept. In the beginning of a project life cycle the issues are defined at a higher level and quantitatively less than in V2 or beginning of V3 while at the end of V3 and V4 fewer issues

should be defined. At the end of V3 the majority of the issues should be closed.

- Industrial property rights might not allow publishing all issues identified in the concept assessment.

Finally it can be discussed if having a higher number of issues also refers to conducting a more profound level of analysis, rather than having well covered the study of a certain concept.

The assumption that the majority of issues and recommendations should be closed by the end of V3 level of maturity, could not be confirmed by the analysis of the data presented in this paper. After interpreting these results it can be questioned if the concept maturity was always at the same level as the declared maturity level of the projects. In addition, it would be interesting to analyze if certain types of issues would rather appear on specific levels of maturity. For example, one would expect that a higher number of issues related to the transition argument would appear with an advanced level of maturity. Also, one would expect that a certain amount of HP issues of an early level of maturity would be resolved with more detailed concept definition. Such a comparison of issues and recommendations across levels of maturity was not possible to be conducted in the scope of this evaluation, it may however be considered for future studies.

Despite the high number of unresolved issues it may be acceptable to continue towards industrialization and deployment, under the condition that a further assessment to confirm the feasibility of the concept in the deployment environment is conducted. This local implementation assessment should resolve relevant issues and implement relevant recommendations and requirements. In case this step is neglected a risk of implementing a concept that is not entirely acceptable to the user is taken, or even worse unknown risks are not controlled.

Finally, this study is limited by its retro-perspective nature. Other evaluation measures such as cost are not considered here, as accessible data cannot be considered being very reliable. To improve the quality of such a study, the use of additional methods like interviews or detailed content analysis of the way of writing the HF elements can be imagined. The assessment could also have been conducted in form of a formative evaluation accompanying the project which would have allowed gathering additional data to assess the value of HPAP and predict the output already during the development to ensure that the adequate HF issues are addressed.

The study presented in this paper allows providing an insight in how HP is addressed in a multinational complex research and development initiative.

V. CONCLUSION AND WAY FORWARD

In summary, this paper presented the assessment of the effectiveness of the application of the Human Performance Assessment Process in SESAR 1. Based on a combination of measures applied, it can be concluded that the HPAP was effectively applied in SESAR projects to the extent possible, as it allowed

- To deploy the adequate competence in form of HP specialists who used a common process to ensure coherence across studies,
- To identify and study HP issues & Benefits, and
- To ensure HP recommendations are managed.

This is considered a major achievement, taking into consideration the complexity of the context and the collaboration over a multi-domain (different industries as well as different operational domains), multi-national and multi-organizational context. As the evaluation of human performance is vital to achieve the future ATM performance, measures will need to be undertaken to ensure a more homogeneous application of HPAP for the next phases. Even though the availability of HPAP is an essential precondition, it may not be sufficient to ensure the quality of the outcome. Additional mitigation means will have to be considered.

To ensure that the contribution of the human in the ATM system will support the expected ATM system performance and with that, that the future ATM concepts will be fit for purpose, the human factors integration from the initial design phase to deployment have to be guaranteed by HF specialists. There is a need to ensure adequate understanding of HP reference material [6] and why it is important to define issues in a certain way to reduce the variety in producing air and ground issues. For the airborne side, HPAP is aligned with the airborne process as required by the certification requirements in CS25.1302 [15] which is defined in the same spirit.

In order to resolve the problem with the variety of issues and produced recommendations, the need to develop a shared understanding on the level of detail for HP assessments is identified. Often assessments were done at exercise level, but exercises very often dealt only with one or two specific concept elements but not with the entire concept. This needs to be better raised for HF specialists to start the awareness assessment at the adequate concept level, including associated technical solutions. Having identified this need the requirement for the application of a HP methodology that facilitates the consolidation of HPAP results on a higher aggregation level can be specified. An attempt was done in SESAR 1 with building the HP case [16] and this attempt will be followed up in SESAR 2020. However this consolidation methodology might reveal that in future the effectiveness of the method application cannot be measured by the number of issues and recommendations but needs other indicators.

HPAP performance is vital to ensure also its future systematic application in projects. However, measures are not systematically developed to ensure that this performance will be achieved by an effective application of HP during the project phase. Also, the risk has to be managed that HPAP may be applied just to show conformance with requirements, without guaranteeing a certain quality of the content of the data produced. No system exists today to ensure that HP content is reviewed according to similar criteria regarding the content that is captured. In order to better harmonize the approaches and

identify potential need to ensure coherence, a need for a formative evaluation approach may be considered. Such an approach would allow through identification of suitable measurement criteria to ensure the final quality of the output.

SESAR 2020 organization has partly acted on this problem by integrating HP in the program's project management principles. Furthermore in SESAR 2020 HP is considered to be a key focus area and is integrated as such into the performance framework. Still, it requires having a lead coordination from project manager point of view. After optimizing the way to apply the process, new challenges arise also that HP shall not only remain focused on providing support for assessments, but HP shall become a real driver to support the ATM change.

Finally, another challenge remains the integration of HF specialists. Having a common framework is useful, as it allows reducing time for defining collaboration, and one can directly focus on content. But it requires experienced HF specialists due to the complexity of the context.

HPAP also allows closing the gap between research and development and operation, however, it was not yet systematically applied for such a transition. The continuous check of the effectiveness of applying HPAP may also be based on data from operational deployment. Evaluation measures will however have to be prepared accordingly.

ACKNOWLEDGMENT

This work is based on the outputs achieved by the partners contributing to the human performance support activity conducted within the SESAR 1 project 16.06.05 Human performance and coordination function (EUROCONTROL, Airbus, Thales, NORACON, DFS, NATMIG, Indra, ENAV, NATS, AENA, Staff association and Airspace User representatives). We would like to thank colleagues from DFS, Skyguide, Novair and EgisAvia for their support during the production of this paper.

REFERENCES

[1] C. Chalon-Morgan et al., Development of an Argument and Evidence based Human Performance Assessment Process for SESAR. EAAP 30 proceedings part 4, 2012.

[2] ISO 9241-210, Ergonomics of human-system interaction Part 210: Human-centred design for interactive systems, Geneva,International Organization for Standardization, 2016.

[3] EUROCONTROL, The Human Factors Case: Guidance for Human Factors Integration Edition 2.0,2007.

[4] FAA Workbench. Available at: http://www.hf.faa.gov/workbenchtools/

[5] MOD Human Factors integration Process (HFI Process). UK Ministry of Defence, 2015.

[6] SESAR Human Performance Assessment Process V1 to V3 – including VLDs; 16.06.05 D27; edition 00.01.00,2016.

[7] eHP repository https://ext.eurocontrol.int/ehp/

[8] SESAR Guidelines for Addressing HP Automation Issues. WP 16.05.01 - Deliverable 04, 2013.

[9] A. Heintz, I. Lévêque, W. Georg, S. Biede, T. Leoff and A. de Beaufort, Interdisciplinary Training for Task and Responsibility Shifts in Future operational ATM Procedures and Systems.Proceedings of the 32nd conference of the European Association for Aviaion Psychology, Cascais, Portugal, 2016.

[10] SESAR Updated Generic SESAR Information Presentation Guide; 16.05.03, D06, edition 00.01.00, 2013.

[11] European Operational Concept Validation Methodology (E-OCVM). Version 3.0. Volume 1 and 2. February 2010.

[12] N. Sproles, The difficult problem of establishing measures of effectiveness for command and control: A systems engineering perspective, Systems Engineering, 2001 (4,2), 145-155, 2001.

[13] D. Nadkarni and M. D'souza, A Study of the Success Value of the Four Approaches to Organizational Effectiveness in 18 Companies in the Indian Service and Industry Sectors, SIES Journal of Management, Vol. 11(1), 59-76, 2015.

[14] https://ext.eurocontrol.int/lexicon/index.php/Operational_Focus_Area retrieved 21.01.2017

[15] EASA, CS-25 Certification Specification for Large Aeroplanes, Brussels.

[16] SESAR Guidance for HP Case building - Initial version,16.04.01, D11; edition 01.00, December 2012.