

SAFE UNDER ANY CIRCUMSTANCES? VALIDATION OF AIR TRAFFIC CONTROLLERS WORKING UNFAMILIAR SECTORS

R. Hunger*, S. Shetty*, M. Finke*

* German Aerospace Center, Institute of Flight Guidance, Braunschweig, Germany

Abstract

Air traffic management (ATM) faces growing demand, mitigating climate impact, adverse weather, and air traffic controller (ATCO) shortages. The Single European Sky ATM Research (SESAR) programme addresses these through research projects like IFAV3, which investigates whether ATCOs can safely manage sectors for which they lack formal unit endorsement, given specified preconditions. This study presents safety results from a validation exercise evaluating operational feasibility when ATCOs manage sectors outside their standard endorsement portfolio, provided they possess comparable experience in similar endorsed sectors. The objective was to determine whether ATCOs in non-endorsed sectors sustain SESAR safety standards derived from Accident-Incident-Models. We applied the SESAR safety assessment methodology defined in SESAR's Extended Safety Reference E-SRM. Average actual and four-minute predicted losses of separation did not significantly differ between reference and solution scenarios. Only three runs involved separation infringements; none were ATCO-induced or showed significant ATCO contributory factors. All success criteria met targets. While ATCOs noted ATM service quality could potentially be affected, they confirmed operational safety remained unimpaired, despite indicating few more subjective safety issues due to comprehensive consideration of potential situations. The study confirms the novel deployment concept maintains safety levels even for ATCOs without full unit endorsement, validating flexible sector staffing approaches. This IFAV3 validation provides empirical evidence that ATCOs can safely extend their operational scope beyond traditional unit endorsements.

Keywords

Air Traffic Control; IFAV; Safety

NOMENCLATURE

Abkürzungen

		MUAC	Maastricht Upper Area Control	-
		NCA	National Competent Authority	-
ACC	Area Control Center	- NN	Non-Nominal	-
AIM	Accident-Incident-Model	- OJTI	On-the-Job Training Instructor	-
ANSP	Air Navigation Service Provider	- OJT	On-the-Job Training	-
ARTAS	ATM Surveillance Tracking and Data Processing System	- SAC	Safety Criteria	-
ATC	Air Traffic Control	- SESAR	Single European Sky ATM Research	-
ATCO	Air Traffic Controller (Officer)	- TMV	Traffic Monitoring Value amount/60 min	-
ATM	Air Traffic Management	- TSA	Temporary Segregated Area/Airspace	-
BRU	Brussels	- UCS	Unit Competency Scheme	-
CBAS	Competency Based Assessment System	- UTP	Unit Training Plan	-
DLR	German Aerospace Centre	-		
E-SRM	Expanded Safety Reference Material	-		
EU	European Union	-		
FL	Flight Level	-		
HAN	Hanover	-		
ICAO	International Civil Aviation Organisation	-		
IFAV	Increased Flexibility of ATCO Validations	-		
LoA	Letter of Agreement	-		
MTCD	Medium Term Conflict Detection	-		

1. INTRODUCTION

Air traffic management (ATM) faces challenging requirements: meeting growing traffic demand while considering climate impact, responding to adverse weather conditions, and operating during a period of acute air traffic controller (ATCO) shortages [1]. One possible solution to mitigate the problem of ATCO shortages is to enable ATCOs to be deployed more flexibly to a wider range of sectors. Previous projects have revealed that major influence on ATCO performance on foreign sectors is the familiarity with the used air traffic control(ATC) system and is dependent on what the ATCO is used to [2,3]. From this, the hypothesis was formulated that with a pre-deployment assessment of ATCOs individual experiences and competences and an

alignment against a predicted situation in a specific sector at a specific time will indicate options to successfully deploy ATCOs deemed as 'suitable' to those sectors under the analysed predicted situation. The corresponding method was established by [4] using a set of 17 parameters to be compared between historical experiences datasets of ATCOs and predicted situation, i.e. traffic, weather and airspace use. The primary objective of the corresponding validation was to demonstrate that ATCOs can effectively operate a sector under conditions determined to be 'suitable' based on measured ATCO competencies and predicted operational conditions, without requiring prior full On-the-Job Training (OJT) or traditional unit endorsement for that sector. The validation defined "suitability to work" as the absence of any significant negative influence on safety, human performance, and environmental impact, while maintaining efficiency standards. However, this paper focuses on the most crucial - the safety impact, as the first and foremost objective of ATM is to prevent collisions between aircraft [5].

This paper is structured as follows. Section 2 describes the Single European Sky ATM Research (SESAR) Project IFAV3, the corresponding part, i.e. solution, the problem addressed, and the solution approach. Section 3 describes the SESAR approach to assess safety and the validation setup. Section 4 presents the results, which are discussed in Section 5. Section 6 concludes the work.

2. BACKGROUND

2.1. Training, Licensing, Endorsements and Rostering

In the European Union (EU), ATCO training and licensing is highly regulated to align basic principles and organisational structures across member states. The corresponding EU Regulations according to [6] are:

- EU Regulation 2015/340 (technical requirements and administrative procedures for ATCOs' licences and certificates)
- Implementing Regulation 2023/893 (amending EU Regulation 2015/340)
- Implementing Regulation 2025/2143 (introduction of competency-based training and assessment and virtual training of ATCOs)

They introduce and prescribe a common foundation to be complied with by all Air Navigation Service Providers (ANSPs) located in EU member states. This section summarises them. National regulations of the national competent authority (NCA) may supplement them.

As a basic principle, an ATCO may work in a defined airspace or at an aerodrome, when they hold

- an ATC license,
- a valid medical certificate,
- appropriate language proficiency,
- a rating suitable for the ATC service,
- a unit endorsement for the airspace, sector group, working position or aerodrome.

For simplification, this paper does not detail the medical certificate and language proficiency requirements.

An ATCO aspirant undergoes initial training at a recognised ATC training facility, first obtaining a student ATC license and then an ATC license. The student ATC license entitles the ATCO to work in a sector or at an aerodrome under supervision of an On-the-Job Training Instructor (OJTI). Usually, in combination with basic training, the aspirant completes a rating course, including comprehensive simulator training in

the ATC environment to obtain the appropriate rating (Aerodrome Control, Approach Control, Area Control).

A unit endorsement is obtained by On-the-job training (OJT) for the appropriate airspace, sector group, working position or aerodrome. An OJT starts with a unit endorsement course, involving theoretical lessons and simulator training. Subsequent supervised work with real air traffic follow, including de-briefings and instructions for developing skills and working routine. The OJT ends with a theoretical and practical exam. For the first OJT in the ATCO career, passing converts the student ATC license into a full one, entitling the ATCO to work at own responsibility according to the obtained unit endorsement. Rostering is strictly limited to sectors covered by the corresponding unit endorsement.

For every ATC unit, the ANSP defines a unit training plan (UTP) and a unit competency scheme (UCS), to be approved by the NCA. These documents detail training practices and contents, the specifications of unit endorsements linked to the ATC unit, and unit-endorsement maintenance requirements. These usually prescribe a minimum number of working hours on the corresponding sector, airspace and working position in a defined time frame, in addition to regular refresher trainings, as prescribed in EU Regulation 2015/340.

The rostering of endorsed ATCOs to working positions nowadays is a complex process, requiring many conditions to be considered for every ATCO, shift and working position to be staffed. Depending on the size of the unit, this process may involve several dozens of working positions and several hundreds of ATCOs in a typical central-European Area Control Center (ACC). The conditions include:

- Operational constraints:
 - active unit endorsements held by the ATCO, determining sectors / airspaces / working positions in which they may be deployed,
 - working positions to be opened, depending on expected traffic demand,
 - unit endorsement maintenance requirements defined in the local UCS (i.e. working hours to be served on a specific working position before a specific date to avoid losing the unit endorsement for this ATCO),
 - requirements due to ongoing OJT (e.g. whether the ATCO holds an OJTI endorsement, or which working positions are excluded due to parallel, often prioritised OJT of a student ATCO)
 - office times or other training lessons of the ATCO
- work-contractual constraints:
 - work-time restrictions due to part-time contracts,
 - maximum number of night shifts and fair balance between early shifts and late shifts,
 - maximum number of weekend shifts,
 - minimum required rest times between shifts
- personal constraints

Creating a rostering schedule for every day and shift, up to months in advance, can be extremely complex for large ATC units. It must consider all constraints for every ATCO while satisfying the demands with sufficient spare personnel to compensate for unplanned absences (e.g. sickness) or unexpected needs to open more sectors. ANSPs therefore use dedicated rostering tools tailored to local needs, local work conditions and agreements with unions. However, due to the large number of constraints to be considered for the actual rostering of ATCOs, the system appears very rigid and restrictive. Therefore, it is not very robust against unforeseen events or temporary / permanent staff shortages. Today this regularly leads to situations where sectors of sector group A cannot be staffed appropriately to meet traffic demand due to an unplanned

or persistent lack of personnel, while sector group B experiences overstaffing. ATCOs from sector group B cannot support sector group A due to missing endorsements, as they need months of training to obtain the second endorsement for sector group A as well. At the same time some ACCs in Europe experience almost no traffic during night in the overall airspace of responsibility. Current sector group endorsements allow a combination of sectors within the same sector group, but not beyond. Therefore, all sector groups need to be staffed with two ATCOs during night, even when all sector groups together control not more than a few flights at the time. As a result, the ATCOs are totally undertasked, but are not available for later shifts of the day with possible understaffing, because of prescribed minimum rest times. Such situations lead to delays in air traffic. According to IATA, delays caused by ATCO staff shortage increased by 192.1 % since 2015 in Europe [7]. Therefore, an increased flexibility in ATCO rostering is an eagerly awaited improvement in ATM, which is for example noted by CANSO as well considering the expected traffic increase in the next years [8].

2.2. SESAR's IFAV3 Project

Increased flexibility of ATCO validations (IFAV) has been a research topic within the SESAR Programme since 2016, and started with the SESAR Project PJ.10-06 Generic / Non-Geographical Controller Validations. It has been continued in projects PJ.10-73 IFAV and PJ.33, and has been brought to a high maturity level within SESAR's IFAV3 project. This recent project started in July 2023, and finished work in June 2026 involving a consortium of 10 European partners, i.e. four ANSPs, five research institutions, two universities and one large ATM supplier, lead by the German Aerospace Centre (DLR). It investigated several strategies towards a more flexible ATCO rostering, involving dedicated tool support, adjusted ATC procedures, smart controller experience assessment and required competency estimation, smart sector grouping, common unit competency scheme framework and increased automation to enable generic controller validations [9]. IFAV3 provided detailed concepts, prototypes of suggested enablers and guidelines for implementation, and performed a comprehensive analysis of needed regulatory changes with strong support from EASA. This project is not only seen as important milestone towards a more resilient ATM system, and as a countermeasure to increasing ATCO staff shortage, but also as an essential enabler for SESAR's virtual centre concept. The project has been awarded in 2026 with the Seamless Sky Runner Up Award of CANSO [10].

2.3. Solution 0542 - Smart Controller Experience Assessment and Required Competency Estimation

The strategy discussed in this paper, called 'Solution 0542', aims at the idea of competency-based rostering, meaning a paradigm shift in ATCO training from 'the endorsed ATCO must be sufficiently trained to be able to handle all thinkable situations' towards 'the IFAV endorsed ATCO must be sufficiently trained to be able to handle exactly the situation that is predicted to happen, plus sufficient training for unpredictable events'. This paradigm shift opens for example the opportunity to deploy ATCOs during night shift to less busy sectors with limited training only, provided the ATCO has sufficient competence to handle a few movements per hour under night conditions. This means she has been theoretically and practically trained just a couple of days to get familiar with sector geometry, night procedures, maybe simulation runs and a few days of OJT, based on a separate 'IFAV endorsement' in

addition to traditional endorsements defined for an ACC. The strategy under investigation foresees two obligatory processes to be considered before an ATCO is actually deployed to a sector with IFAV endorsement to ensure ATCOs are sufficiently trained:

- Process 1: 'Smart Experience Assessment' is an automatic background process, where operational data recordings, such as radar tracks and Air-Ground communication are analysed, and used to calculate a set of defined indicators per ATCO. This scheme of indicators has been designed to quantify the experience the ATCO has gained over an extended period of time - going back at least 90 days - in solving ATC tasks of a specific complexity (considering e.g. traffic diversity), under specific circumstances (e.g. sector geometry and weather situation) with which effort. This is intended for live operations, but can also be supplemented by experience gathered during OJT and simulation runs. These indicators shall correspond to 'observable behaviours' as defined by ICAO, which are in turn an indicator of the ATCO's competence level [11].
- Process 2: 'Required competency estimation' is an automatic background process, where available forecasts and flight plans are used to calculate the same set of defined indicators, but this time per working position, sector and shift to be staffed in the near future. These calculated indicators shall quantify the type and complexity of the ATC tasks expected during a given shift in a sector at a given working position.

The actual rostering is done by comparing the determined experience of an ATCO from process 1, with the predicted demand calculated in process 2. In case of a match, i.e. the predicted demand is equal or lower than the measured experience, the ATCO is classified as 'suitable' and can be deployed to this sector during the considered working period. In case of a mismatch, the ATCO is classified as 'not suitable', and a deployment to this sector during the considered working period is prohibited. In this case, either a deployment to another suitable sector needs to be considered, or the ATCO is deployed according to his/her still maintained traditional endorsement in a specific sector group. As a result, this new concept makes ATCO rostering more complex, as additional constraints for deployment are added. However, the benefit of this approach is the availability of additional deployment options per ATCO to achieve a more flexible ATCO allocation to sectors, compared to the ones offered by traditional endorsements, with only little additional training needed. For implementation, these calculations and resulting constraints for deployment need to be added to existing rostering tools, and an additional display interface for ACC supervisors can be discussed.

3. METHOD

For the validation of the new concept the Expanded Safety Reference Material (E-SRM) guides the safety aspect. Core aspects are presented hereafter before the validation and measurements are described.

3.1. Validation

3.1.1. SESAR Safety Methodology

The E-SRM document [12] is a methodological guide for SESAR projects to ensure that changes to ATM systems comply with strict European safety requirements, by setting out a structured, iterative process across the entire system life cy-

cle. The safety assessment for Solution 0542 was conducted in alignment with the E-SRM as reported in the Safety Report [13]. The most important aspects for understanding the experimental design of the validation and the metrics of the Safety Assessment Report are summarised in the following: For Technology Readiness Level (TRL) 6 (pre-industrialisation phase) as aimed at by the solution E-SRM requires derivation and validation of safety requirements and their assessment under normal, abnormal and faulted conditions, consider effects on safety nets and to check achievable by real time simulation. In this context what is acceptably safe is referred to as the Safety Criterion(a) (SAC). The SACs for Solution 0542 are derived from the relevant Accident-Incident-Models: Mid-Air Collision Risk in En-Route and Airspace Infringements Model. The key SACs include:

- There shall be no increased number of ATC-induced tactical conflicts due to the deployment of IFAV ATCOs based on competency.
- There shall be no increased number of ATC-induced imminent infringements due to the deployment of IFAV ATCOs based on competency.
- There shall be no increase in the number of unauthorised aircraft entering adjacent sectors due to the deployment of IFAV ATCOs based on competency.
- The IFAV concept has no negative impact on the IFAV executive ATCO's perceived level of safety.
- The IFAV concept has no negative impact on the IFAV planner ATCO's perceived level of safety.

3.1.2. Experimental Design and Scenarios

The validation was conducted between 4th and 13th November 2025 in Maastricht Upper Area Control (MUAC) Center's simulator environment which is almost identical to the operational system, enabling TRL 6 validations. MUAC airspace is divided into three sector groups: Brussels (BRU), Hanover (HAN) and DECO (see FIG 1) each consisting of two unit endorsements. ATCOs holding both unit endorsements are allowed to work all sectors and their combinations of the group. Between 6 and 7 participants per sector group were recruited on a voluntary basis for the validation, with some ATCOs holding a third unit endorsement of a second sector group. Each participant completed five 45-minute simulation runs categorised to the following run types:

- Reference runs (2 runs): ATCO holds a unit-endorsement used to establish a baseline performance metric.
- Solution runs (3 runs):
 - Suitable: Conditions predicted as manageable for the specific ATCO.
 - Not-Suitable: Conditions predicted as unmanageable - to test the prediction model's boundaries.
 - Non-Nominal: Either predicted as suitable or not-suitable, but including unforeseen events.

A inter-between subject design was used for the validation, where the reference and solution runs to be compared were conducted by different ATCOs executed the same scenarios, e.g. scenario A could be reference run from ATCO A, solution suitable run for ATCO B and not-suitable run for ATCO C. The prediction of ATCO suitability was based on a comparison of parameters extracted from previous recorded life ATCO shifts on their unit endorsed sectors against these parameters extracted from the various scenarios as described in [4]. The previous experienced shifts included data from 18th July to 15th October 2025 (90 days) leaving an unavoidable gap between last considered life experience and validation execution. The considered ATCO experiences were based on operational traffic and thus not predictable for the validation planning.

Two sectors per sector group (see FIG 1) and two scenarios per sector (see TAB 1) were prepared for simulation use during the validation to ensure as far as possible a comparison outcome per ATCO to satisfy all solution run types, i.e. suitable and not-suitable. The sectors and scenarios per group



FIG 1. Simplified View of Sectors Validated. Coloured similar per Group. Military Temporary Segregated Airspace (TSA) Included in Both Scenarios if Aircraft Symbol is Black and in one Scenario if Aircraft Symbol is Grey.

were selected based on MUAC simulation expert judgement to be as diverse as possible. All sectors extended from Flight Level (FL) 245 to FL 660. Traffic conditions were designed to reflect realistic summer period day shifts. Additionally, sector-specific difficulties were included in one scenario, called 'high' scenario in the following, and excluded in the other, called 'moderate' scenario in the following (see TAB 1). Traffic

TAB 1. Scenario Overview

Scenario	TMV	Military	Number Suitable/ Not-Suitable
BRU West moderate	100%	no	1/5
BRU West high	130%	no	0/3
Luxembourg mod.	100%	yes	4/1
Luxembourg high	130%	yes	0/4
Delta moderate	98%	yes	4/2
Delta high	135%	yes	0/3
Holstein moderate	100%	no	5/1
Holstein high	140%	no	1/2
Ruhr moderate	75%	yes	2/2
Ruhr high	75%	no	0/2
HAN East moderate	87%	no	2/1
HAN East high	120%	yes	0/3

Monitoring Value (TMV) is given relative to the maximum allowed number of aircraft per hour for each sector. Military activity areas were included in half the scenarios oriented on realism (see also FIG 1). The conducted runs per run type for each scenario are given in the last column. Four scenarios were compared to each ATCOs experience measures for selection of individual planned runs on 30th October 2026. If possible one scenario was selected as suitable run and one as not-suitable

depending on satisfaction of requirements by those experiences. One artificial example for sector group BRU is given in TAB 2. The non-nominal (NN) runs were based on the 'moderate' scenarios including two events separated by time: first an emergency descent and second a turbulence report. The NN run was selected to ensure ATCOs are not faced with the same 'moderate' scenarios (except events) twice to minimise training effects.

TAB 2. Scenario Allocation Example

Run Type	Scenario	
	Choice	Selected
Suitable	BRU West moderate ~~BRU West high~~ Luxembourg moderate ~~Luxembourg high~~	Luxembourg moderate
Not-Suitable	~~BRU West moderate~~ BRU West high ~~Luxembourg moderate~~ Luxembourg high	BRU West high
Non-Nominal (NN)	BRU West moderate NN ~~Luxembourg mod. NN~~	BRU West mod. NN

The validation procedure was divided into three distinct phases:

- Pre-validation:
 - Participant's measured experience were compared to scenario requirements and suitability were calculated.
 - Participants completed 2 hours of Computer-Based Training for those sectors they were planned for the validation and did not hold a unit endorsement for.
- During Validation:
 - Gathering subjective data/assessment from participants prior each run (see Section 3.2)
 - Voluntary ATCOs underwent simulation in an exam-like checkout environment
 - Utilising discrete objective data (see Section 3.2)
 - Gathering feedback and subjective data/assessment from different validation roles after each run (see Section 3.2)
- Post-Validation:
 - Analysis and comparison of data

The validation included several distinct roles:

- **Executive ATCO:** ATCO controlling traffic and communicating with simulation pilots, focus of the validation, under assessment, not unit endorsed
- **Planner ATCO:** ATCO coordinating with neighbouring sector, unit endorsed
- **Assessor:** ATCO assessing the executive ATCO, unit endorsed, personal union with Planner ATCO for validation
- **Adjacent:** Controlling neighbouring sector(s) to the sector under validation, unit endorsed, personal union of executive and planner ATCO
- **Simpilots:** simulation pilots executing executive ATCO and adjacent instructions

Each participant conducted her solution runs on one of the two sector groups they had no unit endorsement in the role of the executive ATCO. ATCOs with a third unit endorsement conducted their solution runs on the sector group they had no unit endorsement for.

The focus was set to executive ATCO, while the planner ATCO served as assessor at the same time to reduce supporting staff requirements for the validation. Thus the planner

was unit endorsed for all run types. All runs were supported by simpilots and one adjacent for all neighbouring sectors. Depending on the scenario additionally a unit endorsed adjacent served as executive and planner ATCO for one closely related neighbouring sector, leaving all other sectors for the first adjacent. Each participant served as planner ATCO and assessor five times in addition to the five runs as executive ATCO. The schedules were individually spread over two full days either consecutive or separated by one free day due to overall schedule.

As far as possible the following allocation rules were applied to avoid biases and ordering effects:

- Assessors were not allocated to the same executive ATCO twice; in any case not for two consecutive runs
- Assessors were not allocated to the same scenario twice
- Run type, i.e. suitable, not-suitable and NN, order was varied between participants

3.2. Measurement

Both subjective and objective assessments were conducted to evaluate the outcomes of the validation trials. Subjective assessment included:

- Self assessment:
 - Pre- and post run questionnaire to rate the agreement on a scale from 1 to 6 if they can perform their ATC tasks safely. Before each run they were informed about expected situation, i.e. relative and absolute traffic amount, active military temporary segregated airspace (TSA), and expected weather.
 - adapted Competency Based Assessment System (CBAS) [14]: beneath others rating *safety outcome* on a 4-point-scale from significant below standard to above standard. Operationally used to assess ATCO trainees during OJT.
 - Post validation question to rate the agreement on a scale from 1 to 6
 - Post validation open feedback
- CBAS *Safety Outcome* by assessor and adjacent.

Objective assessments used ARTAS (ATM Surveillance Tracking and Data Processing System) data obtained for all runs. This data consists of various surveillance information such as aircraft callsign, position, altitude, speed and heading at various timestamps throughout the simulation runs.

The following metrics were calculated using the ARTAS data and can be categorised into:

- Safety-critical metrics
 - Separation Infringement: A situation in which the required minimum horizontal and vertical separation between any two aircraft is violated at any point in time. In this work, a minimum horizontal separation requirement of 5 NM (9.26 km) and a minimum vertical separation requirement of 1000 ft (304 m) were applied. Only aircraft above FL 243 were considered. This metric represents existing separation infringements at a given point of time and the minimum distances.
 - Four-minutes to Separation Infringement: Predicts whether two aircraft will violate minimum separation 4 minutes after the observation time. Aircraft positions are extrapolated 4 minutes ahead using their current state, and a separation check is performed at those predicted positions. The metric indicates if a conflict exists at that look-ahead point.
 - Medium-Term Conflict Detection (MTCD) Number and Duration: The MTCD tool included in the simulation environment was used, which indicate potential conflicts over a medium-term time horizon (<20 min), considering an aircraft's projected trajectories over the medium-term.

These metrics indirectly show ATCO induced conflicts that are not already included in the scenario and the duration required by the ATCO to react to resolve them.

- Additional general metrics with potential safety impact
 - Letter of Agreement (LoA) Adherence: Measures the extent to which aircraft transfers between adjacent sectors are carried out according to the specified LoA procedures.
 - TSA Penetration: Represents whether an aircraft violated a restricted military area.

4. RESULTS

The results of the measures from Section 3.2 are reported below in the same order.

4.1. Subjective Measures

The participants were asked to rate their agreement to the two statements: *I expect being/ was able to carry out my ATC tasks safely* before and after the run to capture the perceived safety. FIG 2 shows the averaged rating pre-run in violet and post-run in green for reference runs, solution suitable runs and solution not-suitable runs. The differences between reference

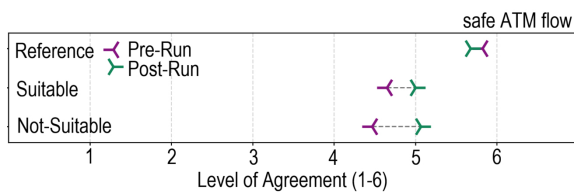


FIG 2. ATCO self-assessment pre- and post-run averaged per run type, adapted from [15]

and solution scenarios (both types) are significant. However, all ratings stay clearly within agreement to safe operation. Interestingly, all ratings increased from pre-to post-run reply. Beside other parameters as reported by [14] the CBAS captures also safety outcome to be rated on the 4-point scale from *significant below standard* to *above standard*. The executive ATCOs as well as the assessors and adjacent ATCOs were asked to rate safety on this scale. FIG 3 shows in total

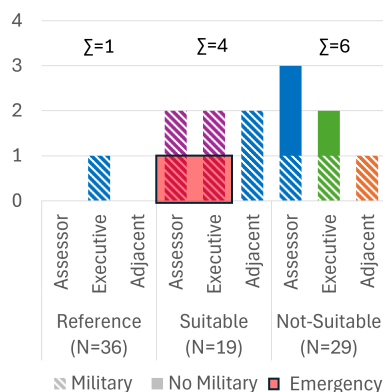


FIG 3. CBAS Safety Outcome. Same colours per run type across roles indicate same runs

only 13 CBAS ratings regarding safety stated *slightly below standard* (of potential 275). Most of the corresponding runs included active military TSA, one included an emergency. Interestingly, the rating between the assessor and the executive matched only twice.

All participants that conducted runs as executive ATCO were asked to rate their agreement to the statement: *I feel that the*

endorsement method has not negatively affected my ability to carry out my ATC tasks safely after their participation on a scale from 0 to 5. Only one participant conducted runs only in the role of the executive ATCO. One participant that conducted runs only in the role of assessor was not asked to rate. FIG 4 groups the results in correspondence to the number of runs the participant had executed under the solution run types suitable and not-suitable.

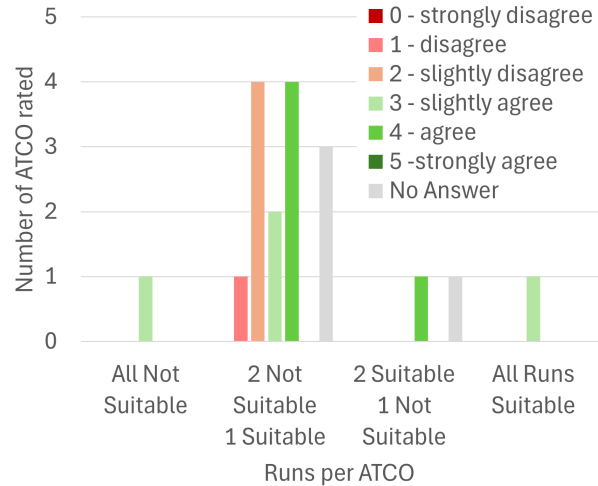


FIG 4. Post-Validation Concept Safety Ratings

Additional feedback noted tight or unsafe clearances solved in time, partly attribute this to a lack of theory knowledge.

4.2. Objective Measures

Logging of aircraft positions enabled post-validation analysis regarding separation infringements as most important safety measure. For technical reasons 6 out of 90 runs could not be recorded successfully. TAB 3 summarises number and severity of all recorded infringements with less than 1000 ft and 5 NM distance between any pair of aircraft. It also states a

TAB 3. Separation Infringements

Metric	Run Type		
	Reference	Suitable	Not-Suitable
Recorded runs	36	19	29
Infringements	4	1	2
Average min. lateral distance	4.64	4.5	4.56
Average min. vertical distance	859	832	845
Reason	Run time finished	Level Bust	Emergency Descend

reason based on review of screen- and voice recordings. *Level Bust* refers to a situation where a pilot is late in transitioning from climb or descend to levelled flight and thus overshoots the instructed FL. FIG 5 shows an example of the separation infringements recorded for the not-suitable run followed by an emergency descend scripted in the NN scenario.

A separation infringement is a seldom event. A more frequent situation that increases probability of infringements with increasing occurrence is used to better assess safety impact: Unresolved conflicts that would lead to an infringement within



FIG 5. Separation Infringements not-suitable Run due to Emergency Descend

the next 4 minutes. That means based on the current aircraft state without the ATCO intervention the distance between any two aircraft will fall below 1000 ft and 5 NM. The lead time of 4 minutes is a value based on MUAC experience with trainees. It could hint that a potential infringement is not identified yet or no solution has been found yet. TAB 4 showing 58 occurrences for reference runs, 24 for solution suitable runs and 45 for solution not-suitable runs.

TAB 4. Unresolved t-4 min Infringements

Run Type	Reference	Suitable	Not-Suitable
Recorded runs	36	19	29
t-4 min conflicts	58	24	45
Averaged per run	1.6	1.3	1.6

Operational ATC System typically include one or even several conflict detection functionalities that should ensure to avoid separation infringements or even unresolved conflicts in the vicinity of 4 minutes as described above. Medium Term Conflict Detection (MTCD) is one of those that was included in the validation system and can be used as safety metric. Comparing same scenarios between run types, i.e. reference, solution suitable and solution not-suitable, for number and duration of MTCD indications can highlight safety impact of the proposed concept. If there is no impact the same number and equal duration of indication for the same scenario can be expected. FIG 6 shows the average count of MTCD indications per scenario per run type. The number of considered

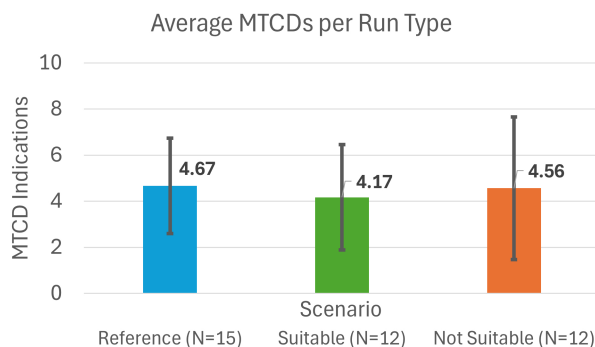


FIG 6. Mean Count of Medium Term Conflict Detection Indications

runs is given at the bottom of each bar. They deviate partly from those given in TAB 1 as not all runs could be recorded successfully and not all scenarios included MTCDs. Runs with military TSA and NN situation are included. The suitable run show fewer MTCD on average, but standard deviations of all run types overlap.

The overall average duration of MTCD indications per run type as shown in FIG 7 are lower for solution - both suitable and not-suitable - than for reference runs.

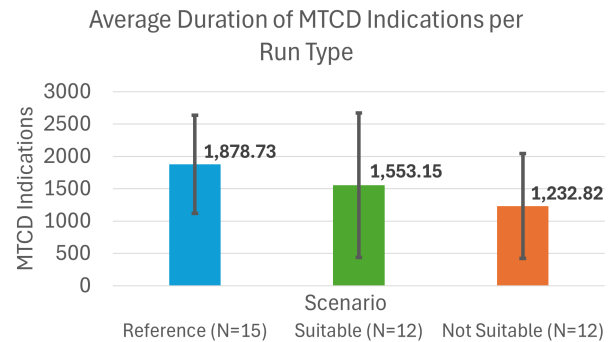


FIG 7. Averaged Medium Term Conflict Detection Indication Duration

Letter of Agreements (LoA) regulate how flights are transferred from one ATC unit to another, be it across ANSPs or sectors within an ACC. A deviation is not by itself safety critical, however it can become so. Preplanning of ATCOs are based on the planned entering conditions of aircraft approaching own sector. A deviation might lead to an infringement in the vicinity of a sector boundary if it is not coordinated and agreed on beforehand.

ARTAS data were recorded for a range between 37 and 97 aircraft per run and allowed for analysis of the adherence to the LoA's prescribed transfer FL or to be more precise the FL at transfer of control point. This metric could only be applied 45 runs for a share of 1 to 25 aircraft, as flights that maintained the FL throughout the sector and flights for which no LoA procedure applied could not be analysed. Flights that met the transfer FL within a range of ± 300 ft and those that were further descended by the following sector after transfer of control were counted as adhering to LoA. A manual check additionally identified those that were not able to leave the airspace within the run time and missed the transfer FL for that reason. All other aircraft were counted as not adhering to LoA. TAB 5 shows that under suitable run type significant ($p < 0.04$) less flights were transferred at a wrong FL. In relative values between 11% for suitable run type and 16% for not-suitable run type of analysed flights deviated from LoA.

TAB 5. Adhering to Transfer FL - averaged per run

Flights	Run Type		
	Reference	Suitable	Not-Suitable
All	55.9	43.4	59
Analysed	13.5	6.8	14.0
Adhering	9.3	5.4	9.1
Not-adhering	1.8	0.8	2.2

An airspace infringement again is by itself not safety critical. However, without coordination for possible crossing or scratching an infringement with an aircraft inside the reserved airspace that could potentially be unknown (and not visible) to the ATCO is possible. TAB 6 shows the number of runs per run type that included infringements with the TSA active during scenarios as indicated in TAB 1, their share of all possible runs, the absolute number of intruding aircraft, and the average per run. The lower half of the table shows the duration of the infringements with at least one aircraft at a

TAB 6. Temporary Segregated Airspace Infringements

Run Type	Runs		Infringing A/C	
	Total	Affected	Total	Average
Reference	18	3 (17%)	15	0.83
Suitable	10	2 (20%)	2	0.20
Not-Suitable	15	4 (27%)	16	1.07

Run Type	Infringement Duration [s]			
	Single Aircraft		Multiple Aircraft	
	Total	Average	Total	Average
Reference	2500	138.9	440	24.4
Suitable	280	28	0	0
Not-Suitable	2469	164.6	160	10.7

time - summed up and averaged per run type - and the part of this duration with more than one aircraft at a time as a sum and as an average per run type.

Considering all measures per individual run and comparing the results to the ATCO suitability assessment the accuracy of the method assessment can be evaluated (see TAB 7). Neg-

TAB 7. Correctness of Method Prediction of ATCO Suitability

Measures Indicating Negative Safety Impact	Prediction Correctness	
	Suitable	Not-Suitable
0	Correct	Wrong
1 if MTCD, t-4min or adjacent assessment	Likely Correct	Wrong
1 if planner assessment	none occur	Likely Correct
1 or 2 if 1 MTCD, t-4min	Undecided	
≥ 3 or self-assessment or $MTCD > \text{Max}(MTCD_{Ref})$	Critical Wrong	Correct

ative safety impact was assumed for any *slightly/significant below standard* assessment of any role, a post-run safety self-assessment of 3 (neither nor) or below, a t-4 min conflict number and MTCD number respectively above average of the corresponding reference runs, any TSA infringement and LoA deviation. FIG. 8 shows only 3% of all runs were assessed as suitable while three or more measures indicate negative safety impact.

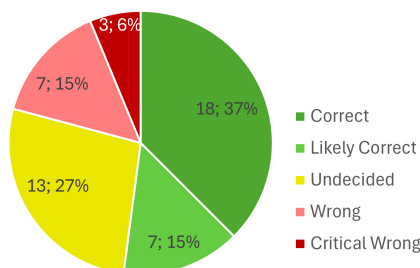


FIG 8. Method Assessment Compared to Validation Results

5. DISCUSSION

5.1. Limitations

Although, 19 ATCOs participating in a human-in-the-loop simulation is a comparable high number, for most measures statistical significance can not be reached. Findings must be interpreted qualitatively, especially if analysing subsets of the data set. Furthermore, generalisability is limited as system, sectors and ATCOs were solely used from MUAC, to fulfil TRL 6 requirements [16]. However, a different system will provide different support that could have influence on safety results. On the other hand, by using several sectors of all sector groups of MUAC and corresponding ATCOs the results are generalisable to a wide range of sectors and ATCOs. The validation itself lead to a major adaption compared to the concept: In focusing on the executive ATCO, the results are limited to this role. Additional feedback draws attention to a potentially more critical deployment of the concept to the planner ATCO. Furthermore, as the planner ATCO always holds a unit endorsement, an influence in terms of rather positive (less safety critical) outcome is possible. The high valuable support of planner ATCO was emphasised by the participants. Planner ATCOs were asked to support the radar ATCOs as usual and to assess them in addition. This double role could have prevented a clean distinguishing. Besides, it will not be represented in objective results. Additionally, the mental load due to that support given to IFV executive ATCO was not measured. In contrary, the very short theory training, as another validation limitation, could have led to worse results than could be expected from the concept. Another validation shortcoming that could not be addressed with reasonable effort is an investigation of what will happen after an ATCO had not seen a sector for several weeks or even longer. The concept considers such situations in principle, however it was not evaluated during the validation as this would have to be performed in two separate time slots with a sufficient distance of several weeks in-between. Furthermore, resource and schedule planning for validation did force a gap between considered past experiences and validation assessment and execution of two weeks. Therefore, immediate past experiences of ATCOs were not considered, which could have had a influence on suitability assessment.

5.2. Subjective Measures Discussion

Perceived safety remained acceptable for both executive and planner ATCOs. FIG 2 shows that ATCOs were typically more sceptical about their own competence before a run. As expected, they felt more comfortable and rated their ability to work safely higher under reference run type, but in both solution run types they agreed to work safely in pre- and post-view as well. ATCOs were not primed with the method's suitability assessment. Assuming a correct assessment, the insignificant difference between suitable and not-suitable results, indicates that ATCOs could not self-assess their suitability. A possible explanation is that executive ATCOs did not self-detect safety issues. However, the method itself is assessed first in this paper. Perception aligned well with CBAS self-assessment: only five times an executive ATCO rated their own run as *slightly below standard*, where standard denotes the level of a trainee receiving unit endorsement at checkout. This low number limits generalisation, but one such rating occurred in a reference run, showing perceived safety can insufficient even in endorsed sectors. CBAS assessments of other roles almost never match those of the executive ATCO when at least one role indicated *slightly below standard* (see FIG 3). This could

be because the other roles respectively did not notice the relevant situations or perceived the safety aspect differently. Only two runs had consistent *slightly below standard* ratings from executive and assessor roles; both included TSA, one included NN situations, and both were assessed suitable. This is critical, especially because the suitable run type had only 60 total assessments, whereas not-suitable had 99 and reference 116. After conducting all runs as executive and assessor the participants were asked for their safety assessment of the concept (see FIG 4). 5 of 18 ATCOs rather disagreed, while 9 agreed and 4 did not answer. However, participants did not know for which runs they were assessed suitable or not. Most had two not-suitable and one suitable run, which may have led to a more negative overall view. Without knowing the method's assessment, a concept-level safety assessment is hardly feasible, because ATCOs also considered runs that the method would not have allowed them to work. They nevertheless conducted these runs during validation to evaluate negative method assessments. Providing the assessment would have influenced run perception and thus was intentionally withheld. A split by number of solution run types conducted showed no tendency to rate the concept worse when more not-suitable runs were faced. The very small subgroup sizes or the perception from runs in the assessor role may explain this.

5.3. Objective Measures Discussion

Assessor safety perception is important for capturing potential negative concept influence that was only prevented by assessors acting as unit-endorsed planner ATCOs. However, a negative perception alone does not mean the traffic was controlled unsafely. The most important safety metric is the number of infringements, i.e., the distance between any two aircraft being less than 5 NM and 1000 ft. Though, all recorded separation infringements (see TAB 3) were non-ATC-induced: they were either recorded because simulation continued after run time ended, or were unavoidable by the ATCO due to a level bust or emergency descend from the (simulation) pilot side. The rather regular occurrence of infringements that would occur in 4 minutes without ATC intervention (see TAB 4) showed similar values per run for all run types, slightly lower for suitable run type. Such unresolved conflicts are not ATC failures and therefore not safety issues in themselves, but they can be a step toward separation infringement. A higher number increases their potential and probability. A lower number may still lead to more actual infringements. The same applies to MTCD indications. Comparing runs with identical scenarios can reveal ATCO actions that led to additional MTCDs beyond the default ones, assumed to be represented by reference run type (see FIG 6). A slight tendency toward lower values was observed for solution suitable run type compared with reference and not-suitable, but the absolute number of considered runs was rather low, and no consistent across all scenarios. Such comparisons require sufficiently many suitable and not-suitable participants for the same scenario. Despite preparatory effort, only 7 of 12 scenarios had both suitable and not-suitable method assessments, and only two had at least two participants in both run types, insufficient for statistical conclusions. The average duration of all MTCD indication is 5 respectively 10 minutes lower for solution suitable and not-suitable run type (see FIG 7). A plausible reason for better performance under solution run types for t-4 min conflicts and MTCD indications is that ATCOs were less relaxed and more eager to detect potential conflicts in sectors for which they did not hold a unit endorsement. In endorsed sectors, ATCOs may be more observant because they know typical flows and conflict points

and when action is required versus when it is possible to wait because it is unlikely to become an infringement. It has to be admitted that most of the high scenarios were assessed as not-suitable and not included here for fair comparison, if no suitable run was available (see TAB 1). Including all not-suitable runs regardless of scenario would substantially worsen results. This could be attributed to the higher traffic. However, this is rather a hint for strong influence of traffic amount and complexity to suitability (and well designed scenarios) than for a less effective and correct assessment by the method. Specifically, there was no increased number of ATC-induced tactical conflicts due to the deployment of IFAV ATCOs based on competency.

TSA infringements and deviations from LoA confirmed better results for solution suitable run type. Both could potentially include correctly coordinated situations that were not fed back to the system and therefore not correctly recorded, but this potential influence was the same across run types. Thus, there was no increase in unauthorised aircraft entering adjacent sectors due to the deployment of IFAV ATCOs based on competency.

In conclusion, no negative safety impact due to the concept was identified, and competency-based rostering functions as intended without degrading existing safety barriers. However, one safety issue caused by the validation process was documented: ATCO confidence levels might initially be affected by reduced traditional sector training, potentially influencing safety perception [15]. ATCOs typically underrated their abilities compared with objective safety results. This should be mitigated through structured feedback, targeted training, and continuous competency monitoring.

Finally, the results presented in this paper are limited to safety aspects, whereby no statement about efficiency could be made. However, the validation took this into account. Meier et al. [14] discussed, amongst others, efficiency results and concluded that there was no indication of reduced traffic handling efficiency, even showing slight improvements with uninterrupted altitude changes increasing from 88.6% to 91.6%. Hagl et al. [15] discussed human performance and revealed subjectively higher mental workload when operating non-endorsed sectors, but it remained within acceptable operational limits with no evidence of mental stress.

6. CONCLUSION

The validation evaluated whether air traffic controllers (ATCOs) can safely manage sectors outside their traditional unit endorsement portfolio when deployment is based on a competency-based experience assessment rather than a fixed sector endorsement. Nineteen voluntary ATCOs from MUAC conducted 90 simulation runs across moderate, high, and non-nominal traffic conditions, including emergency descents and active military temporary segregated airspace. All five safety criteria defined were met: no increase in ATC-induced tactical conflicts, no increase in ATC-induced imminent infringements, no increase in unauthorised aircraft entering adjacent sectors, and no negative impact on the perceived level of safety of either the executive or the planner ATCO. The only three recorded separation infringements were non-ATC-induced (simulation overrun, level bust, scripted emergency descent), and four-minute predicted losses of separation did not significantly differ between reference and solution run types.

Thus, do ATCOs work safe under any circumstances? Yes and No. Despite it was not, and will not be possible to validate 'all' circumstances, the results confirm that the IFAV de-

ployment concept for executive ATCOs maintains established safety levels across the tested operational envelope. Even ATCOs assessed as not-suitable for a given scenario did not perform significantly less safely than the unit-endorsed reference run type: Their averaged separation metrics, unresolved four-minute conflicts, and medium-term conflict detection durations were comparable to, or in trend-level better than, the reference. This suggests that the competency-based assessment provides a meaningful safety margin: even when the method predicts an ATCO to be not-suitable, the resulting safety level remains within acceptable bounds. At the same time only 3% were wrongly indicated to be suitable, while average safety results outperformed not-suitable ATCOs. ATCOs, however, did not always perceive this safety level, consistently underestimating their own competence relative to the objective results, which aligns with findings reported in [15].

The detailed logging of data and assessor feedback and the rigorous counterbalancing of scenarios ensure that the validation provides empirical evidence that competency-based sector deployment can maintain current safety standards. It thereby supports a more flexible and resilient ATCO rostering model, with additional deployment options per ATCO and only little additional training needed. A promising avenue for future work would be to adapt and further refine the present methodology and validate its applicability to a supervisor role within an operational rostering tool environment.

Acknowledgement

The IFAV3 project has received funding from the SESAR 3 Joint Undertaking under grant agreement No 101114683 under European Union's Horizon Europe research and innovation programme. The authors gratefully acknowledge the invaluable contributions of Joachim Keck (Keccon), Andreas Sauerwein, Gratiel Marius Marin, Wiebe Koorevaar, Mariela Konova, Matthew Harrison and Anette Cummings (all MUAC). Their deep knowledge, extensive experience, and the data provided were pivotal to this work's success, and we thank them for the considerable time, effort, and manpower devoted to the project. Special thanks are due to Simas Jegelivičius for his outstanding, dedicated organisational support: coordinating all onsite preparatory meetings, managing the support staff, and continuously supplying clarifications that kept the work moving forward behind the scenes, making the research logistics seamless. We further extend sincere appreciation to all participating Air Traffic Controllers (ATCOs) in the validation phase, whose willingness to engage, insightful feedback, and even friendly dinner tips enriched both the scientific rigor and the collaborative spirit of this study.

Contact address:

robert.hunger@dlr.de

References

- [1] IFATCA. Staff Shortage Survey – EUR Region. <https://ifatca.org/staff-shortage-survey-eur-region/>, 19.09.2024.
- [2] SESAR Joint Undertaking. SESAR Solution PJ.33 W3-01a Validation Report (VALR) Deliverable D2.1.005 Edition 01.00.01, 31 March 2023.
- [3] Michael Finke, Robert Hunger, Lukas Tyburzy, Jana Meier, and M Jameel. Using tool support to enable a more flexible deployment of Air Traffic Controllers to Sectors. Deutsche Gesellschaft für Luft-und Raumfahrt-Lilienthal-Oberth eV, 2024.
- [4] Jana Meier, Marco-Michael Temme, Robert Hunger, and Michael Finke. Unlocking potential: Enhancing air traffic controller deployment through dynamic competency-based rostering. In 2025 AIAA DATC/IEEE 44th Digital Avionics Systems Conference (DASC), pages 1–10, 2025. DOI: [10.1109/DASC66011.2025.11257372](https://doi.org/10.1109/DASC66011.2025.11257372).
- [5] International Civil Aviation Organization. International Standards and Recommended Practices: Annex 11 - Air Traffic Services, 14th ed., 2016.
- [6] SESAR3 Joint Undertaking. SESAR IFAV3 Regulation (REG) Deliverable D2.13 Edition 01.00, 16 January 2026.
- [7] IATA. Air traffic control delays in europe: An overview of the economic consequences, February 2026.
- [8] CANSO. How to manage more traffic in less airspace. <https://canso.org/how-to-manage-more-traffic-in-less-airspace/>, 2024.
- [9] SESAR3 Joint Undertaking. SESAR IFAV3 Safety, Performance Requirements, Interoperability and Operational Service and Environment Definition (OSD) Deliverable D2.14 Edition 01.03, 07 July 2026.
- [10] CANSO. Seamless skies awards: announcing this year's winners and runners-up. <https://airspaceworld.com/blog/news/seamless-skies-awards-announcing-this-years-winners-and-runners-up/>, 2026.
- [11] International Civil Aviation Organization. Manual on Air Traffic Controller Competency-based Training and Assessment, 1st ed., 2017.
- [12] Bruno Rabiller and Octavian Nicolas Fota. Expanded Safety Reference Material (E-SRM) Core Document. Methodology, Eurocontrol, Brussels, Belgium, November 2023.
- [13] SESAR3 Joint Undertaking. SESAR IFAV3 Safety Assessment Report (SAR) Deliverable D2.14 Part II Edition 01.00, 09 March 2026.
- [14] Jana Meier, Marco-Michael Temme, Robert Hunger, Shruthi Shetty, Maria Hagl, Oliver Ohneiser, and Michael Finke. Validation of traditional competency-based assessment and experience-based authorisation for air traffic controllers. In 35th Congress of the International Council of the Aeronautical Sciences, ICAS 2026, Sydney, Australia, 13-17 September 2026. ICAS. accepted.
- [15] Maria Hagl, Robert Hunger, Jana Meier, Shruthi Shetty, Michael Finke, Tim Ruediger, and Marco Temme. Competent, not necessarily confident: Insights from a simulator study on atco readiness on non-endorsed sectors. Human Systems Engineering and Design (IHSED2026): Future Trends and Applications, 2026. DOI: [10.54941/ahfe1008178](https://doi.org/10.54941/ahfe1008178).
- [16] Government Accountability Office. Technology readiness assessment guide. (GAO Publication No. GAO-20-48G). <https://www.gao.gov/assets/gao-20-48g.pdf>, 2020.