

TRAINING AIRCREW COMPETENCY USING EYE-TRACKING (TACET) A VIRTUAL REALITY CASE STUDY

→ CONTENTS

EXECUTIVE SUMMARY _____	3
INTRODUCTION	4
1.1 TACET OVERVIEW	4
1.2 TRAINING PHILOSOPHY	5
1.3 ELE PERFORMANCE ANALYTICS	5
1.4 COMPETENCY BASED TRAINING ASSESSMENT	6
METHOD _____	7
2.1 CASE STUDY CONTEXT	8
2.2 SCENARIOS	8
2.3 TRAINING SETUP	9
RESULTS _____	10
3.1 ELE ANALYSIS	11
3.2 CBTA ANALYSIS	12
3.3 FEEDBACK	13
3.3.1 STRENGTHS	13
3.3.2 LIMITATIONS	14
TESTIMONIALS	14
FUTURE CONSIDERATIONS _____	15
4.1 ADDRESSING FIDELITY LIMITATIONS	16
4.2 BOLSTERING COGNITIVE ABILITIES	17
4.3 POTENTIAL TECHNOLOGICAL ADVANCEMENTS	17
4.3.1 INTEGRATION OF NATURAL LANGUAGE PROCESSING (NLP)	17
4.3.2 REAL-TIME CUSTOMISATION OF SCENARIOS	18
4.3.3 AIRCRAFT AGNOSTIC TRAINER	18
CONCLUSION _____	19
REFERENCES _____	20

 EXECUTIVE
SUMMARY

In this whitepaper, we explore the use case and benefits of using a unique Virtual Reality (VR) tool for pilot competency training and assessment, known as Training Aircrew Competencies using Eye Tracking (TACET). Key to TACET's effectiveness is its incorporation of AI-powered software analytics, which systematically captures and records critical performance and psychometric insights based on Eye-Tracking. This data serves as an objective measure to evaluate various competencies and levels of expertise, enabling targeted development and assessment of proficiency. Over the past two years, TACET has been implemented in training sessions with an international airline to evaluate its potential efficacy as a competency training and assessment tool for pilots. The findings indicate promising results regarding the viability of TACET in enhancing pilot assessment and its integration into training programmes. Moreover, this paper discusses considerations for future development, both of TACET specifically and other immersive solutions, emphasising the potential for continued innovation and improvement in this area.



INTRODUCTION

Human factors encompass a spectrum of psychological, social, and organisational elements that influence individual and collective performance in complex systems. Approximately 75% of flying accidents arise from breakdowns in these 'soft' skills rather than from technical lapses alone (Kharoufah et al., 2018). Recognising and addressing these human factors, including communication, decision-making and situational awareness, becomes paramount in fortifying the overall proficiency of pilots through targeted training. Pilot training has long been situated at the crossroads of conventional classroom methods and state-of-the-art flight simulators. Traditional classroom methods, such as desktop and cockpit procedure trainers, have historically provided a cost-effective and scalable approach to disseminating foundational knowledge. Conversely, full flight simulators are extremely realistic, but they are highly expensive to acquire and operate, difficult to maintain, and require trainees to travel to a fixed location.

The aim of Training Aircrew Competencies using Eye Tracking (TACET) is to use innovative and immersive technology to bridge the gap between classroom-based and full flight simulation-based methods, and to provide a cost-effective yet immersive solution to bolster pilot competency training and assessment. Our training tool aims to augment rather than replace current methods, reduce the need for an instructor and provide highly detailed feedback for pilots. Our goal is to improve training accessibility and enhance elements of the training process, making it more efficient and better for improving pilot human factors.



1.1 TACET OVERVIEW

TACET is a unique software-based competency training and assessment tool for airline pilots, delivered in Virtual Reality (VR). The software takes a crew through several 'flows' and fault scenarios, interacting with each other and the virtual cockpit to overcome each event. The cockpit used in this version of TACET is a digital twin of the Airbus A320. At the core of this training is an intelligent eye-tracking based software tool, named Psych Adaptive Virtual Environments™ (ELE), that provides unique insight into trainee performance and various new opportunities for enhancing skill acquisition.

The ELE software incorporated in TACET could facilitate focused learning experiences, challenging pilots by targeting areas of learning that are commonly difficult, e.g. dealing with complex scenarios or interruptions. Data captured from the eye-tracking is autonomously

processed and analysed to understand how pilots deal with these challenging situations. Understanding this information is important in developing pilots that can adapt well to real-world issues. The TACET software is multiplayer enabled, meaning a 2 person crew is able to play through scenarios together. The faults and flows incorporated into the scenarios are open-ended in such a way that the procedure does not have to be linear, and the pilots are able to discuss all eventualities.



1.2 TRAINING PHILOSOPHY

Simulation training can be a viable alternative to real-world practice methods in safety-critical industries, and we posit that VR holds a distinctive position within pilot training paradigms. While full flight simulators are considered the gold standard, there exists a notable gap between these high-fidelity systems and conventional classroom training methods that are typically used to train non-technical skills or competencies. VR presents itself as an immersive, cost-effective, and efficient intermediary and supplement. The capacity of VR to emulate realistic flight scenarios (e.g., unpredictable events or distractions) renders it a valuable tool for cultivating human factors and 'soft' skills inherent in expert performance. Modern aircraft procedures utilise 'flows', a predetermined pattern of movement to confirm switch positions, and control settings for various stages of flight. Correct completion of these flows is vital to safe operation. Use of simulation can quickly ingrain flow patterns because the trainee is able to spatially orientate themselves and develop muscle memory effectively.



1.3 ELE PERFORMANCE ANALYTICS

The backend of TACET incorporates AI-powered software analytics that captures and records key performance and psychometric data. We have developed closed loop training exercise scenarios using Airbus A320 standard failures that are incorporated into the VR along with distractions and 'flows'. Over the years we have gathered eye tracked data at various expertise levels from which objective performance information can be gathered relating to decision making, motor accuracy, situational awareness etc. This information can then be used to objectively assess the training proficiency and/or expertise of pilots (Harris et al., 2022), as well as targeting areas for development. This information generated via the eye-tracking functionality, particularly the Feedback Eye Movement Training (FBEMT), will allow trainers to align trainee performance against existing operational frameworks for training competency. Current aviation competency assessment procedures require assessors to undertake objective measures of a trainees proficiency and their ability to acquire and/or manage a significant quantity of situational information (e.g., from visual cues and information display systems). Eye-tracking measurements are arguably the only tools that enable such assessments to be undertaken in a truly systematic and reliable manner, while they are also able to provide insight about a trainee's underlying psychological state (e.g., their stress levels, predictions, or upcoming plans/intentions).

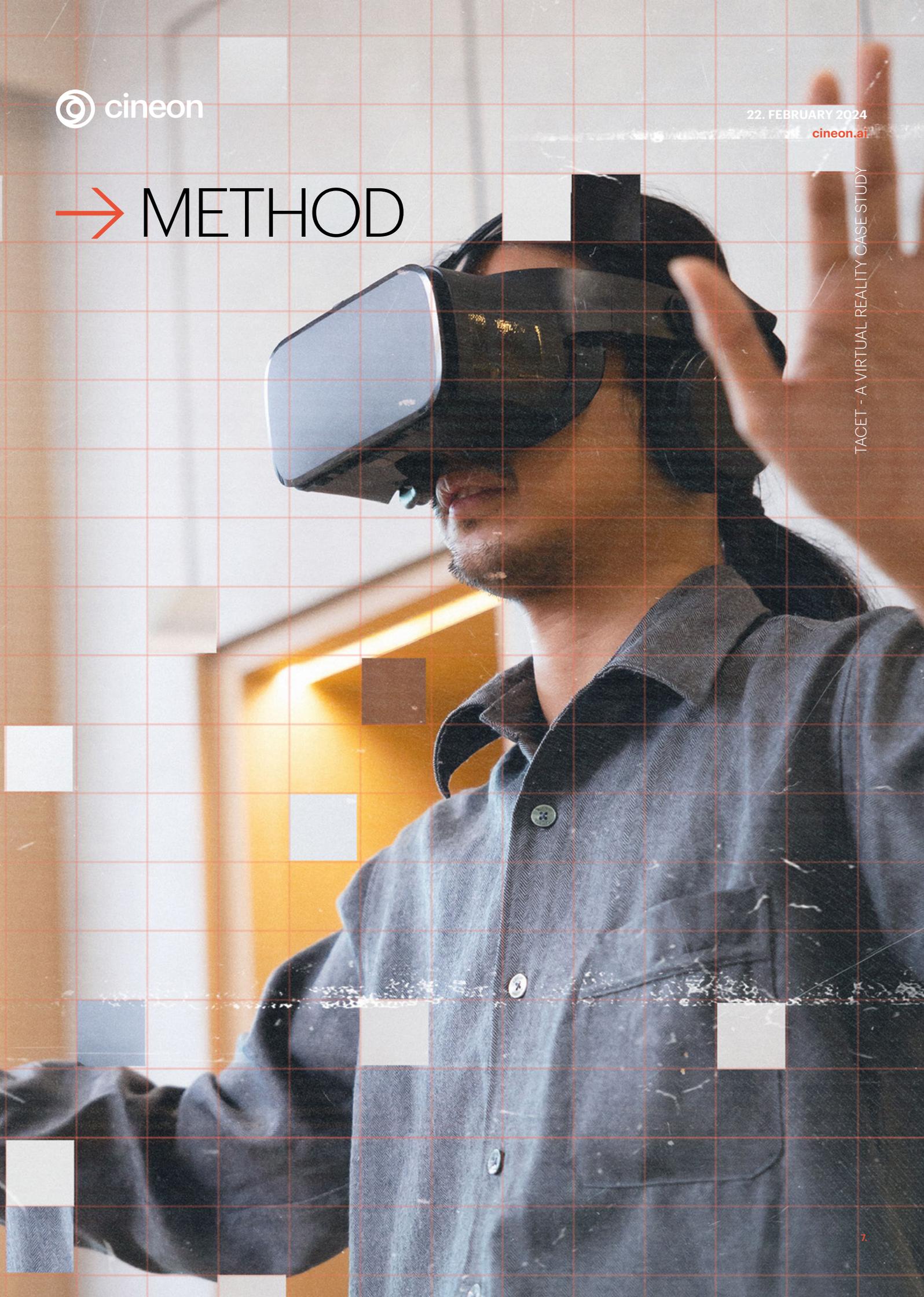


1.4 COMPETENCY BASED TRAINING ASSESSMENT

Competency-based training and assessment (CBTA) is an approach to training and assessment that the aviation industry uses to ensure pilots are qualified (see Civil Aviation Authority). It encompasses assessment techniques that are characterised by a measurement of the standards of performance displayed via specific competencies, and the development of training to the specified performance standards. A competency is manifested and analysed through observable behaviours (OB's) that underpin the relevant knowledge, skills and attitudes to carry out activities or tasks under specified conditions. The eye-tracking and performance assessment systems in TACET are able to pick up on several of these OB's that are related to certain competencies.

The specific competencies that are currently captured by TACET are Workload Management, Problem Solving and Decision Making and Situational Awareness. Usually, these OB's are subjectively captured within the sim by an instructor, but Eye Tracking offers a quantitative and precise way to assess when certain OB's occur. Any areas where a trainee may be lacking can then be picked up by an instructor and be the focus of subsequent training sessions or simulator time. Additionally, the system is able to record the conversation between the crew members, which an instructor is able to playback after the fact and assess proficiency for interpersonal competencies such as Leadership and Teamwork, and Communication.

→ METHOD



2.1 CASE STUDY CONTEXT

Our training partner for this case study was a leading private airline based at Farnborough International Airport. The airline boasts an Airbus A320 in its fleet, mirroring the cockpit configuration utilised in the TACET software. In February 2023, TACET was incorporated into a training day with the company pilots, culminating in an evaluation of the feedback gathered from the session. This input was instrumental in guiding a subsequent round of development for TACET. Notably, a pivotal enhancement emerged with the redevelopment of TACET to accommodate multiplayer functionality, facilitating a collaborative two-person crew experience during training scenarios. A second training day ensued in January 2024 to conclude the case study.

During the session, eight pilots, organised into four sets of two crews, engaged in two training and assessment scenarios. Each crew alternated between the roles of Pilot Monitoring (PM) and Pilot Flying (PF), undergoing a brief two-minute virtual cockpit familiarisation task before entering the scenarios. The duration of each scenario averaged approximately ten minutes per crew, during which ELE analytics were collected from the PM. Post-scenario, an instructor reviewed the generated feedback, conducting a debrief to disseminate key findings to the broader group.

Concluding the training session, a collaborative workshop was conducted with the pilots and instructor to gain insights into their experiences with the tool. Discussions delved into envisioning the tool's integration into a training curriculum and input on directions for further tool development. This engagement not only went towards validating the efficacy of TACET in practical training scenarios but also provided valuable perspectives for evolving this type of training with the needs and preferences of the industry.

2.2 SCENARIOS

Every scenario is designed with predefined events, encompassing faults and distractions, yet they primarily offer an open-ended structure to encourage exploration and discussion.

The scenarios chosen for the case study were the BlueRes Fault and FADEC Fault because they targeted OB's that were specifically relevant to the companies objectives of training for that week. Details of the events and procedures are outlined below.

TYPE	SCENARIO
FLOW	After landing flow
	After start flow
	Taxi flow
FAULT	Rejected takeoff
	Blue resovior hydraulics failure
	Engine 1 FADEC failure
	Air pack overheat

SCENARIO 1 is a Blue Reservoir Hydraulics fault that occurs at the same time as a cabin crew member alerts the pilots to a fire in the galley. The pilots must discuss the situation, prioritise and work through each event. Generally, pilots are required to follow ECAM actions and await further information from the cabin crew.

SCENARIO 2 is a possible FADEC Engine 1 fault; the pilots must confirm whether or not the engine has failed based on the visual information available. If so, they must follow several ECAM procedures. If not, then they proceed to clear ECAM actions and discuss.

2.3 TRAINING SETUP

The training was delivered on the HP Reverb Omnicept headset (see HP), which is tethered to a VR ready laptop or a PC. The use of Tethered VR enables better processing power of the environment, ultimately leading to better visual quality and frame rate, and allows for rapid computation of ELE to deliver timely feedback. Both trainee's are sitting next to each other physically and virtually, as they would in a real cockpit, and are able to communicate verbally as normal.



FIGURE ABOVE: HARDWARE SETUP.

→ RESULTS

Diverse approaches were observed among various crews when addressing the scenarios. For instance, in the case of the BlueRes Fault, one crew opted to prioritise addressing the fire and initiating a diversion, while alternative crews chose to focus on resolving the ECAM fault and wait for further updates from the cabin staff. During the FADEC scenario, the ECAM screen was frozen and unable to be cleared due to an error in the VR procedure. Whilst this was technically a fault with the VR, it led to lengthy discussions about what the pilots would do if this happened in real-life. The occurrence of this ‘black swan’ event posed a significant challenge, catching most individuals off guard and requiring a series of complex problem-solving discussions and operations to be undertaken.

3.1 ELE ANALYSIS

Figures 1 and 2 showcase the heat maps generated from the PM in each scenario, offering an insightful representation of their gaze patterns. The system is able to visualise these areas of focus and insights into strategy, which may then be linked to levels of situational awareness and expertise. The interplay between gaze patterns and eye behaviors becomes particularly apparent in these visual representations, with specific gaze strategies associated with divergent decision-making strategies and cross-checking behaviours.

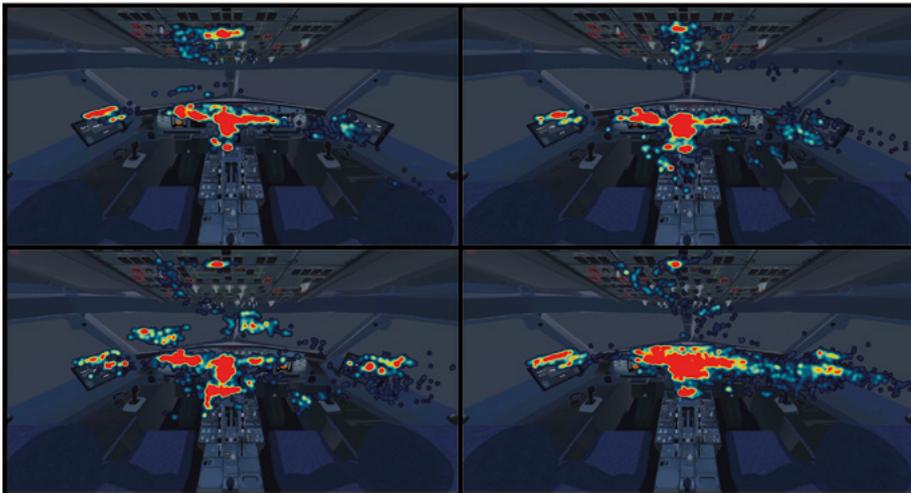


FIGURE 1: PM GAZE HEAT MAPS FROM BLUERESFAULT (CREW 1-4).

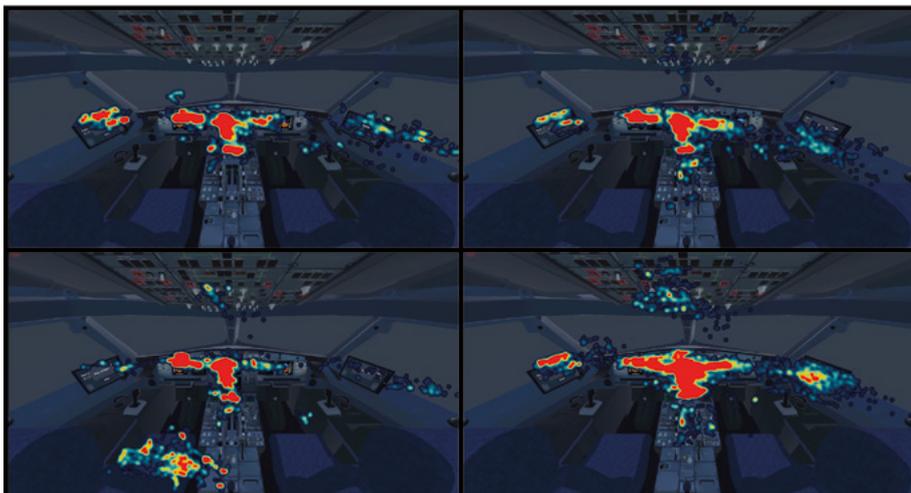


FIGURE 2: PM GAZE HEAT MAPS FROM FADECFAULT (CREW 1-4).

Presented below is a screenshot extracted from the ELE analysis dashboard, where particular eye movements are correlated with varying levels of expertise. This tool empowers instructors to navigate along the time series of the scenario, identifying spikes that indicate instances of heightened psychological responses (e.g., stress, cognitive workload, fatigue) during the assessment. Furthermore, instructors can gauge the correlation between performance and expertise levels by looking at specific eye movement patterns. Notably, a qualitative trend was observed among multiple crews, indicating elevated mental load and stress responses following incidents such as the galley fire cue and the frozen ECAM screen in each scenario.

Overall, this case study stands as testament to the feasibility of using ELE as a feedback tool within competency-based training. We were able to establish that the system is able to detect meaningful changes in user performance states and link them back to key competency-related OB's. On this basis, we are now ready to explore whether this data can be fed-back in real-time to facilitate adaptive training, where we seek to optimise learning experiences according to psychological metrics such as cognitive load, stress and distractibility.

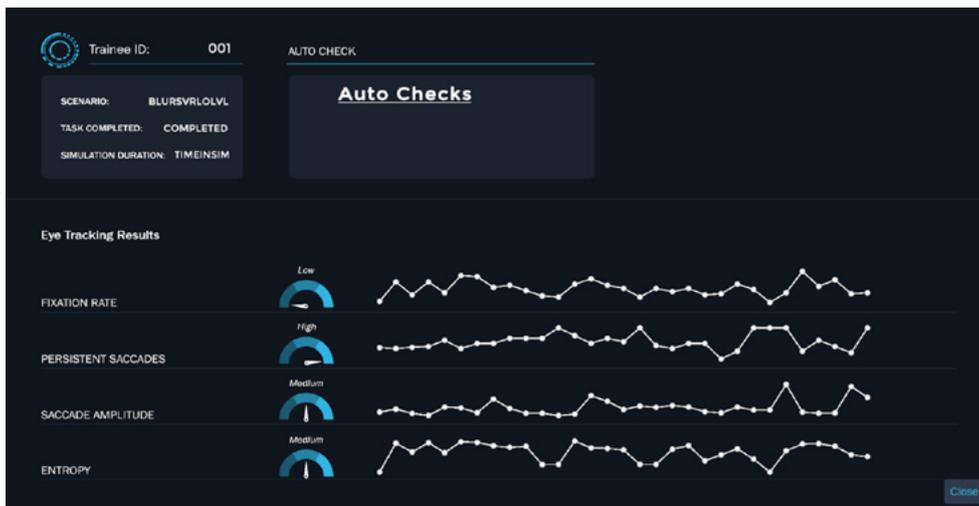


FIGURE 3: THE ELE ANALYSIS DASHBOARD; CREW 1, BLUERESFAULT.

3.2 CBTA ANALYSIS

The prototype of the CBTA system worked well as a proof of concept. Several OB's were captured for each scenario, presented as an overview on the dashboard for instructor evaluation. This system was able to flag potential oversights in crucial information and also provided a high-level assessment of competency scores across various dimensions. This capability could facilitate a targeted approach to address specific weaknesses or learning goals, thereby offering opportunities for tailoring future training. Further development is needed to make the auto-assessment more flexible to allow for extensive exploration during scenario's and facilitate more detailed feedback in the dashboard. This future work could also incorporate ELE-based data metrics into the automated scoring algorithms, to provide a more comprehensive and holistic 'picture' of competency-based training performance.

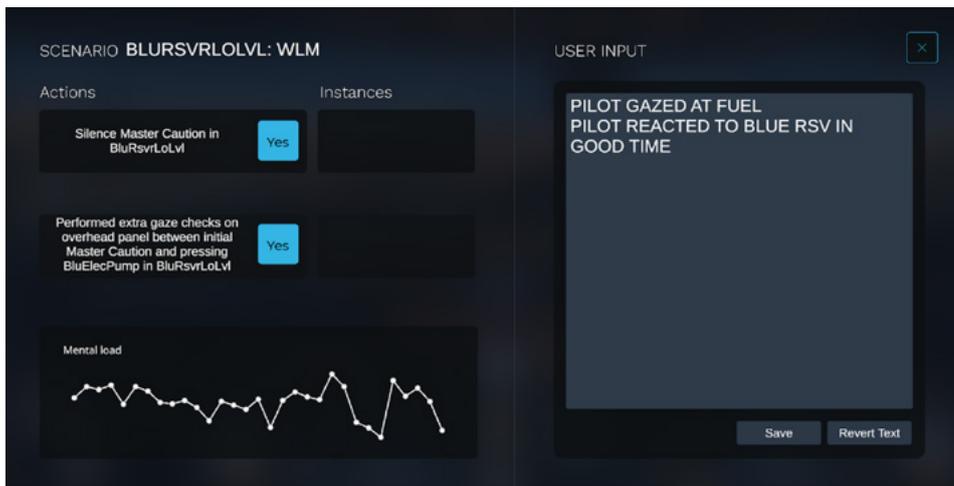


FIGURE 4: THE CBTA DASHBOARD FOR WORKLOAD MANAGEMENT; CREW 1, BLUERESFAULT.

3.3 FEEDBACK

Feedback from the participants has been collated and discussed below. Generally, the introduction of multiplayer was seen as a significant positive, promoting collaboration and in-depth communication. The ongoing challenge lies in addressing fidelity concerns, and ensuring the seamless integration of the CBTA via auto-assessment of the competencies. Participants acknowledged that achieving 100% fidelity of the VR might be a challenge. However, while not replicating completely realistic visual elements, the VR platform was acknowledged to provide the advantage of extended practice time.

3.3.1 STRENGTHS

Notable advancements have been made to the tool when compared to previous training iterations. In particular, the introduction of a second virtual crew member appeared to enhance the training experience, fostering collaborative learning and enhanced communication opportunities. It was noted that TACET, as a training platform, facilitates the utilisation of the competency assessment as an anchoring mechanism, providing a framework for individual skill development. The instructor and trainees expressed appreciation for TACET's capacity to offer enhanced insights, especially in the playback of their gaze behaviour and visual scan paths (which enabled the identification of potential oversight areas during collaborative monitoring activities).

One salient feature highlighted by participants was TACET's capability to induce training scenarios that are non-specific to technical skills and standard operating procedures. Instead, the VR-based approach introduces diverse and unexpected environmental complexities, which uniquely target key competency-related behaviours (e.g., problem-solving skills and/or the management of high-stress workloads). Moreover, the transfer of knowledge acquired in the virtual realm to the physical simulator, coupled with the in-depth feedback provided, could be a valuable tool for the optimisation of simulator time in the future. Prospective work should explore the integration of learning outcomes between VR and conventional simulator systems, to realise the potential that TACET offers for improving training efficiency and cost-effectiveness.

3.3.2 LIMITATIONS

The fidelity of the simulator was identified as a current weakness, with issues relating to visual information (e.g. system pages/dials not being completely accurate to the scenarios) and tactile feedback (e.g., a lack of vibration sensations in the flightdeck) affecting elements of realism in the pilot responses. The inability to observe body language in the VR environment was also noted as a drawback, with the focus being only on verbal cues. It is clear that enhancements in system fidelity should continue to be explored in future iterations and development of the tool.

TESTIMONIALS

"In the scanner you don't have much time for each procedure as there are mandatory things to tick off, but with the VR you have time to explore all eventualities."

"Not being able to clear the ECAM took you out of a normal scenario and you had no idea what to do. You had to think outside the box and work together without procedure and guidance."

"I would find it really useful to chuck on the headset with another crew member to use as a prequel to sim time."

"It's more immersive than the flight sim with more realistic sounds and visuals of flying, even without the feel of movement."



→ FUTURE
CONSIDERATIONS

A crucial aspect of future considerations involves the seamless integration of TACET with other pivotal components of pilot training, namely classroom instruction and flight simulators. Classroom instruction can lay the theoretical groundwork, while VR simulations provide an immersive and practical extension of the learning experience. An essential component of this involves the automation of the training delivery process, aiming to reduce the dependency on personnel and resources. Automating training scenarios and assessments within VR environments can offer a standardised yet adaptive approach to pilot instruction. Intelligent algorithms like those in TACET, can dynamically adjust difficulty levels based on individual progress, providing personalised and efficient training experiences. Additionally, the automation of administrative tasks, such as performance tracking and feedback generation, is able to significantly alleviate the workload on instructors, allowing them to focus on more nuanced aspects of training.

The existing tool has successfully incorporated the multiplayer functionality, accommodating up to two crew members. However, a notable avenue for future enhancement involves expanding this feature to facilitate the engagement of the entire aircraft crew in collaborative scenario simulations. This evolution would not only mirror real-world dynamics but also foster a more comprehensive training environment, enabling various crew members to actively contribute and coordinate in response to simulated challenges. The potential for such an extended multiplayer capacity holds substantial promise for refining team dynamics and overall crew synergy during training sessions.

Furthermore, there exists a necessary requirement for refining the visual aspects of the simulation, aiming to elevate the overall effectiveness of training through enhanced fidelity. Augmenting the quality of critical indicators, such as dials and fuel gauges, through real-time updates could significantly heighten the sense of immersion and/or jeopardy perceived by trainees. This heightened realism would be poised to elicit more authentic and nuanced responses, as the simulated scenarios would closely mirror the complexities and challenges encountered in actual flight situations.



4.1 ADDRESSING FIDELITY LIMITATIONS

A persistent challenge impeding widespread adoption pertains to the perception that VR lacks the fidelity seen in full flight simulators. While it is acknowledged that VR fidelity may not ever attain parity with traditional simulators, it is imperative to underscore that VR's inherent strengths lie beyond the realm of physical fidelity. Notably, VR training excels in facilitating cost-effective repetitions of procedures, enabling the safe simulation of high-risk events, delivering more precise feedback through enhanced data analytics, and adeptly replicating psychologically relevant task features (e.g., adaptive scenarios and unpredictable situational events). These attributes could collectively contribute to optimising simulator time, by strategically and flexibly targeting a pilot's proficiency gaps. Indeed, VR permits extensive exploration of all eventualities of an event, enabling the collaborative practice of procedures prior to costly and standardised simulator assessments.

Emphasising the notion of viability for this kind of training could be a pivotal concept for addressing these fidelity concerns. Viability refers to the practicality and effectiveness of implementing VR training tools in real-world scenarios, assessing whether the technology is able to deliver meaningful training outcomes. Effectively managing the perceptions

regarding fidelity is paramount to prevent immediate rejection of these innovative training approaches by end-users. It is imperative to recognise that physical fidelity, while important, is not the sole determinant of VR's efficacy; rather, it is the amalgamation of its holistic capabilities that render it a potentially valuable tool for prospective pilot training and assessment.



4.2 BOLSTERING COGNITIVE ABILITIES

One of the unique strengths of TACET as a training tool lies in its ability to simulate unexpected interruptions that are not necessarily technically-focused or procedural in nature. These types of scenarios, which can include the often labelled 'Black Swan' events, are less frequently focused on during traditional training programmes and cannot typically be resolved using explicit 'textbook instructions' or standard operating procedures. Incorporating these events into training scenarios can force trainee's to 'think outside the box' and rely on non-technical competency-based skills (e.g., problem-solving, leadership and teamwork). It is argued that a movement towards strengthening these 'soft skill' competencies and socio-cognitive abilities could lead to improved pilot performances and reduced likelihood of accidents in the future (Dapica & Peinado, 2021).



4.3 POTENTIAL TECHNOLOGICAL ADVANCEMENTS

As TACET evolves, several key technological advancements are being explored to enhance training effectiveness. These are outlined in the following sub-sections.



4.3.1 INTEGRATION OF NATURAL LANGUAGE PROCESSING (NLP)

An avenue for advancing the capabilities of TACET (and its underlying ELE software) involves the incorporation of NLP. By incorporating NLP techniques, the training system could firstly gain the capacity to autonomously evaluate competencies related to communication, leadership and teamwork, further reducing the reliance on human instructors and broadening its potential learning applications. Additionally, the incorporation of NLP-driven cabin crew audio and air traffic control cues would facilitate task interactions that operate akin to conversational chatbots. A combination of these two approaches would offer exciting opportunities for competency-based training and assessments of the future. Personalised audio cues and task events could dynamically unfold based on a trainee's dynamic VR interactions and psychometric metrics (e.g., measures of stress and/or cognitive load), enabling autonomous adaptation of the learning environment based on specific programme goals or objectives.



4.3.2 REAL-TIME CUSTOMISATION OF SCENARIOS

Potential next steps for TACET could involve providing trainers and pilots with the ability to tailor scenarios according to specific training or assessment needs. This could materialise through the integration of a dedicated user interface that is used to build the scenario and customise events in real time. This feature could empower trainers to dynamically introduce elements that challenge pilots or target certain observable behaviours, thereby enhancing adaptability and decision-making skills in a controlled, immersive environment.



4.3.3 AIRCRAFT AGNOSTIC TRAINER

Proposing the concept of an aircraft-agnostic trainer marks another potential training advancement. This entails the development of a training environment featuring either a generic cockpit or a team-based setting that is unrelated to aviation. The underlying principle is rooted in the objective of developing competencies that transcend specific aircraft models, fostering generalisability across a spectrum of tasks and skills. By providing a versatile training platform, this concept seeks to build upon core competencies applicable to diverse scenarios, promoting a more holistic approach to pilot training and assessment.

 CONCLUSION

In summary, this case study supported the viability and potential value of TACET as a prospective method of training and assessment within the aviation sector. The software's distinct strengths become evident in its ability to target human factors skills and assess key pilot training competencies, particularly through the simulation of unpredictable and personalised scenario events. A notable use case emerged, highlighting the software's potential for providing actionable feedback on proficiencies and weaknesses before full flight simulator sessions. The results underscore the effectiveness of TACET in eliciting realistic behavioural responses, and affirm the potential of ELE to adapt competency-related training demands and user experiences. Additionally, the proof of concept for the autonomous CBTA system showcases its viability as a valuable assessment tool in the future.

Despite its current achievements, there is acknowledgment of the need for further refinement and development. ELE, in particular, requires enhancements to physically adapt the environment based on competency performance levels and mental load. The CBTA system should undergo further development to allow for more flexible and nuanced assessments based on trainee exploration and discussion. Fidelity improvements are deemed necessary, and the management of user perceptions and/or blended learning methods could be crucial in this regard. A shift towards viability as a measure of training efficacy emphasises the importance of the software's practical utility.

With respect to future advancements, considerations include increasing the number of crew members, either through enhanced multiplayer integration or NLP capabilities. Furthermore, there is a recognised need for real-time customisation of the software by instructors to enhance adaptability and relevance in training scenarios. In essence, while TACET has achieved certain successes in this case study, the outlined areas for further work underscore the unique innovation opportunities that the tool presents for future pilot training and assessment.

 REFERENCES

Dapica, R., & Peinado, F. (2021). Towards a semantic knowledge base for competency-based training of airline pilots. *Procedia Computer Science*, 192, 1208-1217.

Harris, D., Arthur, T., de Burgh, T., Duxbury, M., Lockett-Kirk, R., McBarnett, W., & Vine, S. J. (2022, January 10). Assessing expertise using eye tracking in a Virtual Reality flight simulation. <https://doi.org/10.31234/osf.io/ue58a>

Kharoufah, H., Murray, J., Baxter, G., and Wild, G. (2018). 'A review of human factors causations in commercial air transport accidents and incidents: From to 2000-2016'. *Progress in Aerospace Sciences*, 99, 1-13.