INDUSTRIAL RESEARCH

PJ.W2 PROSA SESAR JOINT UNDERTAKING

D4.2.020-PJ.10-W2-96 AG-TRL6-Final TS-IRS-Part IV-HPAR

Deliverable ID	**D4.2.020**
Dissemination Level:	**PU**
Project Acronym	**PROSA**
Grant:	**874464**
Call:	**H2020-SESAR-2019-1**
Topic:	**Separation Management and Controller Tools**
Consortium coordinator:	**DFS**
Edition date:	**16 February 2023**
Edition:	**00.02.00**
Template Edition	**02.00.02**

Founding Members

EUROPEAN UNION EUROCONTROL

SESAR JOINT UNDERTAKING

Founding Members

EUROPEAN UNION EUROCONTROL

Authoring & Approval

Authors of the document

Name/Beneficiary	Position/Title	Date
NATS	Solution 96 AG HF	20.10.22
NATS	Solution 96 AG HF	15.11.00

Reviewers internal to the project

Name/Beneficiary	Position/Title	Date
Skyguide	Solution 96 AG Member	05.11.00
LDO	Solution Lead	05.11.00
Hungarocontrol	Safety Expert	05.11.00
Skysoft/ATM	Solution 96 AG Member	05.11.00

Approved for submission to the SJU By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
Skyguide	Solution 96 AG Member	02.12.00
LDO	Solution Lead	02.12.00
Hungarocontrol	Safety Expert	02.12.00
Skysoft/ATM	Solution 96 AG Member	02.12.00

Rejected By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date

Document History

Edition	Date	Status	Author	Justification
00.00.01	15.09.2022	Draft	NATS	First Version
00.00.02	24.10.2022	Draft	NATS	Review of first version and additions to results section
00.00.03	28.10.2022	Draft	NATS	Finalised draft version for partner review
00.01.00	30.11.2022	Final	NATS	Final version for submission

00.02.00	16/02/2023	Final	NATS	Word Template removed from the front page

PJ.10-W2 PROSA

SEPARATION MANAGEMENT AND CONTROLLER TOOLS

This document is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 874464 under European Union's Horizon 2020 research and innovation programme.

Abstract

This document contains the Human Performance (HP) assessment report for the Attention Guidance concept- PJ.10-W2-Sol.96 Attention Guidance (AG). It consists of the HP assessment plan, the results of the HP activities conducted according to the HP assessment process, newly identified issues and the HP recommendations & requirements. It corresponds to the completion of the four steps of the Human Performance assessment process, namely: Step 1 – Understand the concept: Baseline, Solution and Assumptions, Step 2 – Understand the Human Performance Implications, Step 3 – Improve and Validate the concept and Step 4 – Collate findings & conclude on transition to next phase.

The PJ.10-W2-Sol.96 AG focuses on the Human Machine Interface (HMI) of the Controller Working Position (CWP). The project is driven by an enabler for Controller productivity enhancements by Attention Guidance at the CWP/HMI.

This report describes the findings of the validation exercise held at Skyguide in June 2022. The findings suggest that the Attention Guidance solution benefit overall human performance by reducing visual clutter and helping controllers to focus their scan in the most relevant areas of their HMI. Further work should ensure that with increased levels of automation the human operators are kept in the loop appropriately.

Table of Contents

List of Tables

List of Figures

1 Executive Summary

This Human Performance Assessment Report (HPAR) summarises the results of the completed HF activities for PJ10-Sol.96 Attention Guidance completed in accordance with those set out in the Human Performance Assessment Plan (HPAP).

The validation activity was executed in the Skyguide Operation, a very high En-Route operational environment above FL355. Attention Guidance introduced a fade-out algorithm which was expected to drive controllers' visual and cognitive focus to more relevant tracks (i.e. requiring interaction) and less on non-conflictual flights (i.e. unlikely to require interaction).

The main HP areas from the HPAP of interest were:

- Maintaining of situation awareness with increased levels of automation
- Workload management with increased levels of automation
- Task distribution between human and machine
- Appropriate alerting

Evidence was gathered using both quantitative and qualitative methods. This included questionnaires at end of run and end of participation in the study, system data (eye tracking) and interviews/ debriefs.

Results showed that participants were able to perform their tasks appropriately. They reported good levels of situation awareness and were able to manage their workload to acceptable levels with the introduction of Attention Guidance. The visual scanning data shows a trend towards more focussed scanning and increased fixation times. These results are based on a small sample size and need to be interpreted with caution at this stage. During debriefings controllers reported that they found the solution very promising and would like to see further development on it to eventually introduce it into live operation as they perceived a great decrease in visual clutter. This led to high user acceptance reports and trust in automation.

Going forwards, HP work should focus on ensuring that all information is displayed at the right time and that alerts and task drivers are in line with a coherent design philosophy. Whilst very promising results in the area of user trust in automation were achieved it remains a key area to drive development work and eventual operational deployment.

2 Introduction

2.1 Purpose of the document

The purpose of this document is to describe the result of the activities conducted to date according to the Human Performance (HP) assessment process [1] for Solution 96 AG. It also aims to present the HP activities completed to address the HP arguments relevant for the Solution during the simulation exercise. HP recommendations and requirements are presented as the result of activities conducted in this project.

2.2 Intended readership

The intended audience for this document is the following:

- SESAR 2020 PJ.10-W2-96 AG partners
- The key stakeholders targeted by the Solution, i.e.
    - Airspace Users who will benefit from the deployment of the fade-out algorithm in upper En Route airspace included in a very high complexity environment.
    - Air traffic Controllers who will be directly impacted by the Solution to enable the fade-out algorithm permanently or temporary in a very high complexity environment.

2.3 Structure of the document

This document is Part IV of the overall TVALP. It describes the expected changes within the defined HP arguments, identified issues and benefits and the derived HP objectives and HP activities that need to be considered within the validation exercises.

The document consists of six chapters:

- **Chapter 1: Executive Summary.** Provides a summary of the key information and elements contained in the Human Performance Assessment Report.
- **Chapter 2: Introduction.** Provides the document purpose and structure, intended readership and the background activities to be considered.
- **Chapter 3: The Human Performance Assessment Process: Objective and Approach.** Provides the general background for the Human Performance Assessment process.
- **Chapter 4: Human Performance Assessment.** Provides the results of the validation exercises outlined in the Human Performance Assessment Plan as well as HP recommendations and requirements.
- **Chapter 5: References**. Lists all references and applicable documents that have been considered in the production of the HPAR.

2.4 Acronyms and Terminology

Term	Description
Human Factors (HF)	HF is used to denote aspects that influence a human's capability to accomplish tasks and meet job requirements. These can be external to the human (e.g. light & noise conditions at the work place) or internal (e.g. fatigue). In this way, "Human Factors" can be considered as *focussing on the variables that determine Human Performance*.
Human Performance (HP)	HP is used to denote the human capability to successfully accomplish tasks and meet job requirements. In this way, "Human Performance" can be considered *as focussing on the observable result of human activity in a work context*. Human Performance is a function of Human Factors (see above). It also depends on aspects related to Recruitment, Training, Competence, and Staffing (RTCS) as well as Social Factors and Change Management.
HP activity	An HP activity is an evidence-gathering activity carried out as part of Step 3 of the HP assessment process. An HP activity can relate to, among others, task analyses, cognitive walkthroughs, and experimental studies.
HP argument	An HP argument is an HP claim that needs to be proven through the HP Assessment Process.
HP assessment	An HP assessment is the documented result of applying the HP assessment process to the SESAR Solution-level. HP assessments provide the input for the HP case.

HP assessment process	The HP assessment process is the process by which HP aspects related to the proposed changes in SESAR are identified and addressed. The development of this process constitutes the scope of Project 16.04.01. It covers the conduct of HP assessments on the Solution-level as well as the HP case building over larger clusters of Solutions.
HP benefit	An HP benefit relates to those aspects of the proposed ATM concept that are likely to have a positive impact on human performance.
HP case	An HP case is the documented result of combining HP assessments from Solutions into larger clusters (SESAR Projects, deployment packages) in SESAR.
HP issue	An HP issue relates to those aspects in the ATM concept that need to be resolved before the proposed change can deliver the intended positive effects on Human Performance.
HP impact	An HP impact relates to the effect of the proposed solution on the human operator. Impacts can be positive (i.e. leading to an increase in Human Performance) or negative (leading to a decrease in Human Performance).
HP recommendations	HP recommendations propose means for mitigating HP issues related to a specific operational or technical change. HF recommendations are proposals that require additional analysis (i.e. refinement and validation). Once this additional analysis is performed, HF recommendations may be transformed into HF requirements.
HP requirements	HP requirements are statements that specify required characteristics of a solution from an HF point of view. HP requirements should be integrated into the DOD, OSED, SPR, or specifications. HF requirements can be seen as the stable result of the HF contribution to the Solution, leading to a redefinition of the operational concept or the specification of the technical solution.

Table 1: Acronyms and Terminology

Acronyms	
AG	Attention Guidance
HPAP	Human Performance Assessment Plan
FL	Flight Level
TVALP	Technical Validation Plan

TOD	Top Of Descent
XFL	Exit Flight Level
SSR	Secondary Surveillance Radar
ATM	Air Traffic Management
HMI	Human Machine Interface
ATCO	Air Traffic Control Officer
EC	Executive Controller
PC	Planner Controller
SPR-INTEROP/OSED	Safety and Performance Requirements – Interoperability Requirements / Operational Service and Environment Definition
SESAR	Single European Sky ATM Research Programme
VHC	Very High Complexity
CWP	Controller Working Position

Table 2: Acronyms

3 The Human Performance Assessment Process: Objective and Approach

The purpose of the HP assessment process described in detail in Human Performance Guidance document [1] is to ensure that HP aspects related to SESAR Solution technical and operational developments are systematically identified and managed. The SESAR HP assessment process uses an 'argument' and 'evidence' approach. An HP argument is an 'HP claim that needs to be proven'. The aim of the HP assessment is to provide the necessary 'evidence' to show that the HP arguments impacted have been considered and satisfied by the HP assessment process. This includes the identification of HP requirements and recommendations to support the design and development of the concept, which will be defined in the HP Assessment Report.

The HP assessment process is a four-step process. Figure 1 provides an overview of these four steps with the tasks to be carried out and the two main outputs (i.e. HP plan and HP assessment report). In addition, an HP Log is recommended to be maintained throughout the lifecycle of the Solution in which all the data/ information obtained from all HP activities conducted as part of the HP assessment is documented. This HP Log is a living document and is continuously updated and / or added to as the SESAR Solution progresses.

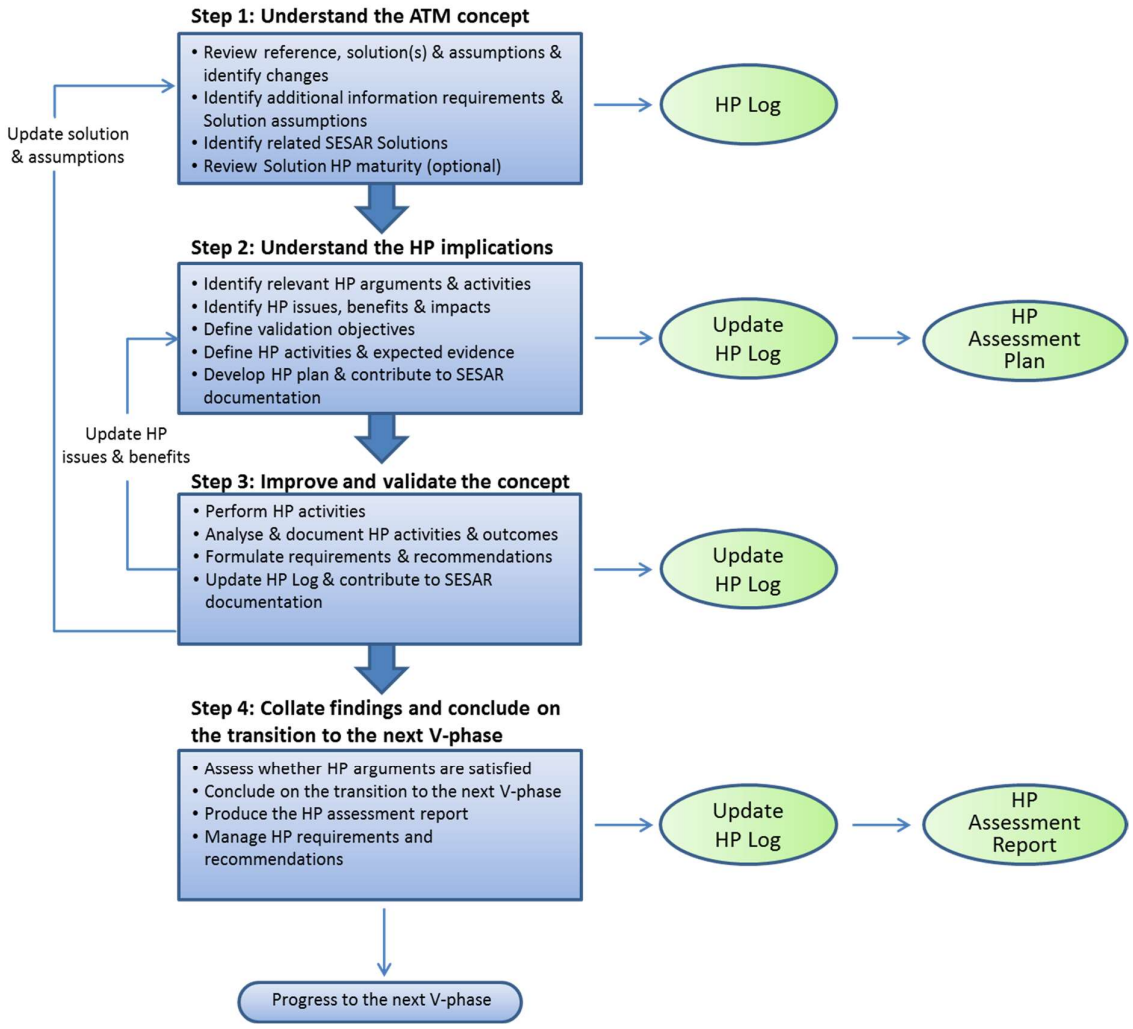


Figure 1 Human Performance Assessment Process

4 Human Performance Assessment

4.1 Step 1 Understand the ATM concept

4.1.1 Description of reference scenario

The baseline, respectively the reference traffic scenario is derived from real air traffic data from Skyguide's Geneva Upper Airspace 2017.
The simulation within the reference scenario ran without the attention guidance functionalities and was used to address the En-Route operational environment. This baseline was used to compare against the Solution scenario.
For further details of the reference scenario refer to section 5.1.4.1 of Part I of the TVALP.

4.1.2 Description of solution scenario

The solution scenario is the same as the reference scenario except for the added logic of fade-out with the following evolutions:

- Profiles extended until the XFL for the detection of conflicts.
- Implementation of the fade-out algorithm for sectors above level 355.
- Information of the ATCO that an action is needed when the track is in fade-out in case of:
    - transfer of Control,
    - Top of Descent 1 (Tod1) and Top of Descent 2 (Tod2) has been reached,
    - ECAT conflict
    - Reception of an E-coordination that do not imply a conflict with another flight,
- Alert of the ATCO that a flight is newly in its scan after a previous fade-out in case of:
    - Minimum lateral distance below 20 NM and flight is vertically intercepting another one,
    - RAM alerts,
    - CLAM / EHS CLAM alerts,
    - Emergency tracks or special SSR codes (e.g.: 7500, 7600 or 7700),
    - Reception of an E-coordination implying a conflict with another flight

For further details on the solution scenario refer to section 5.1.4.2 of Part I of the TVALP.

4.1.3 Consolidated list of assumptions

Identifier	Title	Description	Justification	Impact on Assessment
EXE-PJ.10-96-AG-TRL6-01	Realistic traffic	The scenario contains a realistic amount of traffic	Representativeness of the exercise	High
EXE-PJ.10-96-AG-TRL6-01	Training and competencies	Human Performance	To validate the algorithm it is important that the controllers are familiar with the scenario and tools	High

4.1.4 List of related SESAR Solutions to be considered in the HP assessment

Not applicable.

4.1.5 Identification of the nature of the change

The following table systematically identifies and captures the nature of the change that results due to the introduction of the AG concept in terms of, the ATM actors impacted as well as the potential changes to their work. This assessment forms the basis of the planned HP activities.

HP argument branch	Change & affected actors
1. ROLES & RESPONSIBILITIES	
1.1 ROLES & RESPONSIBILITIES	It is not expected that Attention Guidance will introduce any changes to the roles and responsibilities of controllers. There is a chance that changes to responsibilities could be required for operational supervisors. Attention Guidance could require high level control across the ops room, meaning that supervisors should be in the position to enable or disable the functionality across all positions. However, the controllers at the individual workstations might need to be able to turn off the functionality even if the supervisor enabled it. This will need to be investigated during validation exercises.

1.2 Operating Methods	New operating methods will need to be developed as part of the concept. The system will display tracks in different states which require individual responses not existing in today's operation. Operating methods will be required for when to use the functionality and when to turn off the algorithm. For example, severe weather scenarios could require the functionality to be switched off as all tracks will be relevant during weather avoiding scenarios. Another scenario might be very low traffic load which could lead to (almost) all tracks being presented in fade-out status. Generally, the threshold (e.g. traffic count, workload,…) for appropriate usage envelope should be identified during validations. In current operations controllers can choose whether they want to probe a clearance before issuing it. Given that tracks which used to be 'fade-out' could require interaction when not anticipated by the controller they might need to rely on the probe functionality in order to not miss any conflicts. Therefore, a change in procedures could be required to mandate the use of the probe functionality. Operating methods during abnormal operating conditions will need to be revisited in general (e.g. weather, emergencies, etc.). New operating methods need to address degraded modes of the system. They should cover appropriate controller actions when the functionality or parts of it fail.
1.3 Tasks	Visual scanning as a task will be impacted by the introduction of Attention Guidance. Controllers' visual search could be narrowed down (not geographically but in terms of numbers of aircraft). This could lead the controller into thinking that they have got more spare cognitive capacity than they actually have given the number of aircraft in their sector. A new task will be to detect & acknowledge/reject 'intermediate' status of tracks. This is a new HMI functionality which requires controllers to learn a new task. It will need to be carefully considered what will

	happen in case the controller fails to detect or react appropriately to an 'intermediate' status. Dealing with abnormal operating situations might potentially require controllers to do established tasks differently. For example, an aircraft which was in 'fade-out' status could declare an emergency. The controller will need to quickly integrate the aircraft into their mental picture before executing any actions to support the aircrew. In general, controllers will need to learn how to trust the system and understand the reasons for particular tracks being in 'fade-out' or 'normal' status to be able to act appropriately.
2. HUMAN & SYSTEM	
2.1 ALLOCATION OF TASKS (HUMAN & SYSTEM)	Attention Guidance will introduce changes to the way that tasks are distributed between human and machine. Those changes affect decision making, monitoring and conflict detection. In current operations the controller has to assess whether an aircraft entering their sector needs to be interacted with (i.e. clearances need to be issued). With Attention Guidance in place the system will do this check for the controller and suggest whether a track should be foreground or background based on defined parameters. The system will present tracks in different states which the controller will need to understand and assess its appropriateness. An additional task on the human will be to ensure that all the tracks are in the correct state and if required switch between them. As new alerts are introduced as part of Attention Guidance appropriate reactions will be required by the controllers to ensure situations can be resolved safely. In current operations controllers are not mandated to probe clearances before issuing them. This might need to change as they are required to integrate 'fade-out' tracks very quickly if the traffic scenario changes unexpectedly. Controllers' decision making will be affected by this change in working methods. More reliance will be put onto the system when it comes to conflict detection.

	To ensure trust in the system controllers need to be able to understand how the system assesses in what state the tracks are presented (e.g. updates are based on clearances but also changes in flightpath independent from controller instructions - emergency/weather avoidance). In summary, it is expected that controllers "forget" about fade-out tracks (i.e. not include into regular scan) as per design of this functionality. This means that the machine takes on a more active role by taking part of the monitoring task away from the controller. Given that this is a significant shift in task distribution between human and machine the effects of this on overall human performance will need to be investigated carefully.
2.2 PERFORMANCE OF TECHNICAL SYSTEM	Information provided to the ATCO needs to be accurate. This means that there will need to be checks of the accuracy of the algorithm's adherence to set parameters (e.g. display of tracks closer than 20nm). Information provided to the ATCO needs to be timely. This means that any lag in the system needs to be reduced to an acceptable level (see limits below): 1-2 s Completion of user input to display of error indication 2 s Request for next page of information to completion of one-page change, Completion of user input to completion of simple process. Completion of display manipulation request to completion of display change (e.g. open a window; zoom). 5 to 10 s Completion of user input to completion of frequently performed complex operation. >10 s Completion of user input to completion of infrequent, complex process.
2.3 HUMAN – MACHINE INTERFACE	New HMI elements will be introduced to support controllers' focus. This creates new modes which need to be considered carefully and their suitability assessed to avoid potential mode confusion. Generally, all scenarios where a track changes its state could potentially cause controller confusion and therefore inappropriate controller actions. All these situations will need to be assessed during validation.

	There is no change to input devices, but the display of information will change as well as an increase of potential interactions with the tracks (e.g. acknowledgments/rejection of status changes). In Skyguide controllers can only use their mouse for system inputs. Attention Guidance will not impact input devices. Effects on support information and conflict detection will need to be assessed. Currently there are interpersonal differences in the way that controllers display their speedlines (not at all, turn on/off, displayed permanently, etc). Given that speedlines will be a crucial indication for 'intermediate' states there could be a requirement for a standardised way to have them displayed. New alerts will be introduced by AG for example flashing (for one minute) if a controller fails to acknowledge track status changes and indications when an aircraft reaches top of descent. All these new indications will need to be assessed during validation.
3. TEAMS & COMMUNICATION	
3.1 TEAM COMPOSITION	No change to team composition is expected by the introduction of Attention Guidance.
3.2 ALLOCATION OF TASKS	No major changes are expected. However, it needs to be explored whether supervisors should be able to control the functionality on a higher level. There could be situations (e.g. bad weather) that require AG to be switched off. This could be the responsibility of the supervisor(s). The allocation of tasks between controllers should not be affected by AG.
3.3 COMMUNICATION	Potential changes in communication between EC and PC will need to be assessed as both ATCOs can use AG independently on the same sector. Additionally, as mentioned above the communication between controllers and supervisors might be affected when it comes to decisions about appropriateness of using the functionality or not. Furthermore, if the functionality can be turned on/off at individual workstations this will need to become part of the handover procedure.
4. HP RELATED TRANSITION FACTORS	

4.1 ACCEPTANCE & JOB SATISFACTION	Attention Guidance is expected to be acceptable if it supports controllers' focus and contributes to a perceived reduction in workload and increase in efficiency. It will be unacceptable if controllers don't trust the algorithm which decides whether a track should be displayed as foreground or background. There could also be issues if reaction to abnormal situations is perceived to be more demanding/confusing than in today's operation. This needs to be managed carefully and explored during validation exercises.
4.2 COMPETENCE REQUIREMENTS	No significant changes to competence requirements are expected as a result of the introduction of Attention Guidance.
4.3 STAFFING REQUIREMENTS & STAFFING LEVELS	Attention Guidance is not expected to affect staffing requirements or levels.
4.4. RECRUITMENT AND SELECTION	No changes to recruitment and selection are expected
4.5. TRAINING NEEDS	Controllers will need to be trained in the changes that Attention Guidance introduces (HMI, procedures, etc.) but no overall change to ATCO training is foreseen.

Table 2: Description of the change

4.2 Step 2 Understand the HP implications

4.2.1 Identification of relevant arguments, HP issues & benefits and HP activities

Arg.	Issue ID	HP issue / Benefit	HP/Valid. Obj. ID	HP validation objective	Activity
1.2	HFI-ARG1.2-10.96-TRL6-001	Procedures to reflect response to track status - AG will present tracks in different states to the controller. As a result, each status will require specific interaction which controllers are not familiar with in today's operation. There is a risk for controller confusion as to how to interact with tracks in certain states (especially 'intermediate' status).	OBJ-ARG1.2-10.96-TRL6-001	Validation activities should assess actors understanding of appropriate interactions with aircraft for each possible track status.	Real Time Simulation
1.2.2	HFI-ARG1.2-10.96-TRL6-002	Procedures to reflect abnormal weather situations - The use of Attention Guidance might not be appropriate in all weather and traffic scenarios. When aircraft start to avoid weather, all tracks in the sector suddenly become relevant. When traffic levels are very low all tracks could be in fade-out status. If the switch between using AG and not using it is not clearly defined there is a risk for controllers using the functionality inappropriately.	OBJ-ARG1.2-10.96-TRL6-002	Validation activities should assess actors understanding of the use/switch off of AG in abnormal weather situations and varying traffic scenarios.	Real Time Simulation
1.2	HFI-ARG1.2-10.96-TRL6-003	Procedures around probing of clearances - AG might require a change in the way that controllers use the probe functionalities on their workstations. Currently probing clearances is up to the ATCO and not	OBJ-ARG1.2-10.96-TRL6-003	Validation activities should assess the level of appropriate use of probing functionalities in the AG context.	Real Time Simulation

		mandatory. Given that a/c could move from 'fade-out' to becoming relevant quickly ATCOs might not have them integrated in their mental picture as quickly as the situations requires. There is a risk of issuing inappropriate clearances due to a lag of updating their mental model.			
1.2.2	HFI-ARG1.2-10.96-TRL6-004	Procedures around alerting philosophy - AG might introduce new alerts. Clear procedures around the appropriate responses to alerts need to be introduced. Otherwise there is a risk of controllers not reacting in the desired manner which could lead to safety critical situations.	OBJ-ARG1.2-10.96-TRL6-004	Validation activities should assess the actors understanding and appropriate reaction to alerts in the AG context.	Real Time Simulation
1.2.2	HFI-ARG1.2-10.96-TRL6-005	Procedures around dealing with emergencies - AG could add a layer of complexity when ATCOs are required to handle an aircraft emergency. If the system fades out a track and this aircraft has an emergency the controller needs to be able to integrate the track into their mental model very quickly. If a track is displayed as 'normal' and declares an emergency the ATCO has to take all surrounding tracks ('normal', 'fade-out' 'intermediate') into consideration when formulating a plan for the aircraft.	OBJ-ARG1.2-10.96-TRL6-005	Validation activities should assess the actors understanding of procedures around aircraft emergencies.	Real Time Simulation
2.1.2	HFI-ARG2.1-	AG will introduce changes to the task distribution between human and machine.	OBJ-ARG2.1-10.96-TRL6-001	Validation activities should assess the understandability of	Real Time Simulation

	10.96-TRL6-001	Their appropriateness will need to be assessed during validation exercises. The main change will be from human assessment whether a track needs interaction to the machine performing this assessment. The human actors will need to understand the logic of the algorithm to decide whether the assessment is appropriate and can be trusted.		the system algorithms and the suitability of task allocation between human and machine	
2.1.2	HFI-ARG2.1-10.96-TRL6-002	AG could affect the interaction with existing functionalities on the CWP. Wider use of probing functionality could be required. This would transfer the task of pre-clearance conflict detection from the human to the machine.	OBJ-ARG2.1-10.96-TRL6-002	Validation activities should assess the understandability of the system algorithms and the suitability of task allocation between human and machine, especially in the context of the use of probe functionality.	Real Time Simulation
2.	HFI-ARG2.1-10.96-TRL6-003	The introduction of AG might require a more standardised way of displaying information on the CWP. Currently ATCOs can chose if and how to display speed lines depending on personal preference. With AG speed lines will become an (even more) important indication for situation awareness. They will be highlighted/fade-out when a track goes into 'intermediate' status. ATCOs need to able to pick up on this cue regardless of their standard way to display speedline	OBJ-ARG2.1-10.96-TRL6-003	Validation activities should assess the suitability of ways to display speed lines	Real Time Simulation

2.1.5	HFI-ARG2.1-10.96-TRL6-004	There is a potential risk that controllers overestimate their cognitive capacity using AG. Tracks in fade-out status are still the controller's responsibility but could be neglected from their mental model as they are less salient and don't require immediate interaction. In case of e.g. abnormal weather developing, AG could get switched off which then means a sudden increase in salient tracks which could overwhelm the controller. Aircraft emergencies could lead to the same outcome (e.g. mass diversion).	OBJ-ARG2.1-10.96-TRL6-004	Validation activities should assess the threshold at which switching off AG leads to controllers being overwhelmed and struggle to integrate prior fade-out tracks into their mental model.	Real Time Simulation
2.1.6	HFI-ARG2.1-10.96-TRL6-005	It will be important for controllers to understand the logic of the algorithm which determines whether a track is displayed as 'normal' or 'fade-out'. If the underlying logic is not clear to the user, they are less likely to trust it.	OBJ-ARG2.1-10.96-TRL6-005	Validation activities should assess the level of trust into the system.	Real Time Simulation
2.2.1	HFI-ARG2.2-10.96-TRL6-001	To support user trust, it will need to be checked whether the system performs accurately against pre-defined parameters such as separation minima.	OBJ-ARG2.2-10.96-TRL6-001	Validation activities should assess the accuracy of the system algorithms	Real Time Simulation
2.2.2	HFI-ARG2.2-10.96-TRL6-002	To support user trust, it will need to be checked whether the system performs according to defined lag maxima.	OBJ-ARG2.2-10.96-TRL6-002	Validation activities should assess the adherence to pre-defined lag maxima.	Real Time Simulation
3.2.1	HFI-ARG3.2-	It is not clear yet whether the use of AG needs to be managed on a higher level than individual workstations. There could	OBJ-ARG3.2-10.96-TRL6-001	Validation activities should identify the appropriate task distribution between	Real Time Simulation

	10.96-TRL6-001	be a necessity for the supervisors to be responsible for deciding when AG can be used and when it needs to be switched off. Tasks to enable this must be split appropriately between controllers and supervisors to avoid confusion. Consequently, controllers might want to keep the possibility to switch AG off at their workstation even though the supervisor has enabled it.		supervisors and controllers in relation to turning on/off the Attention Guidance functionality.	
3.3.1	HFI-ARG3.3-10.96-TRL6-001	Following on from HFI-ARG3.2-10.96-TRL6-001 it needs to be ensured that efficient communication is possible between controllers and supervisors when it comes to decision making about using AG or switching it off.	OBJ-ARG3.3-10.96-TRL6-001	Validation activities should assess whether an acceptable level of team communication can be established between controllers and supervisors.	Real Time Simulation

Table 4: HP Arguments, related HP issues and benefits, and proposed HP activity **Step 3 Improve and validate the concept**

4.3.1 Description of HP activities conducted

HP Activities include Situation awareness measurement, Mental Workload measurement, Usability/ User confidence measurement and others as shown in the tables below.

ACTIVITY 1.	
Description	Situation Awareness Measurement
Related Arguments	ARG 1.2; ARG 2.1

HP objectives	OBJ-ARG1.2-10.96-TRL6-001 OBJ-ARG1.2-10.96-TRL6-002 OBJ-ARG1.2-10.96-TRL6-003 OBJ-ARG1.2-10.96-TRL6-004 OBJ-ARG2.1-10.96-TRL6-005 OBJ-ARG2.1-10.96-TRL6-001 OBJ-ARG2.1-10.96-TRL6-002 OBJ-ARG2.1-10.96-TRL6-003 OBJ-ARG2.1-10.96-TRL6-005
Tools/Methods selected out of the HP repository	Assessment of situational awareness | HP repository (eurocontrol.in
summary of the HP activity	Situation Awareness measures from HP repository used after each to assess whether acceptable levels of situation awareness ca achieved.

Table 3: Description of Activity 1

ACTIVITY 2.	
Description	Mental Workload Measurement
Related Arguments	ARG 1.2; ARG 2.1
HP objectives	OBJ-ARG1.2-10.96-TRL6-001 OBJ-ARG1.2-10.96-TRL6-002 OBJ-ARG1.2-10.96-TRL6-003 OBJ-ARG1.2-10.96-TRL6-004 OBJ-ARG2.1-10.96-TRL6-005 OBJ-ARG2.1-10.96-TRL6-001 OBJ-ARG2.1-10.96-TRL6-002

	OBJ-ARG2.1-10.96-TRL6-003 OBJ-ARG2.1-10.96-TRL6-005
Tools/Methods selected out of the HP repository	Assessment of workload | HP repository (eurocontrol.int)
summary of the HP activity	Mental Workload measures from HP repository to be used after each run to assess whether acceptable levels of workload can be maintained. Real Time Simulation exercises completed November 2021.

Table 4: Description of Activity 2

ACTIVITY 3.	
Description	Reaction Time Measurement
Related Arguments	ARG 1.2
HP objectives	OBJ-ARG1.2-10.96-TRL6-004 OBJ-ARG1.2-10.96-TRL6-005
Tools/Methods selected out of the HP repository	CWP logs to show reaction times to alerts
summary of the HP activity	Duration of alert display measured to understand latencies in reaction time which are expected to provide additional objective situation awareness data.

Table 5: Description of Activity 3

ACTIVITY 4.	
Description	Usability/User Confidence Measurement
Related Arguments	ARG 2,1; ARG 2.2
HP objectives	OBJ-ARG2.1-10.96-TRL6-006 OBJ-ARG2.2-10.96-TRL6-001 OBJ-ARG2.2-10.96-TRL6-002
Tools/Methods selected out of the HP repository	Assessment of Acceptance | HP repository (eurocontrol.int)
summary of the HP activity	Usability/ user confidence measures from HP repository to be used after each day to assess whether AG is acceptable to the end user.

Table 6: Description of Activity 4

ACTIVITY 5.	
Description	Visual Scan Measurement
Related Arguments	ARG 1.2
HP objectives	OBJ-ARG1.2-10.96-TRL6-004 OBJ-ARG1.2-10.96-TRL6-005
Tools/Methods selected out of the HP repository	Tobii Visual Scanning equipment

summary of the HP activity	Controller scan patterns will be collected using Tobii eye tracking equipment to understand the influence of AG onto visual scanning and to provide another objective measurement for situation awareness.

Table 7: Description of Activity 5

ACTIVITY 6.	
Description	Teamwork and Communication Measurements
Related Arguments	ARG 3.2; ARG 3.3
HP objectives	OBJ-ARG3.2-10.96-TRL6-001 OBJ-ARG3.3-10.96-TRL6-001
Tools/Methods selected out of the HP repository	Assessment of teamwork and communication | HP repository (eurocontrol.int)
summary of the HP activity	Teamwork and communication measures from HP repository to be used after each day to assess if and how AG impacts those HP areas.

Table 8: Description of Activity 6

ACTIVITY 7.	
Description	Debriefings
Related Arguments	

HP objectives	All HP Objectives
Tools/Methods selected out of the HP repository	Debriefing Template
summary of the HP activity	A debriefing template will be followed to gather qualitative feedback on all aspects of AG.

4.4 Step 4 Collate findings & conclude on transition to next phase

4.4.1 Summary of HP activities results

Data collected during the simulations at Skyguide in Switzerland consisted of Questionnaire data for both End of Run and End of Participation questionnaires, and Eye Tracking Data.

A reference run and a solution run were completed for each traffic scenario. A key to match up the Reference vs. Solution runs is shown in Table 10.

Table 10: Reference and Solution Runs Key

	Reference		Solution		Eye Tracking Data
Sector	**L57**	**L5/67**	**L57**	**L5/67**	*N/A*
FRA1	Run 12	Run 9	Run 4	Run 7	Y
FRA2	Run 2	Run 8	Run 10	Run 5	Y
FRA3	*N/A*	Run 11	Run 3	Run 6	N (Run 6)

4.4.1.1 Questionnaire Results

Self-reported data was collected by the participants of the Attention Guidance Simulations in two questionnaires, End of Run and End of Participation. End of Run was completed after every run in the simulation and End of Participation when the participant had completed the study. Results from questionnaire data are illustrated in Sections 4.4.1.1.1 and 4.4.1.1.2.

4.4.1.1.1 End of Run

The questionnaire results from end of run can be seen in Figures 2&3. The data shows that certain areas of workload were affected by the solution differently. For example, teamwork seems to take up less cognitive capacity during the solution runs than during reference runs (questions 10, 12 &13) as well as scanning and interpreting flight information data and prioritising tasks (questions 1, 14 &15). Another positive effect can be seen in conflict management (questions 2, 3 & 4). The data also shows an increase in workload for anticipating future traffic and evaluating the consequences of a plan (questions 5 & 8).

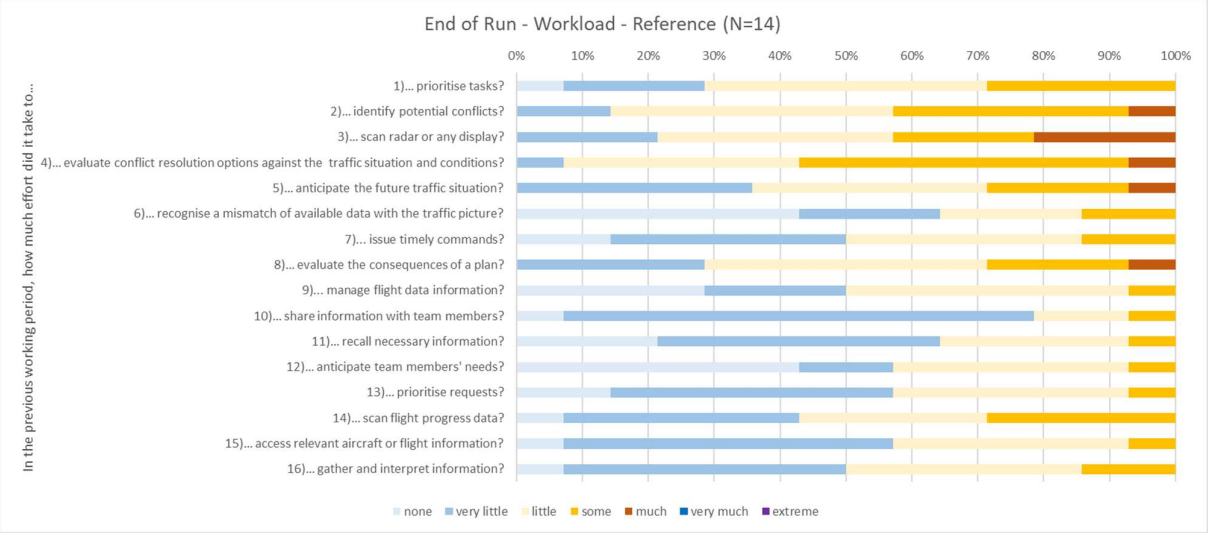


Figure 2: End of Run Workload Reference

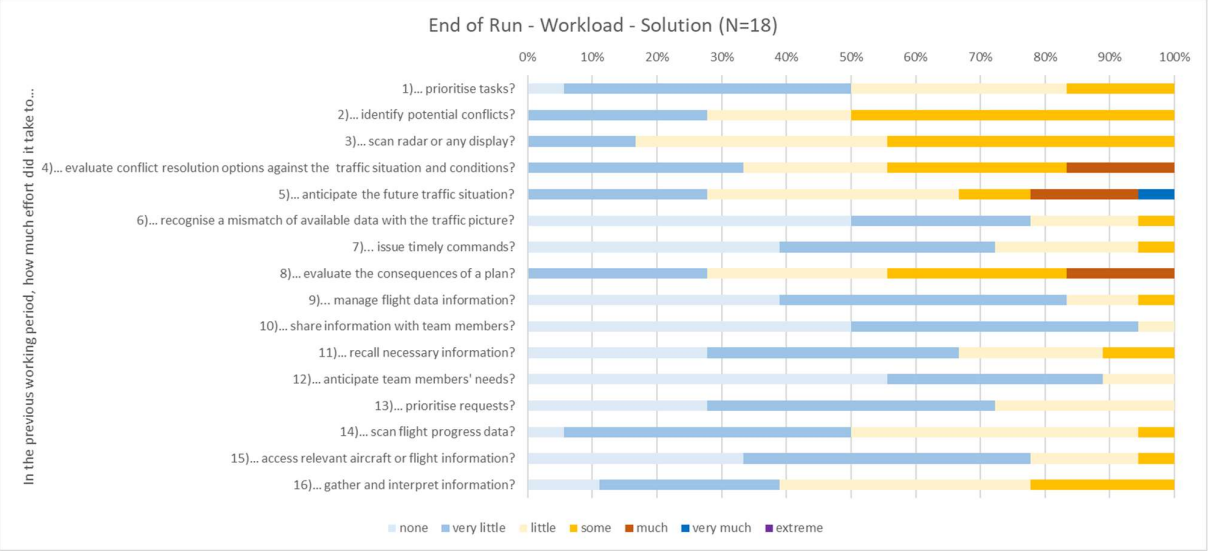


Figure 3: End of Run Workload Solution

Situation awareness (SA) data below shows that participants were able to maintain good SA throughout both conditions. However, a small improvement can be seen as ATCOs found it easier to search for particular information (question 6) and they were less likely to focus on a single problem (question 2). This aligns with the aim of the solution to focus the controllers' attention and decrease the cognitive effort related to visual scanning.

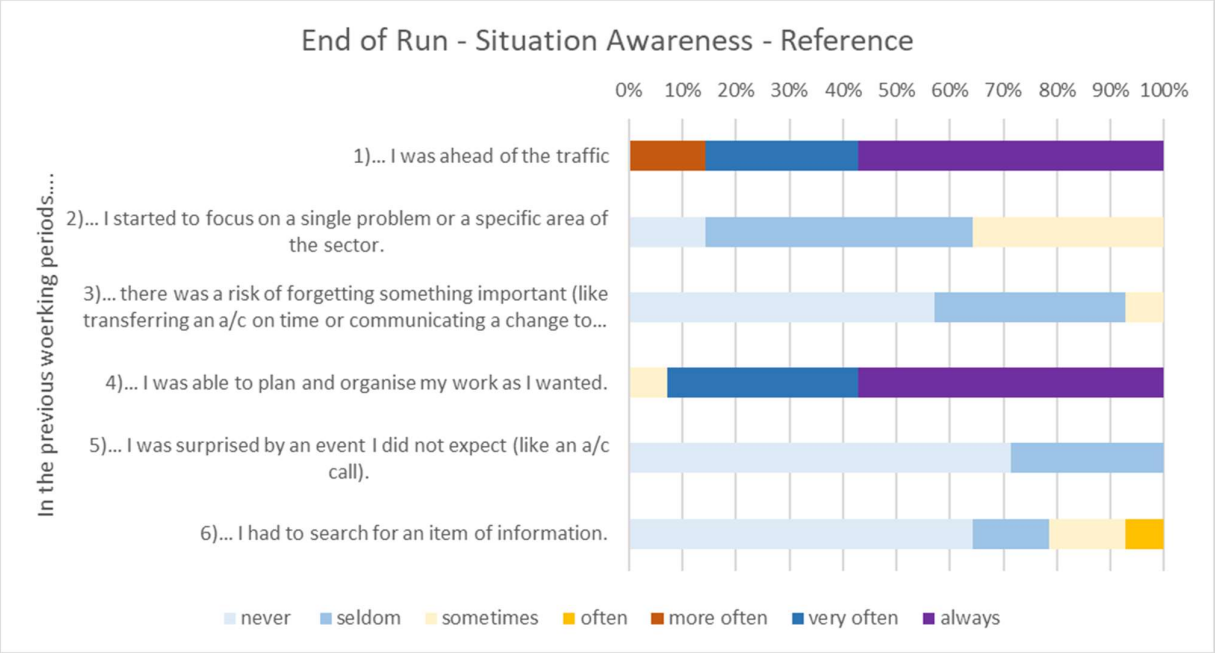


Figure 4: End of Run Situation Awareness Reference

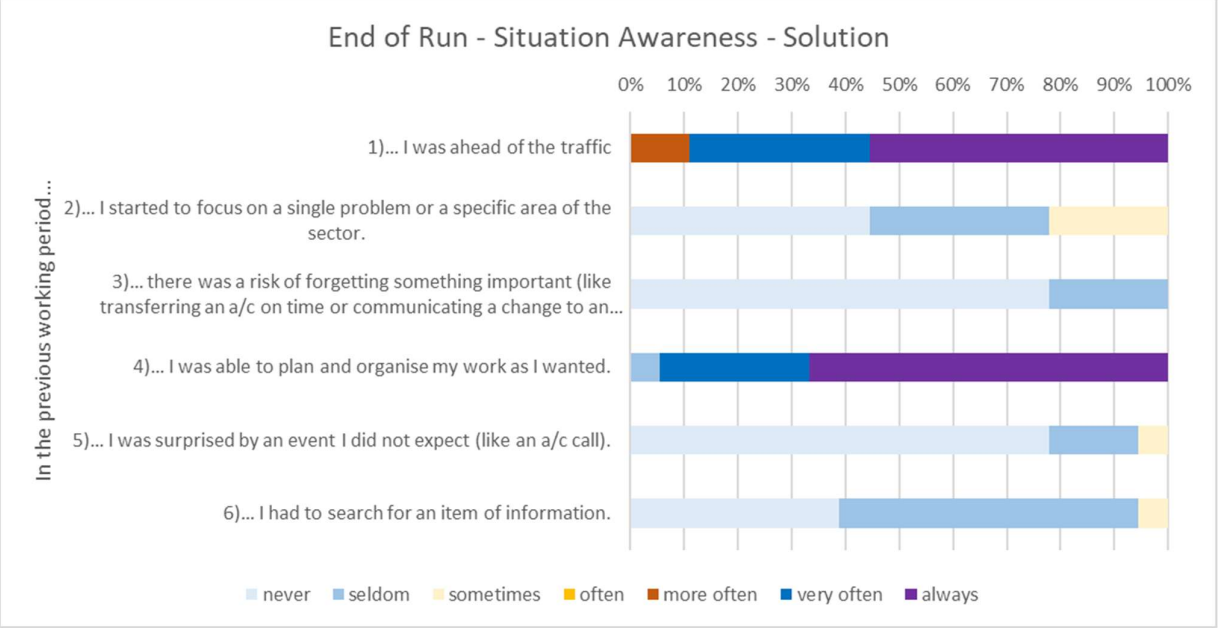


Figure 5: End of Run Situation Awareness Solution

4.4.1.1.2 End of Participation

Section 1: Assessing the Impact of Automation on Mental Workload (AIM-I)

This section was designed to assess the impact of various ATC tasks in the previous working period(s) on the workload of the ATCO, at the end of their participation in the study. The 32 questions asked

in this set were related to information processing, organising tasks and dealing with the flow of information. The ATCO was asked 'In the previous working period(s), how much effort did it take to… For each question, there were seven possible answers, which ranged from 'none' to ''extreme' (none; very little; little; much; very much; extreme). Participants were asked to focus on their experience overall using Attention Guidance rather than individual runs to understand their global attitude towards the solution.

-Note that the term system in some of the questions refers to the technical system they were working with in the previous working period(s), including all the displays, functionalities and tools.

The AIM-I Results for all 32 questions asked regarding the impact of automation on mental workload were mostly answered 'none', 'very little' 'little' or 'some'. Average answers were 'very little' and 'little'. There were 6 questions which contained 'much' (3) or 'very much' (3) responses. All 6 of these questions had the same percentage response answering 'much' or 'very much' (20% of responses for each 6 questions). Overall, the ATCOs responded positively as 'little'/ 'very little' effort was required for most of the questioned asked.

Overall, results confirm the findings from the end of run questionnaire in that the tasks related to conflict resolution and predicting future traffic scenarios take up most cognitive capacity. This is to be expected given that these tasks rely on the third building block of situation awareness (perception, understanding, prediction). During the debriefing it became clear that the solution did not increase the perception of the ATCOs' workload when using Attention Guidance. Therefore, the higher cognitive effort associated with the tasks mentioned is likely down to the complexity of the traffic scenario.

It can be seen that only 20% of participants report these tasks to take "much" or "very much" effort. In further development work it will need to be decided what the acceptable threshold for cognitive effort on specific tasks is to ensure the solution can be assured from a HF perspective.

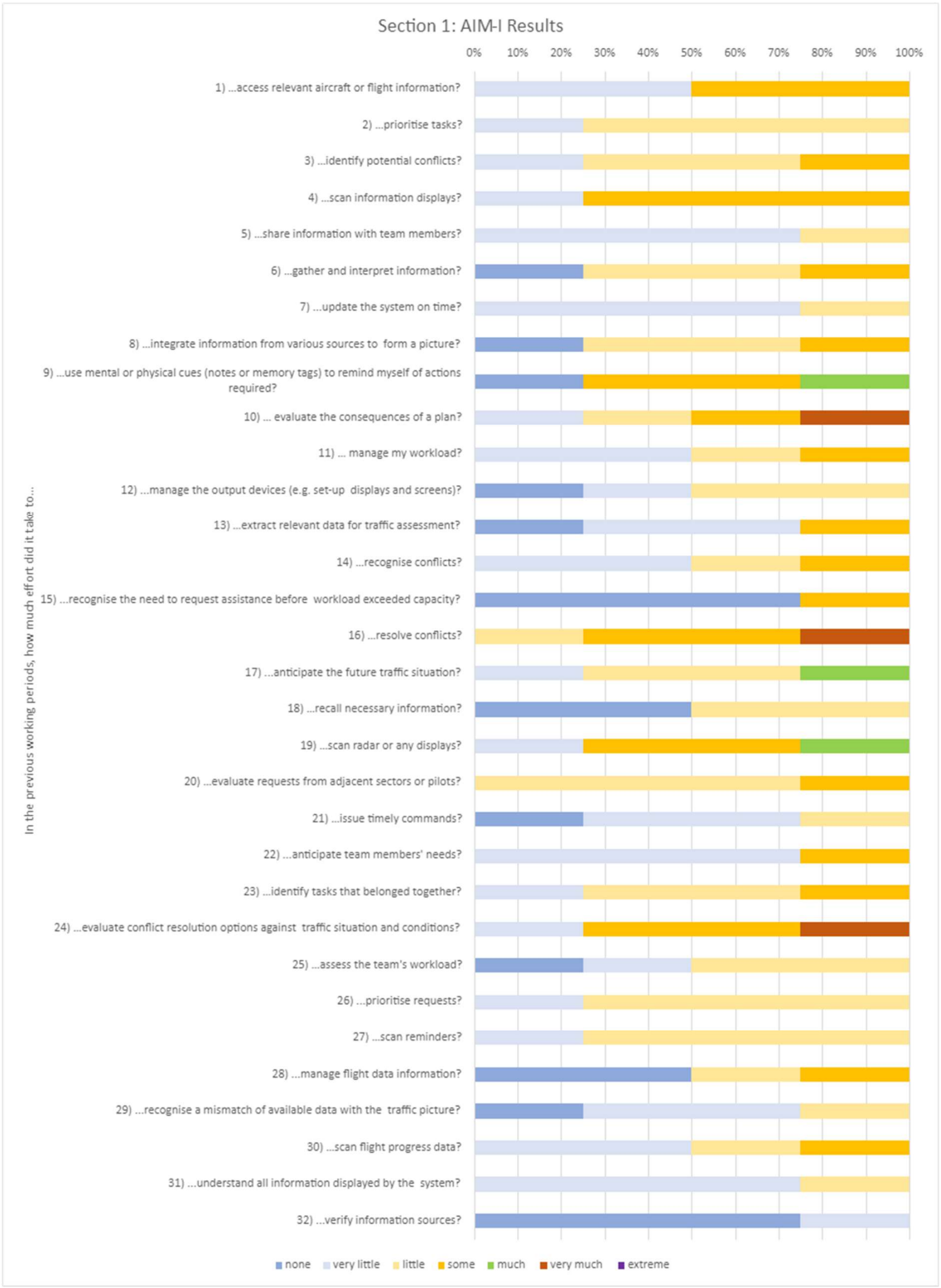


Figure 6: AIM-I End of Participation Results

Section 2: SATI (SHAPE Automation Trust Index)

This section within the end of participation questionnaire was designed to assess the level of the ATCOs trust in the system. Trust is the extent to which a user is willing to act on the basis of external information, recommendations, actions and decisions of another person, a computer based tool or a decision aid. The user was asked 'In the previous working periods, I felt..' then answered 6 statements, with seven possible ratings (none; very little; little; much; very much; always).

The term system in the statements refers to the technical system the ATCO was working with in the previous working period(s), including all the displays, functionalities and tools.

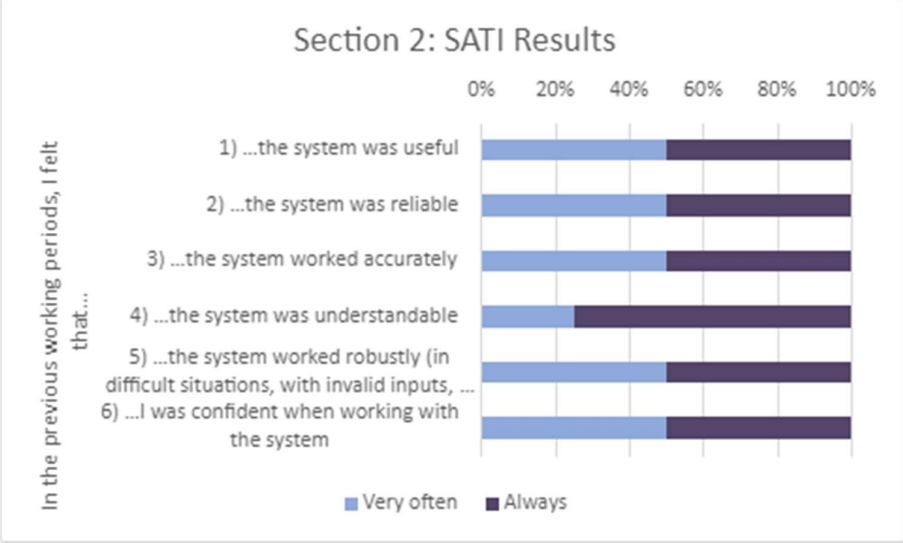


Figure 7: SATI End of Participation Results

The SHAPE Automation Trust results were very positive with all responses 'very often' or always' when asked that the user felt the system was accurate, useful, reliable, etc. No respondents answered neutrally or negatively to these questions. For 5/6 questions, 50% answered 'very often' and 50% 'always'. For 1 question (Question 4) the response result was 25% 'very often' and 75% 'always'.

It can therefore be concluded that the solution is very transparent to the users which will support acceptance into the operation should the functionality be deployed in live ops. For any further development, particular focus should be put on ensuring the system is transparent and can therefore be trusted to support successful deployment and delivers predicted benefits.

Section 3: Safety

For completeness the results from the Safety section of the end of participation questionnaire are reported below. For detailed interpretation of the results refer to section 4.2.3 of the TVALR.

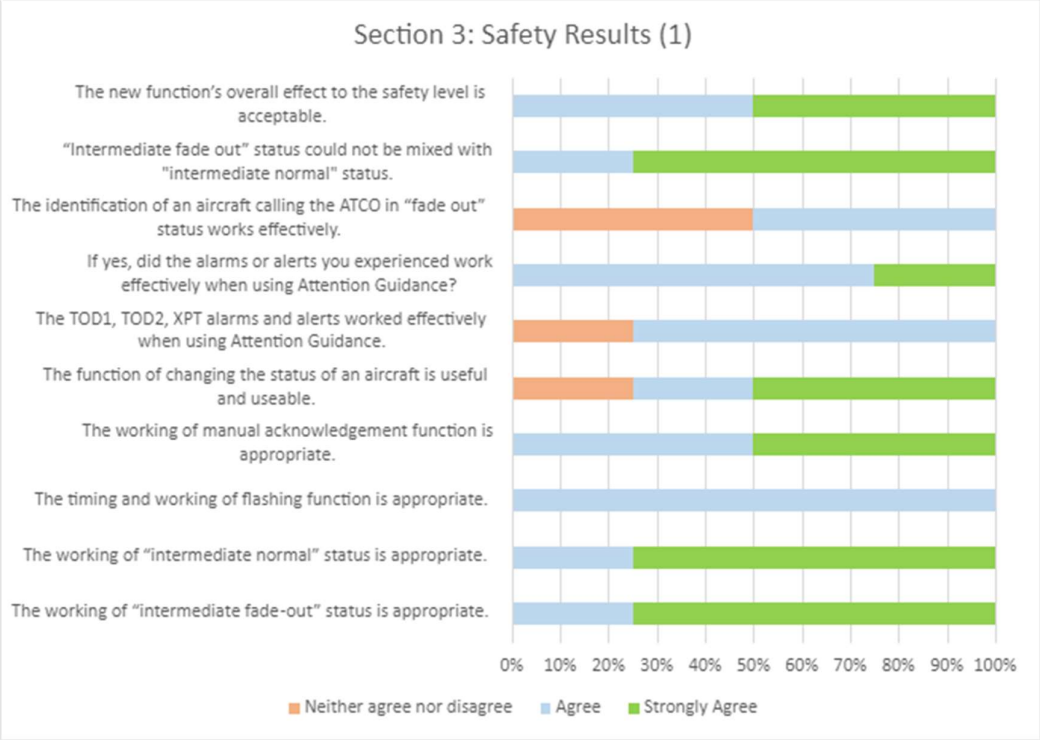


Figure 8: Safety End of Participation Results (1)

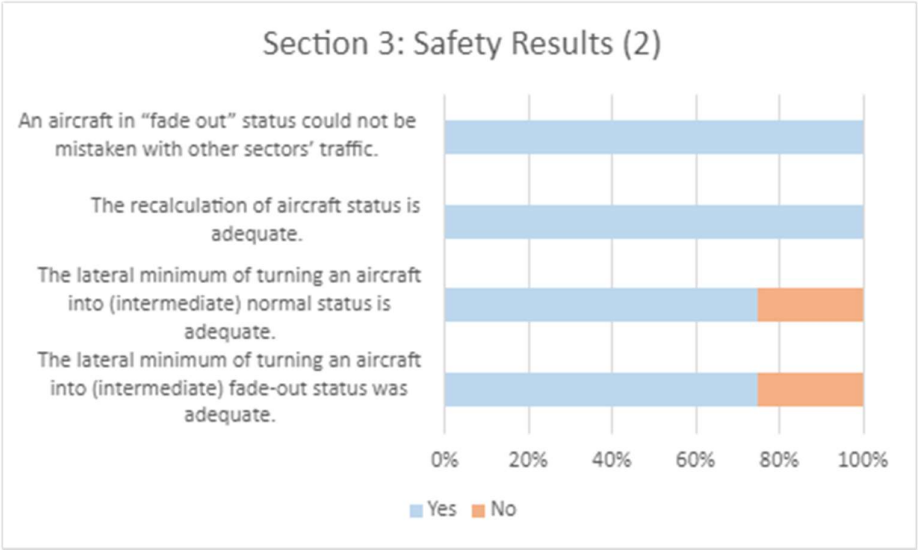


Figure 9: Safety End of Participation Results (2)

Section 3: Safety -Comments

The results of the questionnaires for both End of Run and End of Participation were generally positive. The ATCO's who participated in the study perceived the Attention Guidance tool to reduce radar screen clutter and the tool was popular. The ATCO's made remarks that the tool worked well, and they became comfortable using it. Some ATCO's thought some fine-tuning it would be beneficial to further improve the functionality of the tool. Overall, Attention Guidance was received very well by participants.

4.4.1.2 Debrief- Comments

The debriefings were started by asking controllers to give a general overview of how they found working with the Attention Guidance functionality. Comments were very positive overall with ATCOs reporting a perceived decrease in visual clutter and consequently visual workload. They stated that, when the concept was first introduced to them, they were sceptical as the solution meant that an important task was being taken from the human and given to the AG algorithm. However, after working with the new functionality controllers built up trust into the system very quickly as the system was transparent and easy to interpret. The solution did not cause any surprise to the controllers.

Generally, controllers reported that they forgot about tracks in fade-out status which was intended by the concept. They reported that the integrated task drivers and alerts appear predominantly at the right time to allow them to formulate appropriate plans. These task drivers appear early enough to not cause surprise or create time pressure. There was a concern that it would be difficult to re-integrate fade-out tracks into the mental model once they become relevant again. In the simulated scenarios this was not the case. Tracks that "came back" from fade-out status were easy to re-integrate as the system behaviour was very transparent to the ATCOs. In any further development this aspect should be tested further to ensure this remains as it is after this first simulation exercise.

After completion of run 12 (reference) the ATCOs asked whether they could try this run with the solution as they were convinced that it was a new, very busy scenario. They were surprised when they were told that they had already worked this scenario with the solution in run 4. The decrease in visual clutter meant their perception of workload was reduced significantly to the point where they thought it was a different traffic sample. They reported that they would not work with such high traffic volumes in real operation and that they would normally split out the sector. In run 4 this was not the case. The controllers worked the same scenario without mentioning that they would normally split the sector in this high traffic load. This is a promising result for the concept in terms of controllers being able to handle higher volumes of traffic. However, it also means that any system failure or degraded modes of operation need to be managed carefully if controllers start to take on more traffic using attention guidance than they would without.

Some further comments included:

- Timing of frequency pop up could be changed to closer to the sector boundary as it caused some additional clutter
- ATCOs reported it "felt good" to turn tracks into fade-out status indicating that they truly excluded them from their mental mode. This behaviour is in line with the concept but needs to be monitored carefully to ensure that they can re-integrate fade-out tracks when needed.

- The intermediate status was positively received as it gives a timely heads-up that actions will be required within a manageable time frame
- The thresholds for track status changes could be changed when trust in system is established and settled. ATCOs appreciated that the 20nm buffer provides them with enough time to react to status changes but could be reduced in future when the implications of doing so are fully understood.
- ATCOs were generally surprised as to how quickly they got used to the new human-machine task distribution and how quickly they were able to understand and trust the system.

4.4.1.3 Eye Tracking Results

During 8 of the simulation runs eye tracking data was collected. ATCOs wore eye tracking glasses, the 'Tobii Pro Glasses 2' (Figure 1). Video capture via the scene camera was collected, as well as eye movement via the 2 cameras per eye. The glasses were connected via a cord to the recording device, which was either attached to the controller, or resting on the desk. The data was saved onto a SD card within the recording device. The glasses were chosen for use due to their practicality in the set up, and with minimal obstruction to controllers. The software 'Tobii Glasses Controller' was used to calibrate and start/stop the recordings.

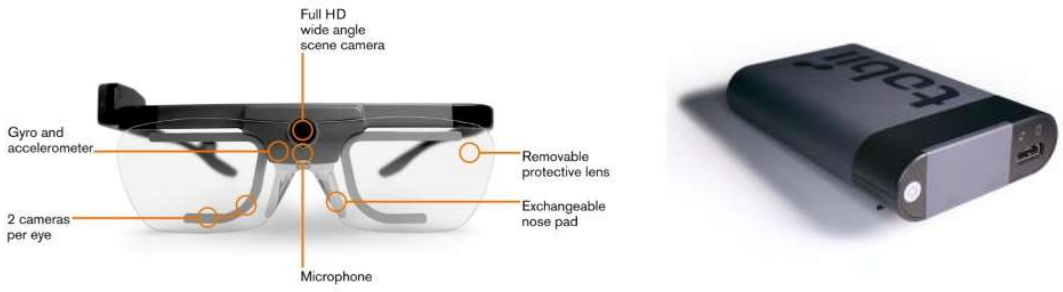


Figure 10 Tobii Eye Tracking Equipment

Data was collected during:
FRA 1

- Run 12 (reference) – Run 4 (solution)
- Run 9 (reference) – Run 7 (solution)

FRA2

- Run 2 (reference) – Run 10 (solution)
- Run 8 (reference – Run 5 (solution)

No Data was collected during:
FRA3, Sector L57

- Run 3 (Solution)

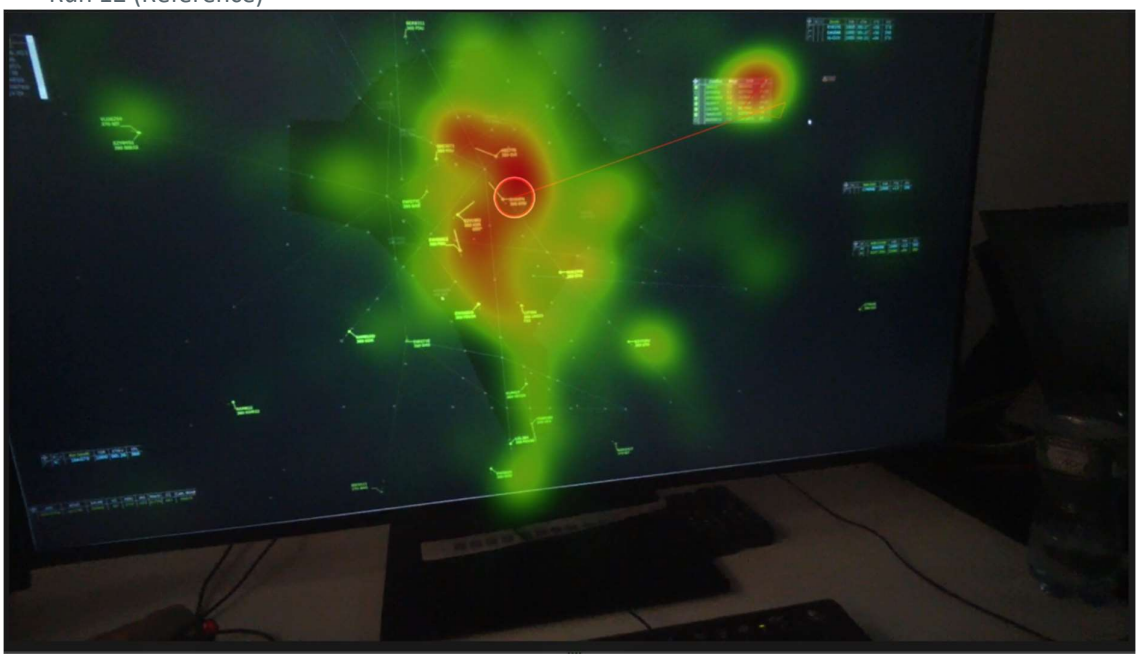
FRA3, Sector L5/57

- Run 11 (Reference)
- Run 6 (Solution)

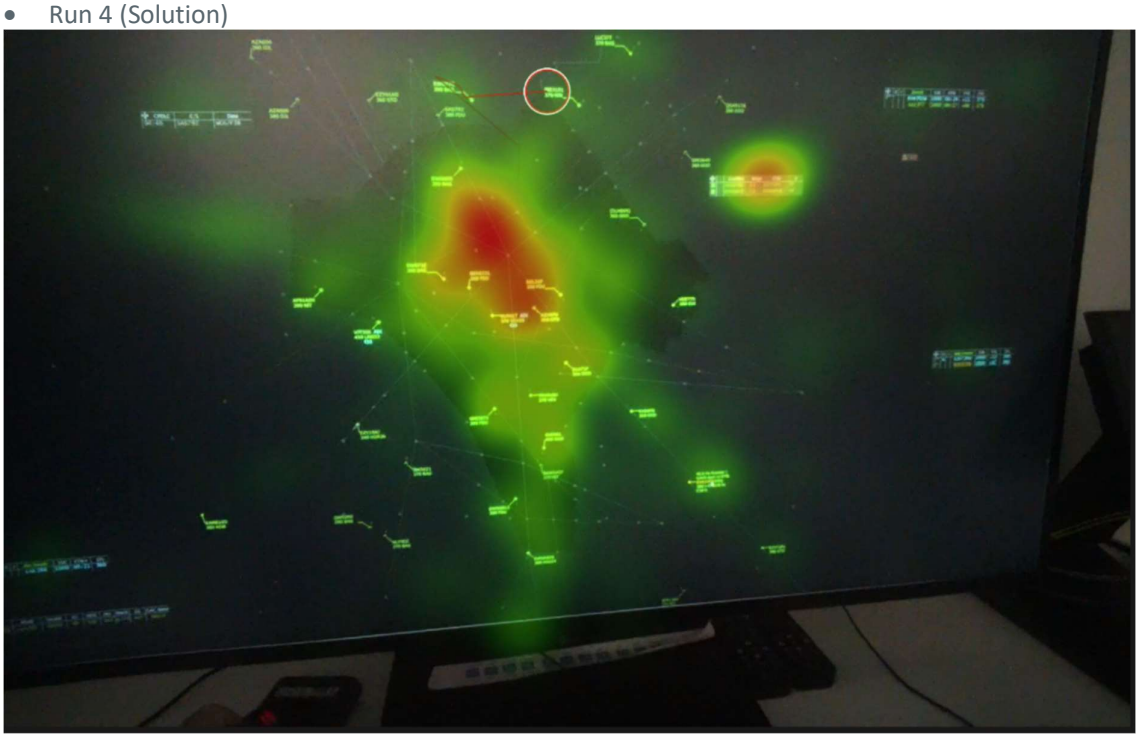
In the figures below the results of the eye tracking data collection are displayed as “heatmaps”. These heatmaps are a simple way to visualise at which parts of the screen the controllers focused more frequently (red areas) compared to less frequent scanning (green areas). The pictures can therefore be interpreted similarly to a weather radar display. A brief interpretation of the results can be found with every reference-solution run pairing.

FRA1, Sector L57

- Run 12 (Reference)



- Run 4 (Solution)



In this FRA1 bandboxed scenario it can be seen that the focus area is more condensed in the solution run than it is in the reference run. This is according to what was expected from a concept perspective as the controller is able to focus their attention to the relevant areas of the screen.

In the table below the metrics for Run 12 and 4 is shown in more detail. A distinction was made between scanning behaviour inside the sector boundaries versus outside the sector boundaries to understand whether the conflict resolution (inside the sector) or planning horizon (outside the sector) were affected differently by the solution or reference scenario. Note: a fixation is defined as any period of time greater than 50ms of the eye resting on a piece of information, fixations are connected to a deeper level of cognitive processing and therefore linked to intentional attention and perception (e.g. forming a plan for conflict resolution); a saccade is defined as an eye movement from one piece of information to the next lasting between 20 and 40m; a visit is defined as the time spent in an area of interest combining fixations and saccades, visits are linked to a different way of cognitive processing compared to fixations only, they tend to be connected to pattern matching and quickly updating an existing mental model rather than for example conflict resolution which takes more dedicated effort.

Table 5 Visual Scanning Metrics (R4&12)

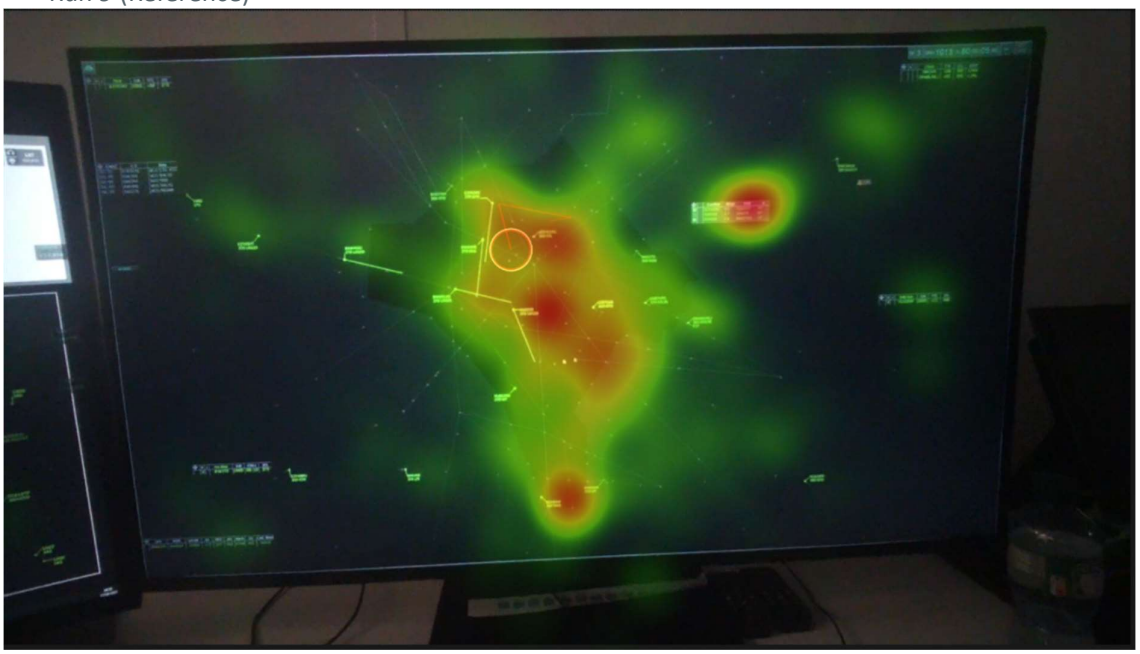
Scenario	Area of interest	% age of fixation time	Average fixation duration (sec)	% age of fixation count	Average visit duration (sec)	% of visit count
Reference (R12)	Inside sector boundary	59	0.26	60	2.32	47
	Outside sector boundary	41	0.27	40	1.44	53
Solution (R4)	Inside sector boundaries	58	0.30	55	1.92	45
	Outside sector boundaries	42	0.28	45	1.25	55

Definitions: %age of fixation time = relative amount of fixations in each area of interest excluding other areas (e.g. communication panels, support screens, etc.); average fixation duration = average length of individual fixations; %age of fixation count = similar to fixation time but based on absolute numbers of fixations in each are of interest; average visit duration = similar to average fixation duration but including saccades; %age of visit count = similar to fixation count but including absolute number of saccades.

The parameters in the table suggest that the ATCO shifted their attention from inside the sector to a more balanced scan between inside and outside the sector (%age of fixation count shifts by 5% points) and therefore focussing more on planning traffic into the sector during the solution run compared to the reference scenario. The average fixation duration increases (from 0.26 to 0.30 seconds inside the sector) simultaneous to a decrease in visit duration (from 2.32 to 1.92 seconds inside the sector) indicating the ATCO focused longer on relevant tracks while identifying them quicker per visit in the area of interest.

FRA1, Sector L5/57

- Run 9 (Reference)



- Run 7 (Solution)

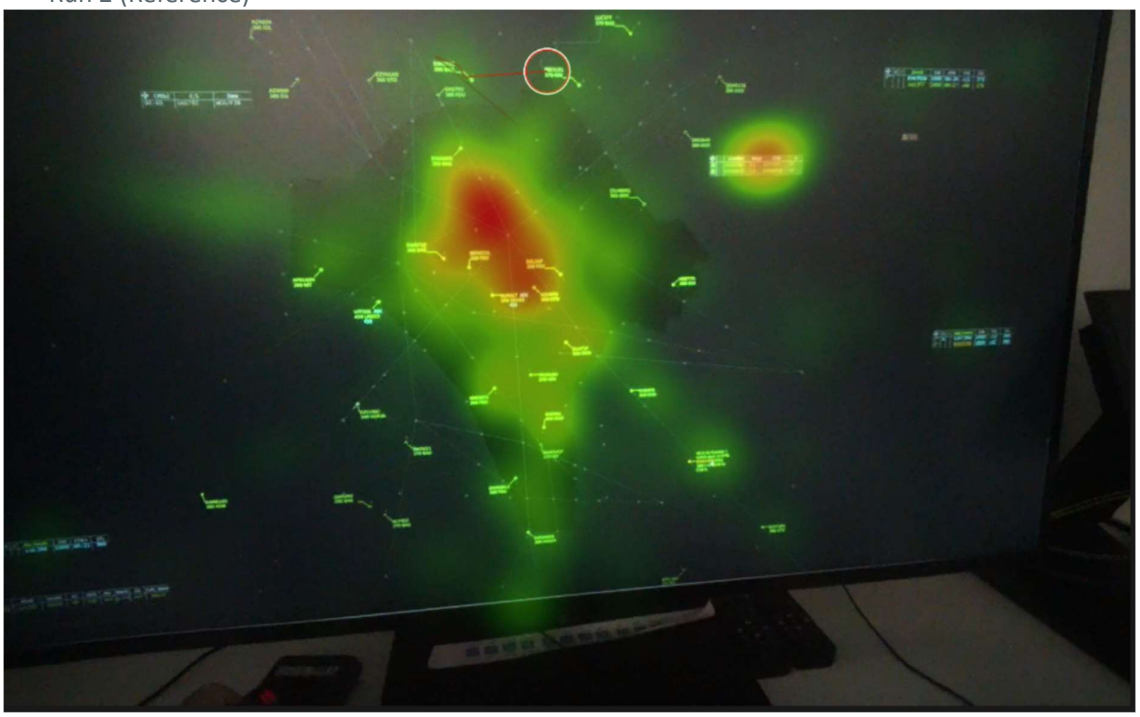
No Data

During run 7 the controller sat at an unexpected angle to the workstation as they were discussing the run and solution with the controller to their right. This meant that the analysis software was unable to run the necessary data mapping and therefore this run could not be analysed.

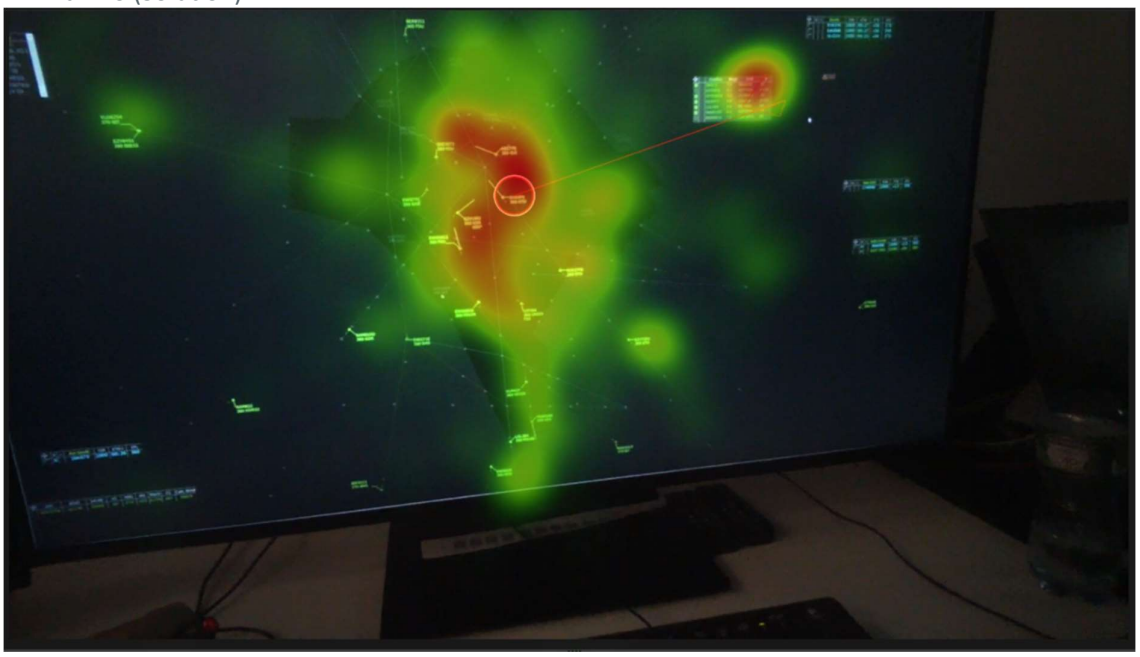
From the reference scenario alone (run 9) no conclusion as to how the solution affects the scanning pattern can be drawn. The analysis could be repeated during further development work.

FRA2, Sector L57

- Run 2 (Reference)



- Run 10 (Solution)



The heatmaps for this run pairing do not follow the expected pattern of showing a more focused scan in a particular are of the sector. Due to the small sample size it is difficult to say whether this is a genuine outlier or whether there is not enough evidence to confirm the trend observed in the other run pairings. Looking at the table below, showing the metrics, it seems as if this particular run could

be an outlier as the numbers deviate from the trend observed in other run pairings as well. However, any conclusion would need to be tested on a larger sample size.

Table 6 Visual Scanning Metrics (R2&10)

Scenario	Area of interest	% age of fixation time	Average fixation duration (sec)	% age of fixation count	Average visit duration (sec)	% of visit count
Reference (R2)	Inside sector boundary	81	0.27	79	2.86	56
	Outside sector boundary	19	0.25	21	0.88	44
Solution (R10)	Inside sector boundaries	75	0.33	73	2.78	49
	Outside sector boundaries	25	0.29	27	0.91	51

Definitions: %age of fixation time = relative amount of fixations in each area of interest excluding other areas (e.g. communication panels, support screens, etc.); average fixation duration = average length of individual fixations; %age of fixation count = similar to fixation time but based on absolute numbers of fixations in each are of interest; average visit duration = similar to average fixation duration but including saccades; %age of visit count = similar to fixation count but including absolute number of saccades.

The parameters in the table suggest that the scanning pattern of this participant is more extreme towards fixating within the sector boundaries rather than fixating at traffic coming into the sector. However, the controller does visit both areas of interest for roughly the same number of times. This could indicate that this individual takes in most information outside the sector very quickly and puts more focus on conflict resolution within the sector. There is a small tendency for a more even split during the solution run but not significantly so.

FRA2, Sector L5/57

- Run 8 (Reference)

- Run 5 (Solution)

The results for this reference and solution run are in line with expectations. During the solution run the controller focussed their attention in a more concise area of the screen. This indicates that the solution is indeed guiding the controllers' attention to the most relevant part of the sector. In further work it could be explored whether the areas of interest correlate to specific sector characteristics (e.g. common conflict points, TODs, planning horizon, etc.).

Table 7 Visual Scanning Metrics (R5&8)

Scenario	Area of interest	%age of fixation time	Average fixation duration (sec)	%age of fixation count	Average visit duration (sec)	% of visit count
Reference (R8)	Inside sector boundary	67	0.22	69	2.76	51
	Outside sector boundary	33	0.24	31	1.28	49
Solution (R5)	Inside sector boundaries	69	0.38	67	3.24	51
	Outside sector boundaries	31	0.36	33	1.52	49

Definitions: %age of fixation time = relative amount of fixations in each area of interest excluding other areas (e.g. communication panels, support screens, etc.); average fixation duration = average length of individual fixations; %age of fixation count = similar to fixation time but based on absolute numbers of fixations in each are of interest; average visit duration = similar to average fixation duration but including saccades; %age of visit count = similar to fixation count but including absolute number of saccades.

The parameters in the table above suggest that during the solution run the participant started to focus on individual tracks for a longer period of time suggesting that conscious cognitive effort was placed on those tracks. This is in line with the assumption that with Attention Guidance controllers only focus on "relevant" tracks. Additionally, the average visit duration increased pointing to the same conclusion. The controllers spent more time interrogating relevant tracks.

In summary all run pairings are in line with the hypothesis of Attention Guidance changing controllers' visual scanning behaviour towards longer fixation times in both areas of interest but in particular inside the sector. this indicates that the algorithm does lead them to scan relevant tracks for a longer period of time to achieve goals that require greater cognitive effort (e.g. conflict resolution).
The heatmaps show that in two out of three run pairings the scan changed to a narrower focus area in line with the concept aim. It should be understood in further work whether these areas correlate to specific sector characteristics to ensure that Attention Guidance leads controllers to the most important parts of their sector.

Not part of the analysis was reaction times to unexpected events such as emergency squawks. The equipment used does not provide the necessary resolution to identify these events in the analysis. A more specific experimental set up would be needed to investigate this further. It would provide insights to what extent the algorithm changes the situation awareness and provides assurance that the human is still in the loop and can integrate fade-out tracks into their mental model quickly enough compared to today's operation.

4.4.2 Summary of HP activities recommendations / requirements

Issue ID	HP issue / Benefit	HP Issue/ Benefit Status	HP/ Valid. Obj. ID	activity conducted	results / evidence	recommendations	requirements
Arg. 1.2: Operating methods are exhaustive and support human performance.							
HFI-ARG1.2-10.96-TRL6-001	Procedures to reflect response to track status - AG will present tracks in different states to the controller. As a result, each status will require specific interaction which controllers are not familiar with in today's operation. There is a risk for controller confusion as to how to interact	Closed	OBJ-ARG1.2-10.96-TRL6-001	Real Time Simulation	Recognising conflicts scored well with little effort required by ATCOs to do so (responses varied from 'very little'- 'some'). Resolving conflicts required more effort, with most responses 'little' or 'some', but a small proportion (25%) required 'very much' effort. The comments provided by the controllers in the questionnaires do not fully explain the reasons for the high cognitive load. Some comments state that during the run not a lot of tracks were eligible for fade-out status meaning the controllers applied their current ways of working. It is expected for controllers to express that conflict resolution	N/A	N/A

	with tracks in certain states (especially 'intermediate' status).				does take up a significant proportion of the cognitive capacity as it involves integration and understanding of all available information and formulating of a suitable plan to resolve the situation without causing further conflictions. The comparison between reference and solution, looking at questionnaire data, suggests that controllers require less effort for tasks such as taking in information and not focusing on single problems. Acceptance rates for 'Intermediate normal' and 'Intermediate fade-out' status were good. 25% of ATCOs rated these as 'agree' 75% 'Strongly agree[d]' that the working of these statuses are appropriate. Therefore, no further recommendations are needed as the Attention Guidance tool has scored positively and helped ATCOs identify different track statuses.		

Arg. 1.2.2: Operating methods cover operations in abnormal operating conditions.

HFI-ARG1.2-10.96-TRL6-002	Procedures to reflect abnormal weather situations - The use of Attention Guidance might not be appropriate in all weather and traffic scenarios. When aircraft start to avoid weather, all tracks in the sector suddenly become relevant. When traffic levels are very low all tracks could be in fade-out status. If the switch between using AG and not using it is not clearly defined there is a risk for controllers using	Open	OBJ-ARG1.2-10.96-TRL6-002	Real Time Simulation	In the end of participation questionnaire on conflict resolution; higher effort was required in the question asked to evaluate conflict resolution options against the traffic situation and conditions (25% very little, 50% some, 25% very much effort) in comparison to most of the other questions asked to assess the impact of automation on mental workload. At this maturity level challenging weather situations were not part of the validation scenarios. This should be included in future development work.	It is recommended that bad weather situations as well as very low (but still realistic) traffic scenarios are included in further real time simulation work to understand the threshold of the usefulness of Attention Guidance.	Further development work shall develop procedures to decide when and how Attention Guidance is switched off/ not appropriate to use.

	the functionality inappropriately.						
Arg. 1.2: Operating methods are exhaustive and support human performance.							
HFI-ARG1.2-10.96-TRL6-003	Procedures around probing of clearances - AG might require a change in the way that controllers use the probe functionalities on their workstations. Currently probing clearances is up to the ATCO and not mandatory. Given that a/c could move from 'fade-out' to becoming relevant quickly ATCOs might not have them integrated in their mental	Open	OBJ-ARG1.2-10.96-TRL6-003	Real Time Simulation	Mean acceptance scores for functionality and operating methods of the system scores were high, with a 100% positive response to 7/10 Safety functionality questions of 'agree' or 'strongly agree' responses (to appropriateness of functionality and alarms). The other 3/10 questions contained some neutral scores (25-50% 'neither agree nor disagree'). As covered in HFI-ARG1.2-10.96-TRL6-001 above, 20% of controllers reported that it took "much" or "very much" effort to resolve conflicts. Whilst it is not clear whether these scores were related to the use of AG or to today's way of operating the question around use of probing functionality is not fully answered at this stage. Nevertheless, during debriefings controllers were confident that the	It is recommended to continually monitor the appropriate use of existing functionality such as probing when introducing new functionality such as AG.	The use of the probing functionality should be considered during further development work of AG.

	picture as quickly as the situation requires. There is a risk of issuing inappropriate clearances due to a lag of updating their mental model.				use of the probe functionality was appropriate and would not need to change as part of the implementation of this solution. However, when changes to procedures are introduced in order to integrate the functionality in the existing operating producers the use the probe functionality should be reviewed to ensure they are effective and appropriate.		
Arg. 1.2.2: Operating methods cover operations in abnormal operating conditions.							
HFI-ARG1.2-10.96-TRL6-004	Procedures around alerting philosophy - AG might introduce new alerts. Clear procedures around the appropriate responses to alerts need to be introduced. Otherwise there is a risk of controllers not reacting in the desired manner	Open	OBJ-ARG1.2-10.96-TRL6-004	Real Time Simulation	Acceptance of effectiveness of alarms and alerts was good (12.5% neutral, 87.5% agree/ strongly agree). Situation awareness improved or did not change on average with use of the Attention Guidance tool when comparing the reference and solution run data, which indicates that alerts appropriately supported the controllers in maintaining their situational awareness. During debriefings controllers stated that the alerts were appropriately driving their	It is recommended that a clear design philosophy is applied when further developing and refining alerts and task drivers in the AG functionality.	Attention Guidance alerts and task drivers shall be in line with alerting philosophy guidelines.

	which could lead to safety critical situations.				attention towards relevant flights and acted as task drivers to trigger correct actions at the correct time. It should be considered during further development to distinguish between alerts and task drivers and follow suitable design principles.		
HFI-ARG1.2-10.96-TRL6-005	Procedures around dealing with emergencies - AG could add a layer of complexity when ATCOs are required to handle an aircraft emergency. If the system fades out a track and this aircraft has an emergency the controller needs to be able to integrate the track into their mental model very quickly. If a	Open	OBJ-ARG1.2-10.96-TRL6-005	Real Time Simulation	A number of different emergency scenarios were integrated in the simulation. Controllers were able to spot them and react appropriately. They reported that even when a flight is in fade out the buffer (20nm) provides them with enough time to react to any emergency scenario. Whilst this is a promising result, emergency scenarios should still be integrated and considered during further development work to ensure sufficient safety buffers are kept for each possible track status to give controllers enough time to react and formulate appropriate plans.	It is recommended that emergency scenarios remain a key use case for further development work.	Emergency scenarios shall be considered as part of further Attention Guidance development work.

	track is displayed as 'normal' and declares an emergency the ATCO has to take all surrounding tracks ('normal', 'fade-out' 'intermediate') into consideration when formulating a plan for the aircraft.						
Arg. 2.1.2: Changes to the task allocation between human and machine support human performance..							
HFI-ARG2.1-10.96-TRL6-001	AG will introduce changes to the task distribution between human and machine. Their appropriateness will need to be assessed during validation exercises. The main change will	Open	OBJ-ARG2.1-10.96-TRL6-001	Real Time Simulation	The trust in automation questionnaire indicated that the overall trust in the system and its transparency are very good. This means the system's "decision making" can be easily interpreted by the controller and does not lead to surprises/startle responses. Almost 80% of participants reported that the system was "always" understandable, just over 20% of participants reported that	It is recommended that trust in automation scales remain a key parameter during further development work to ensure Attention Guidance remains	N/A

	be from human assessment whether a track needs interaction to the machine performing this assessment. The human actors will need to understand the logic of the algorithm to decide whether the assessment is appropriate and can be trusted.				the system was "very often" understandable. 50% of controllers "always" felt confident the use of the system and the other 50% "very often" felt confident in the use of the system. there were no negative responses in the trust scale. During debriefings controllers reported that when first introduced to the concept they were sceptical and expected to not be comfortable handing over more control to the system. They were surprised that only after a small number of runs they already started to feel comfortable and started to trust the algorithm's decisions. This is a very promising result for the concept and indicates that controllers can see real benefit in this solution.	transparent to the user.	
HFI-ARG2.1-10.96-TRL6-002	AG could affect the interaction with existing functionalities on the CWP. Wider use of probing	Closed	OBJ-ARG2.1-10.96-TRL6-002	Real Time Simulation	In terms of identifying potential conflicts, there was a moderate improvement between the reference run and the solution run using the AG tool. Identifying conflicts in the reference run	N/A	N/A

	functionality could be required. This would transfer the task of pre-clearance conflict detection from the human to the machine.				scored 15% 'very little'; ~43% 'little'; 35% 'some' and 7% 'much' effort; In the solution run scores were moderately better (equating to lower effort) at 27.5% 'very little'; 22.5% 'little' and 50% 'some'. It is not clear whether this improvement is due to the visual de-clutter AG brings or whether the users changed/adapted the way they used the probe functionality. This questions was covered during debriefings but users did not report that they would change the way they probe conflicts due to AG.		
Arg. 2: Technical systems support the human actors in performing their tasks.							
HFI-ARG2.1-10.96-TRL6-003	The introduction of AG might require a more standardised way of displaying information on the CWP. Currently ATCOs can chose if and	Closed	OBJ-ARG2.1-10.96-TRL6-003	Real Time Simulation	ATCOs were asked during the debrief whether they had used the speed lines any differently when working with AG. They responded that their way of using speed lines did not change due to the introduction of AG.	N/A	N/A

	how to display speed lines depending on personal preference. With AG speed lines will become an (even more) important indication for situation awareness. They will be highlighted/fade-out when a track goes into 'intermediate' status. ATCOs need to able to pick up on this cue regardless of their standard way to display speedline						

Arg. 2.1.5: Human actors can acquire an adequate mental model of the machine and its automated functions.

HFI-ARG2.1-10.96-TRL6-004	There is a potential risk that controllers overestimate their cognitive capacity using AG. Tracks in fade-out status are still the controller's responsibility but could be neglected from their mental model as they are less salient and don't require immediate interaction. In case of e.g. abnormal weather developing, AG could get switched off which then means a sudden increase in salient tracks which could	Open	OBJ-ARG2.1-10.96-TRL6-004	Real Time Simulation	Mean aggregated situation awareness scores from the end of run questionnaires indicate a slight improvement in SA with use of the Attention Guidance tool in comparison to the reference data for 3/6 Situation Awareness Questions (2, 4 and 5). There was no change for one question (1) and no significant mean change for two questions (6 and 7). The themes emerging from the questionnaire that were positively affected by AG are related to not tunnelling on individual problems and being able to manage and organise tasks "as they wanted" which increased a feeling of autonomy and control. Therefore, all questions recorded either an improvement in SA; or no significant change in comparison to the baseline (reference) data. This was even true when individual aircraft declared emergencies. This is a promising result but more extreme situation should be considered in further simulations	It is recommended that more extreme scenarios (e.g. mass deviations) are simulated to ensure the controllers are able to manage sudden loss of AG functionality and having to re-integrate greater numbers of fade-out tracks and detect conflicts against fade-out tracks.	Further development simulations shall incorporate scenarios where ATCOs quickly have to re0-integrate fade-out tracks into their mental model (e.g. weather or mass diversion scenarios).

	overwhelm the controller. Aircraft emergencies could lead to the same outcome (e.g. mass diversion).				to fully explore the capability of the human to remain in the loop and take back full control when required.		
Arg. 2.1.6: The level of trust in automated functions is appropriate							
HFI-ARG2.1-10.96-TRL6-005	It will be important for controllers to understand the logic of the algorithm which determines whether a track is displayed as 'normal' or 'fade-out'. If the underlying logic is not clear to the user, they are less likely to trust it.	Closed	OBJ-ARG2.1-10.96-TRL6-005	Real Time Simulation	Aggregated acceptance scores for the SATI (SHAPE Automation Trust Index) Section; designed to assess the level of trust in the system by the user was included within the end of participation questionnaire indicate the system scored highly in all 6 categories (useful, reliable, accurate, understandable, robustly, confidence when working with the system). There was a 50/50 split between 'very often' and 'always' ratings for 5 of the 6 questions. The result increased slightly for the 'the system was understandable' to 75% (3) ATCOs giving it the highest available score ('always'). Overall, acceptance and trust in the system was very good.	N/A	N/A

					No further recommendations are therefore deemed to be necessary as it is already captured under HFI-ARG2.1-10.96-TRL6-001.		
Arg. 2.2.1: The accuracy of information provided by the system is adequate for carrying out the task.							
HFI-ARG2.2-10.96-TRL6-001	To support user trust, it will need to be checked whether the system performs accurately against pre-defined parameters such as separation minima.	Closed	OBJ-ARG2.2-10.96-TRL6-001	Real Time Simulation	Results from the end of participation questionnaire showed ATCOs answered positively ('very often' or 'always') on all questions regarding reliability or accuracy, e.g. 'the system was reliable' and 'The system worked accurately'. There were no negative responses for any of the questions related to the accuracy of information provided, indicating that the accuracy of the information provided by the Attention Guidance tool was adequate. No further recommendations are therefore necessary.	N/A	N/A
Arg. 2.2.2: The timeliness of information provided by the system is adequate for carrying out the task.							
HFI-ARG2.2-10.96-TRL6-002	To support user trust, it will need to be checked whether the	Open	OBJ-ARG2.2-10.96-TRL6-002	Real Time Simulation	The system scored positively in the end of participation questionnaire on workload-related questions involving timeliness of information.	It is recommended that the timeliness of certain information	Timeliness of information displayed shall remain a key

	system performs according to defined lag maxima.				Approximately 100% (4 ATCOs) self-reported a minimal increase in workload levels ('very little'-'some') from relevant questions; e.g. 'how much effort did it take to access relevant aircraft or flight information?' indicating the timeliness of information displayed by the system was not significantly increasing ATCO levels of workload. However, within the comments section of the same questionnaire, 50% (Two) ATCO's made comments that 'frequencies pop up too early' and 'the frequency on fade out could appear later', indicating they would prefer information to be displayed slightly later than the current system displays frequencies.	provided is reviewed (e.g. frequency pop ups)	consideration during further development work.
Arg. 3.2.1: Changes to the task allocation between human actors do not lead to adverse effects on human tasks.							
HFI-ARG3.2-10.96-TRL6-001	It is not clear yet whether the use of AG needs to be managed on a higher level than	Open	OBJ-ARG3.2-10.96-TRL6-001	Real Time Simulation	The supervisor position was not included in the simulation exercises. Further development work should include the supervisory role as part of the	It is recommended that the role of the supervisor is considered for procedure	N/A

	individual workstations. There could be a necessity for the supervisors to be responsible for deciding when AG can be used and when it needs to be switched off. Tasks to enable this must be split appropriately between controllers and supervisors to avoid confusion. Consequently, controllers might want to keep the possibility to switch AG off at their workstation even though the supervisor has enabled it.				development of adequate procedures to enable/disable AG.	development around enabling/disabling the AG functionality.	

Arg. 3.3.1: Intra-team and inter-team communication supports the information requirements of team members.

HFI-ARG3.3-10.96-TRL6-001	Following on from HFI-ARG3.2-10.96-TRL6-001 it needs to be ensured that efficient communication is possible between controllers and supervisors when it comes to decision making about using AG or switching it off.	Open	Validation activities should assess whether an acceptable level of team communication can be established between controllers and supervisors.	OBJ-ARG3.3-10.96-TRL6-001	See HFI-ARG3.2-10.96-TRL6-001	It is recommended that the role of the supervisor is considered for procedure development around enabling/disabling the AG functionality.	N/A

Table 8: Summary of the HP results and recommendations/ requirements for each identified issue & related argument

4.4.3 Maturity of the Solution

Maturity checklist for finalising the V3 assessment			
ID	**Question**	**Answer** *Fill in 'yes' or 'no'.*	**Comments** *Please substantiate your answer.*
1	Has a Human Performance Assessment Report been completed? Have all relevant arguments been addressed and appropriately supported?	Y	*This document is the Human Performance Assessment report. All arguments have been addressed in section 4.4.2*
2	Are the benefits and issues in terms of human performance and operability related to the proposed solution sufficiently assessed (i.e. on the level required for V3)?	Y	*A good and holistic picture could be obtained of all benefits and issues during the validation exercise. All relevant HP evidence could be collected during the exercise.*
3	Have all the parts of the solution/concept been considered?	Y	*A detailed description of the solution was available (section 4.1.2). All HP arguments were considered in light of the changes that this solution would bring to the operation (section 4.1.5) and all results are discussed related to all parts of the solution (section 4.4.2).*
4	Have potential interactions with related projects/concepts been considered and addressed?	N	*No related projects were identified in the planning phase of the project.*
5	Is the level of human performance needed to achieve the desired system performance for the proposed solution consistent with human capabilities?	Y	*All HP metrics collected indicated good human performance. Controllers were able to manage their workload appropriately whilst maintaining good situation awareness. User acceptance and trust in automation was consistently high with good task performance.*
6	Are the assessments results in line with what is targeted for that concept? If not, has the impact on the overall strategic performance objectives/targets been analysed?	Y	*Section 4.4.2 shows consistently good results which are in line with expectations.*
7	Has the proposed solution been tested with end-users and under sufficiently realistic conditions, including abnormal and degraded conditions?	N	*Several emergency scenarios (abnormal scenarios) were created during the simulation exercise. Data suggests that the solution does not have negative influence on handling these scenarios.* *Degraded modes have not been tested yet.*

8	Do validation results confirm that the interactions between human and technology are operationally feasible, and consistent with agreed human performance requirements?	Y	*The evidence collected confirms that the interactions between human and machine are feasible. Participants scored very high on user acceptance and trust in automation scores. Workload and situation awareness were acceptable.*
9	Have all relevant SESAR documentation been updated according to the HP activities outcomes (OSED, SPR)?	Y	*HP contribution was made to other documents*
10	Do the outcomes satisfy the HP issues/benefits in order to reach the expected KPA?	Y	*The evidence collected was sufficient to address the issues and benefits and shows promising HP results for the solution to be taken further.*
11	Have HP recommendations and HP requirements correctly been considered in HMI design, procedures/documentation and training?	Y	*All recommendations and requirements are captured in Appendix B & C. They relate to all aspects of HP.*
12	Have the major factors that can influence the transition feasibility (e.g. changes in competence requirements, recruitment and selection, training needs, staffing requirements, and relocation of the workforce) been addressed? Are there any ideas on how to overcome any issues?	Y	*The only relevant transition factor is training controllers on the new solution. This is not anticipated to be an issue given how intuitive the HMI is to use and how quickly controllers accepted to solution into their existing tool set.*
13	Have any impacts been identified that may require changes to regulation in the area of HP/ATM? This includes changes in roles & responsibilities, competence requirements, or the task allocation between human & machine.	Y	*Changes to task allocation between huma and machine are foreseen. They have been acceptable to the human actors at this stage but will need to still be considered going forwards.*
14	Has the next V-phase sufficiently been prepared (additional testing conditions, open HP issues to be addressed)?	Y	*Recommendations have been made and captured in Appendix B.*

5 References

Human Performance

[1] SESAR 2020 Solution 96: TRL6 TVALP - Attention Guidance - Part IV - Human Performance Assessment Plan

[2] SESAR 2020 Solution 96: TRL6 TVALP - Attention Guidance - Part II – Safety Assessment Report

[3] SESAR 2020 Solution 96: TRL6 TVALR – Attention Guidance

Appendix A – Additional HP activities conducted

Not applicable.

Appendix B – HP Recommendations Register

HP Recommendations Register									
Reference	Type of recommendation	Recommendation	Rationale	Assessment source + Reference report	Scope (Air, Air/Ground, Ground)	Concept/ solution Involved	Recommendation status	Rationale in case of rejection	Comments
REC-ARG1.2-10.96-TRL6-HP-001	Human Performance	It is recommended that bad weather situations as well as very low (but still realistic) traffic scenarios are included in further real time simulation work to understand the threshold of the usefulness of Attention Guidance.	A threshold at which AG should be disabled needs to be determined	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

REC-ARG1.2-10.96-TRL6-HP-002	Human Performance	It is recommended to continually monitor the appropriate use of existing functionality such as probing when introducing new functionality such as AG.	The introduction of new functionality needs to be considered against all affected existing functionality to ensure consistency for the user.	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-
REC-ARG1.2-10.96-TRL6-HP-003	Human Performance	It is recommended that a clear design philosophy is applied when further developing and refining alerts and task drivers in the AG functionality.	An alerting philosophy should be followed to ensure consistency of alerts and task drivers in the system.	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-
REC-ARG1.2-10.96-TRL6-HP-004	Human Performance	It is recommended that emergency scenarios remain a key use case for	With increased automation it is important to ensure	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

		further development work.	that huma actors are able to take back full control when required (e.g. during emergency scenarios).						
REC-ARG2.1-10.96-TRL6-HP-005	Human Performance	It is recommended that trust in automation scales remain a key parameter during further development work to ensure Attention Guidance remains transparent to the user.	Only if trust in the system is ensured controllers will accept and use the functionality as intended.	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-
REC-ARG2.1-10.96-TRL6-HP-006	Human Performance	It is recommended that more extreme scenarios (e.g. mass deviations) are	See REC-ARG1.2-10.96-TRL6-HP-004	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

		simulated to ensure the controllers are able to manage sudden loss of AG functionality and having to re-integrate greater numbers of fade-out tracks and detect conflicts against fade-out tracks.							
REC-ARG2.2-10.96-TRL6-HP-007	Human Performance	It is recommended that the timeliness of certain information provided is reviewed (e.g. frequency pop ups)	Based on feedback received on the timeliness of information presentation this should be reviewed during further development work.	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

REC-ARG3.2-10.96-TRL6-HP-008	Human Performance	It is recommended that the role of the supervisor is considered for procedure development around enabling/disabling the AG functionality.	It needs to be established whether the supervisors are required to make decisions around the overall use of AG in different scenarios.	Attention Guidance HPAR	Ground/ Ground	PJ10-Sol.96	<in progress>	-	-

Table 9: HP recommendations

Appendix C – HP Requirements Register

HP Requirements Register									
Reference	Type of requirement	Requirement	Rationale	Assessment source + Reference report if available	Scope (Air, Air/Ground, Ground)	Concept/ solution Involved	Requirement status	Rationale in case of rejection	Comments
REQ-ARG1.2-10.96-TRL6-HP-001	Human Performance	Further development work shall develop procedures to decide when and how Attention Guidance is switched off/ not appropriate to use.	A threshold at which AG should be disabled needs to be determined	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

REQ-ARG1.2-10.96-TRL6-HP-002	Human Performance	The use of the probing functionality should be considered during further development work of AG.	The introduction of new functionality needs to be considered against all affected existing functionality to ensure consistency for the user.	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-
REQ-ARG1.2-10.96-TRL6-HP-003	Human Performance	Attention Guidance alerts and task drivers shall be in line with alerting philosophy guidelines.	An alerting philosophy should be followed to ensure consistency of alerts and task drivers in the system.	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-
REQ-ARG1.2-10.96-	Human Performance	Emergency scenarios shall be considered as part of	With increased automation it is important to	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

TRL6-HP-004		further Attention Guidance development work.	ensure that huma actors are able to take back full control when required (e.g. during emergency scenarios).						
REQ-ARG2.1-10.96-TRL6-HP-005	Human Performance	Further development simulations shall incorporate scenarios where ATCOs quickly have to re-integrate fade-out tracks into their mental model (e.g. weather or mass diversion scenarios).	See REC-ARG1.2-10.96-TRL6-HP-004	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

REQ-ARG2.2-10.96-TRL6-HP-006	Human Performance	Timeliness of information displayed shall remain a key consideration during further development work.	Based on feedback received on the timeliness of information presentation this should be reviewed during further development work.	Attention Guidance HPAR	Air/Ground	PJ10-Sol.96	<in progress>	-	-

Table 10: HP Requirements

Appendix D – HP Log

Not applicable.

-END OF DOCUMENT-

Insert beneficiary's logos below, if required and remove this sentence