Entertainment Computing 43 (2022) 100510

Contents lists available at ScienceDirect

Entertainment Computing

journal homepage: www.elsevier.com/locate/entcom

Who trains the trainers? Gamification of flight instructor learning in evidence-based training scenarios

Rubén Dapica [a], Alfredo Hernández [b], Federico Peinado [b,*]

[a] European University of Madrid, Department of Mechanical and Aerospace Engineering, Tajo s/n, Urb. El Bosque, 28670 Villaviciosa de Odón, Madrid, Spain
[b] Complutense University of Madrid, Department of Software Engineering and Artificial Intelligence, Prof. José García Santesmases, 9, 28040 Madrid, Spain

ARTICLE INFO

Keywords:
Aerospace Industry
Flight Simulator
Gamification
e-Learning
Competency-Based Training

ABSTRACT

Despite all the technological improvements currently available in aeronautical training, most accidents still have non-technical skills as contributing factors. In recent years, the International Civil Aviation Organization has been promoting a new training paradigm for crew performance that integrates all kinds of relevant skills, both technical and non-technical. The content of each training session is adapted to the specific improvement needs of each pilot, based on the data obtained in both real operations and simulations, instead of using stereotyped training scripts. This paradigm is called Evidence-Based Training. The assessment of non-technical pilot performance in these changing scenarios is complex, being highly dependent on the experience of the flight instructor. In this paper, we present a Flight Instructor Simulator that aims to gamify the learning process of instructors who have to work under this new training paradigm. A proof of concept of such an application has been created and tested by aviation personnel directly involved in this type of training. The results show that the use of gamification for flight instructors training in the assessment of pilot's non-technical competencies is a promising line of work. Furthermore, some recommendations are extracted from this study that can serve as design guidelines for similar projects.

1. Introduction

The role of flight instructors in simulations is to set up and monitor training sessions oriented to assessing and improving pilot performance. Traditionally, these training sessions focused on mastery in applying maneuvers and procedures through repeated exposure to known emergencies. This type of training is currently being superseded by a new concept in competency-based learning called Evidence-Based Training (EBT). This approach is promoted by the most important aeronautical institutions, such as the International Civil Aviation Organization[1] (ICAO) [1] and the International Air Transport Association[2] (IATA) [2,3]. It is essentially a competency-based approach. Unlike traditional training, this new approach aims to expose pilots to unforeseen scenarios so that they can develop skills that allow them to face situations for which they is not specifically trained. This learning paradigm has received enormous interest in many fields in recent decades, although applying it effectively has proven to be very complex, especially with regard to the role of the trainer. Hence, the training of these professionals constitutes a challenge [4].

In another vein, gamification, i.e. the use of typical elements from games in learning environments, has proven to be a strategy that can contribute to improved learning, in aspects such as engagement and motivation [5]. However, the use of gamification in the simulation of flight instructors' work is currently an unexploited field. The purpose of this work is to explore the viability of a gamified simulator for flight instructors based on EBT scenarios as well as to identify possible strengths and challenges of such an approach.

The rest of the paper is structured as follows. Section 2 reviews the types of aeronautical simulators and previous work on gamification in the field, providing a justification for the approach followed in this paper. Section 3 presents a model of the work of a flight instructor, defined within an EBT program and following the guidelines set by civil aviation authorities. The general criteria that should guide instructors for the assessment of pilots according to this approach are detailed here.

* Corresponding author.
E-mail addresses: ruben.dapica@universidadeuropea.es (R. Dapica), ahburgos@ucm.es (A. Hernández), email@federicopeinado.com, fpeinado@ucm.es (F. Peinado).

[1] https://www.icao.int
[2] https://www.iata.org

https://doi.org/10.1016/j.entcom.2022.100510
Received 31 January 2021; Received in revised form 27 April 2022; Accepted 23 June 2022
Available online 27 June 2022

In Section 4 the methodology used in this paper is briefly discussed. Section 5 describes the rationale and process followed to design the proof of concept in terms of choice of game elements. The foundations for the design of simulator gamification as well as the research questions that are intended to be elucidated are established in this section. Section 6 gives a description of the proof of concept and Section 7 details the steps followed and instruments used in order to expose this approach to the scrutiny of the experts. In Section 8, the information obtained from the experts is exhibited and analyzed in order to validate the proof of concept and to establish some design principles applicable to this type of gamification. Finally, in Section 9, we conclude that a system like the one presented is promising, and future guidelines for this line of research are discussed.

2. Related Work

Simulators are virtual learning environments that allow for the acquisition, application and practice of the knowledge and skills necessary for satisfactory job performance. The use of this technology allows the learning curve to be smoothed while offering a safe environment and reduced costs at the same time [6].

In the aerospace field there is an longstanding tradition of the use of simulators for training. The first simulators date back to 1910 [7]. One of these early trainers, the so-called *apprentice barrel*, consisted of a barrel split in half, resting on a flat surface on which a seat was mounted. The instructors manually moved the barrel to represent the pitch and roll of the plane. Today, we could hardly associate this artifact with a modern flight simulator, due to the highly technological operation of current systems.

It was not until the 1940s that flight simulators based on computers able to solve flight mechanics equations appeared. Along these lines, in 1948 *Curtiss-Wright* developed the first flight simulator that was used by an airline [8]. From here, different elements were added to improve the experience until the first modern simulators for commercial aviation were developed in 1954, incorporating motion [9].

Flight simulators would continue to evolve, reaching our days, where they have real-time high quality 3D graphics and the ability to recreate aircraft and environmental features with great fidelity. Today simulators can be differentiated based on their purpose:

- **Professional use**. The objective is the training of pilots, allowing them to acquire and improve both knowledge and skills. This type of simulator must be approved by the organizations responsible, such as the Federal Aviation Administration (FAA)[3], the European Union Aviation Safety Agency (EASA)[4] or the civil aviation authorities in the different countries.
- **Recreational use**. The objective is the entertainment of the user and although it can prompt him/her to acquire some knowledge, in no case is the motivation training the user to become a real pilot. Some games, like the case of X-Plane, may be approved for professional use, in combination with specific hardware, by organizations such as the FAA. Nowadays, the possibilities offered by online mapping services and cloud computing have raised the visual fidelity of this software to the highest level. The most recent example is *Microsoft Flight Simulator 2020* [10].

Nevertheless, the suitability of a flight simulator training session is not limited to the ability to represent environmental stimuli (visual, auditory, kinesthetic and tactile) accurately. First, correct environment representation does not guarantee cognitive fidelity. With current technology, practically any situation that may arise in flight can be simulated, but the way in which events occur must conform to consistency criteria. Otherwise, the proposed training scenario would not be credible and could spoil the utility of the session, regardless of the simulator's technical sophistication [11]. Second, simulated events must be adapted in level and nature to each pilot's strengths and weaknesses, so as to stimulate the specific skills that must be developed.

The approach based on exposing the student to credible situations suitable for assessing and improving their abilities is called Scenario-Based Training (SBT). This approach focuses on the cognitive aspect of training, making use of the possibilities offered by simulation technology. Although the case at hand is that of flight simulators, it is noteworthy that such an approach has also been used successfully in the medical [12] and military fields [13].

However, no matter how well a scenario is chosen and simulated, the ultimate goal of the session is for learners to be aware of their own mistakes and weaknesses in order to improve. For this task, the role of assessment and feedback provided by the instructor is essential. The most complex part of this work is the assessment of non-technical skills (NTSs). These competencies have a huge impact on safety. Failures in these types of skills such as teamwork and communication have contributed to numerous accidents [14–16]. Although training in non-technical aviation competencies has a long history, there is a lack of systematic approaches in evaluating said skills [17].

Although the use of simulators for pilot training has a very extensive history, there are no precedents for similar tools for the training and assessment of instructors. While there have been precedents of computer-aided tools for scenario design, their suitability as well as the skill in their application were always evaluated with real pilots [18]. Even today, instructors' performance on the simulator is directly monitored by another human evaluator. In this way, the instructor is assessed while performing his/her work in a session of the physical simulator with real pilots. The need for computer tools that provide support in this regard has already been identified in the specialized literature [19].

Facilitation of training and assessment of trainers using tools designed for this purpose may be even more justified in the case of adaptation to the EBT framework. Adapting to this new teaching–learning framework means that instructors must become deeply familiar with the competencies and Observable Behaviors (OBs) on which it is based [20]. Furthermore, the EBT approach involves the integration of NTSs in the evaluation of pilot performance in the SBT. As has been consistently found, experiential learning methods are the most effective for training these types of skills [21]. Simulating the work of a flight instructor assessing pilots facing an EBT training session could allow the instructor to become familiar with this training philosophy, while avoiding the costs associated with acquiring such experience in training sessions with real pilots.

For all these reasons, the creation of simulation tools for flight instructor training is currently a very interesting line of research. However, the mere use of simulation does not guarantee the learning of a discipline. Indeed, if the simulation is not engaging, there will be a loss of interest and therefore an abandonment of the educational tool [22].

To avoid this problem, one of the strategies that has aroused great interest in recent years is gamification. Although simulators tend to fall into the category of serious games [23], rather than creating a new game, our approach will be introducing game mechanics into an environment that was not originally designed as such. Therefore, this work aims more at gamification, the use of game design elements in non-game contexts [24], in this case a learning application.

The use of gamified simulators can be an adequate solution to the need for experiential learning in a wide range of educational domains [25–27]. In the aerospace sector, as research works along the lines of this paper we can mention the following. Very recently, Badea [28] applied game elements to the work environment of air traffic controllers. The application consisted of representing fictitious airplanes that were superimposed on real traffic, generating conflicts that promoted the need to maintain attention on the supervision of automatic traffic

[3] https://www.faa.gov
[4] https://www.easa.europa.eu

control systems. Peng et al. [29] developed a gamified simulation of the Apollo missions. The application was focused on increasing motivation and promoting STEM learning. More closely related to professional training, Cornelissen *et al.* [30] evaluated an application of gamification for astronaut training. Specifically, they introduced a series of game elements aimed at increasing motivation and promoting self-study for the preparation of manned space missions, with positive results. In the field of military pilot training, Noh [31] explored the potential benefits of gamification in an existing military flight simulator in terms of motivation and usability, based on the judgment of a set of Korean Air Force pilots. Finally, sticking to the scope of airline pilot training, although some possible advantages of gamification have been identified [32], at the time of writing, this paper's authors have not found any relevant research work following the approach proposed here.

This work aims to design and evaluate a proof of concept for a gamified simulator aimed at improving the motivation and engagement of flight training instructors who are becoming familiar with the EBT framework.

3. Modeling Flight Instruction in EBT

Pilot training programs in airlines have traditionally been focused on the repeated execution of maneuvers and procedures in the face of known events that produced accidents in the past. However, this paradigm has proven to be insufficient. In the first place, the number and variety of possible events that could lead to an aviation accident is incalculable, which makes it impossible to train each situation in order to automate a specific behavior for each case [1]. Second, accident analysis has consistently shown that between 70 and 80% of air accidents have human error as a contributing factor. These accidents render poor performance in cognitive processes such as decision-making, communication, leadership or workload management as common patterns. All these characteristics imply NTSs, the assimilation of which cannot be ensured by the mere repetition of procedures [2].

The role of the instructor in EBT programs can be summarized in the following areas of work [20]:

- **Evaluation.** Assessing pilots' competencies.
- **Scenario design.** Preparing training sessions that expose pilots to situations that give them the opportunity to work on the specific skills to be developed.
- **Instruction.** Conducting and managing training sessions in such a way as to optimize the achievement of learning objectives.
- **Feedback and coaching.** Interacting with trainees so that they can recognize their strengths and weaknesses when coping with problems and how they could be handled more efficiently.

All these instructor functions are present in an SBT training session. Pilots are exposed to flight scenarios that contain unexpected events and situations. The scenarios are composed of operational situations that constitute a challenge for the pilots, and which will require the application of knowledge, skills and attitudes beyond the mere application of standard procedures. How pilots manage threats and errors in real time is the behavioral basis for determining whether the necessary competencies have been acquired and used correctly.

In this study we will only focus on the assessment of pilots' non-technical competencies. Thus, the instructor will be given some virtual pilots to train who have a specific competency profile. Based on this information, a scenario will be generated and displayed by the tool. The instructor will have to observe the pilots' behavior in said scenario and assess their performance in relation to the acquisition of the competencies detailed in Section 3.2.

3.1. Scenario Selection

SBT is an approach that is rooted in Situated Learning Theory, according to which meaningful learning cannot be separated from experience in concrete contexts, and is therefore strongly linked to the activity where it is used [33]. A scenario in flight simulation is understood as equivalent to narrative scenarios in video games; that is, the set of elements such as context and events that make up the story in which participants are immersed while playing the game [34].

In the case of flight simulation scenarios, elements can be classified into:

- **Conditions.** Static elements such as the operational environment (flight plan route, aircraft load conditions, air traffic, flight phase, etc.) or the aircraft model.
- **Events.** Occurrences that appear during the flight, such as malfunctions, meteorological phenomena, instructions from air traffic control, etc.

The concept of simulator that we propose in this work allows the instructor to select some parameters for the scenario that will be represented during the session, and the program will configure the specific conditions and events in the scenario based on the parameters provided (see Section 6.1).

3.2. Pilot Competencies and Observable Behaviors

The pilot competencies identified as relevant by ICAO are defined in their guide [1]. They are listed and defined below:

- **Application Procedures (APK).** Identifies and applies procedures in accordance with published operating instructions and applicable regulations using directed knowledge.
- **Communication (COM).** Demonstrates effective oral, non-verbal, and written communications under normal and non-normal conditions.
- **Aircraft Flight Path Management, Automation (FPA).** Controls the aircraft's flight path through automation, including the use of flight management systems and guidance.
- **Aircraft Flight Path Management, Manual Control (FPM).** Controls the aircraft's flight path through the flight manual, including the use of flight management systems and systems' flight guide.
- **Leadership and Teamwork (LTW).** Demonstrates effective leadership and teamwork.
- **Problem Solving and Decision Making (PSD).** Accurately identifies risks and solves problems. Uses the right decision-making processes.
- **Situation Awareness (SAW).** Perceives and understands all relevant information available and anticipates what could happen that may affect the operation.
- **Workload Management (WLM).** Manages available resources efficiently to prioritize and perform tasks in a timely manner in all circumstances.

The main goal of our simulator is for the user to observe the performance of virtual pilots while they face a new scenario and learn to correctly identify those OBs associated with the possession of the competencies to be developed. These OBs are defined and detailed in Table 1. For this work, purely technical competencies (APK, FPA and FPM) have been discarded.

3.3. Instructor Competencies and Observable Behaviours

Training in NTSs has a long tradition in aviation [35]. However, many of the soft skills training programs are flawed because of metrics issues. Although technical competencies have easy-to-measure parameters (airspeed, track accuracy, etc.), the same does not occur in non-technical competencies, and variability in assessment criteria can be considerable [2]. For this reason, standardization in instructor

Table 1
Pilot OBs for non-technical competencies. The correct assessment of pilot competencies during the simulation and therefore the user's (flight instructor's) performance is based on these OBs. The user must observe the pilots' behaviors during the simulation and detect whether they reflect the OBs associated with each competency positively or negatively.

Competency	Observable Behaviors
Communication (COM)	• Ensures the recipient is ready and able to receive the information • Appropriately selects what, when, how and with whom to communicate • Conveys messages clearly, accurately and concisely • Confirms that the recipient correctly understands important information • Listens actively and demonstrates understanding when receiving information • Asks relevant and effective questions • Adheres to standard radiotelephone phraseology and procedures • Accurately reads and interprets required company and flight documentation • Accurately reads, interprets, constructs and responds to datalink messages in English • Completes accurate reports as required by operating procedures • Correctly interprets non-verbal communication • Uses eye contact, body movement and gestures that are consistent with and support verbal messages
Leadership and Teamwork (LTW)	• Understands and agrees with the crew's roles and objectives • Creates an atmosphere of open communication and encourages team participation • Uses initiative and gives directions when required • Admits mistakes and takes responsibility • Anticipates and responds appropriately to other crew members' needs • Carries out instructions when directed • Communicates relevant concerns and intentions • Gives and receives feedback constructively • Confidently intervenes when important for safety • Demonstrates empathy and shows respect and tolerance for other people • Engages others in planning and allocates activities fairly and appropriately according to abilities • Addresses and resolves conflicts and disagreements in a constructive manner • Projects self-control in all situations
Problem Solving and Decision Making (PSD)	• Seeks accurate and adequate information from appropriate sources • Identifies and verifies what and why things have gone wrong • Employs proper problem-solving strategies • Perseveres in working through problems without reducing safety • Uses appropriate and timely decision-making processes • Sets priorities appropriately • Identifies and considers options effectively • Monitors, reviews, and adapts decisions as required • Identifies and manages risks effectively • Improvises when faced with unforeseeable circumstances to achieve the safest outcome
Situation Awareness (SAW)	• Accurately identifies and assesses the state of the aircraft and its systems • Accurately identifies and assesses the aircraft's vertical and lateral position, and its anticipated flight path • Accurately identifies and assesses the general environment as it may affect the operation • Keeps track of time and fuel • Maintains awareness of the people involved in or affected by the operation and their capacity to perform as expected • Accurately anticipates what could happen, plans and stays ahead of the situation • Develops effective contingency plans based upon potential threats • Identifies and manages threats to the safety of the aircraft and people • Recognizes and effectively responds to indications of reduced situational awareness
Workload Management (WLM)	• Maintains self-control in all situations • Plans, prioritizes and schedules tasks effectively • Manages time efficiently when carrying out tasks • Offers and accepts assistance, delegates when necessary and asks for help early • Reviews, monitors and cross-checks actions conscientiously • Verifies that tasks are completed to the expected outcome • Manages and recovers from interruptions, distractions, variations and failures effectively

evaluation criteria is so important in EBT programs.

In the same way that pilots must demonstrate the possession of a series of competencies by virtue of some OBs, instructors must be evaluated based on the possession of their own competencies. The EBT instructor and evaluator competencies (IEC) are also defined by the authorities [20], and listed below:

- **Pilot competencies (IEC 1).** Demonstrates the possession of pilot skills (see Section 3.2).
- **Management of the learning environment (IEC 2).** Ensures that instruction, assessment and evaluation are conducted in a suitable and safe environment.
- **Instruction (IEC 3).** Conducts training to develop the trainee's competencies.
- **Interaction with trainees (IEC 4).** Supports trainee's learning and development. Demonstrates exemplary behavior (role model).
- **Assessment and evaluation (IEC 5).** Assesses the trainee's skills. Contributes to continuous training system improvement.

The flight instructor, who must first of all be an experienced pilot of the aircraft on which the simulation is being carried out, must possess the same skills to be assessed in the pilots, as well as know-how to identify those OBs that demonstrate their possession, hence the importance of the first instructor competency listed above (IEC 1). This skill is the most fundamental of all, since the entire training process, including feedback, interactions with the trainees and assessment, is based on the pilots developing their own competencies. Therefore, the third, fourth and fifth competencies are based on the first.

Consequently, in this first stage of designing an instructor training simulator, only this first and fundamental competency, IEC 1, has been taken into account.

Fig. 1 shows the conceptual model of an instructor's work during an SBT session, following the design of our simulator concept. There are three stages of action. In the first one, the flight instructor designs a training scenario for the virtual pilots. In the second, the instructor observes the pilots' behavior during the simulation, gives feedback if necessary and evaluates their skills. Finally, the system records all the instructor's actions and evaluates his/her skills as a flight instructor, providing a an interactive display of the results.

4. Methodology

The process followed in this paper is defined within a user-centered

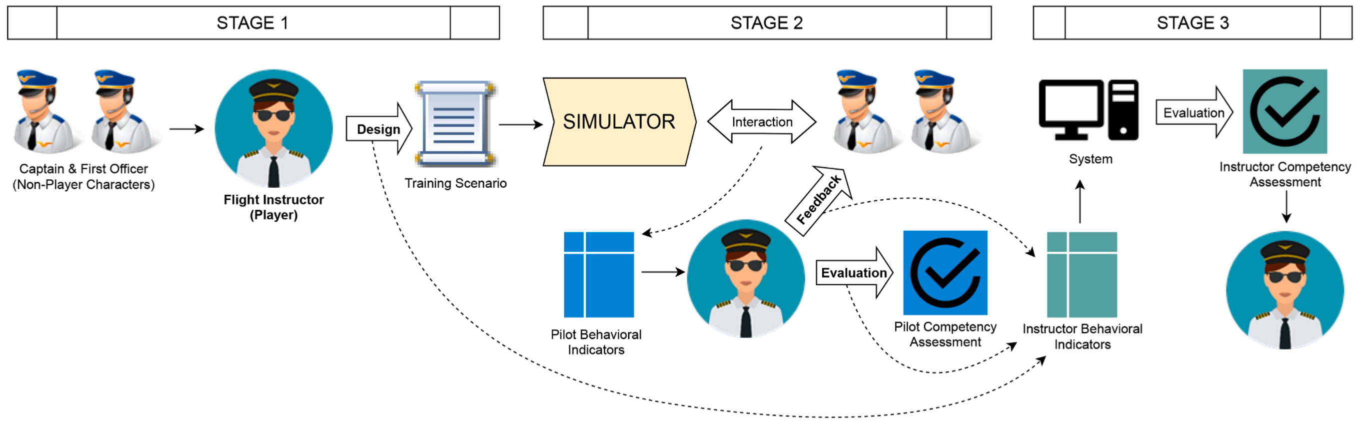

Fig. 1. Model of the work of a flight instructor according to our concept. The simulator is structured in three stages: session setup, simulation and assessment.

methodology applied to gamification, based on 4 phases, adapted from Mora et al. [36]:

- **Declaration**. In which the problem is defined, as well as the actions to be taken and the results expected. Within this phase, a delimitation of the gamification concept to be used, the setting of goals and the definition of research questions were carried out. These actions are described in Section 5.
- **Creation**. Where gamification must be customized by carefully choosing the number and nature of the game elements according to the user's characteristics in order to optimize learning. In the design of the proof of concept (Section 6) these considerations have been taken into account, such as the demographic characteristics of the potential users and the mechanics of the application.
- **Execution**. In this phase, user interaction is sought, resulting in user feedback that will be used for the next phase. This is the objective of the experimental design (Section 7), based on exposing the proof of concept to the experts and the choice of specific surveys to evaluate the effect of gamification.
- **Learning**. That is, about measuring and analyzing the results. In this paper, it is mainly developed through the evaluation of the results and the answer to the research questions (Section 8) as well as the analysis of the feedback given by the experts for future improvement actions (Section 9).

5. Gamification

One of the most complex and striking aspects of the development of a game-based learning application is its design approach. The correct balance between gamification and functional design is critical in the quality of an educational tool [37].

Oftentimes, the concepts of gamification and serious games in learning are confused. In this work, we have considered both term in the sense in which it is used by Landers [38]: *"Serious games incorporate all game elements, but to varying degrees; in contrast, gamification involves the extraction and application of particular elements or meaningful combinations of elements to non-game processes"*. We have chosen a gamification approach to give priority to realistic simulation, as it is aimed at a future tool to be used by professionals.

5.1. Objective and Research Questions

The goal of gamification in educational applications is to impact learning positively. Numerous studies have reported that gamified formative assessment systems that give effective feedback improve learners' performance compared to conventional assessment [5]. Yet, as empirical evidence has shown, the merely incorporating many game elements does not necessarily lead to learning improvements [39]. For motivational purposes, gamification design elements that provide feedback on one's progress and track one's own performance and that of others have been found to have yielded the best results [40]. The most used strategy for this is to combine game elements of the three types from the so-called PBL triad (points, badges, leaderboards) [41]. However, the use of items for competitive purposes can have negative effects on learning, especially in the case of leaderboards [42]. The objective of our gamification process is to obtain the positive effects on user experience and usability that game elements can provide, while avoiding the potential negative effects on learning that the most competitive aspects of games can produce.

Our hypothesis is that in the professional field in which this application is circumscribed, the simple incorporation of non-competitive game elements for feedback may be enough to improve motivation and the usability of the learning tool. The research questions posed are as follows:

- **RQ1**: *Can the incorporation of non-competitive game elements contribute to a positive user experience in an EBT simulator for instructors?*
- **RQ2**: *Can the implementation of these game elements result in the positive usability of the tool?*

Both questions address the viability of the concept in terms of gamification design. The validity of a simulator based on this concept in terms of learning transfer compared to a traditional non-gamified tool will be left for future work.

5.2. Design Elements

With the aim of finding a conceptual and methodological consensus regarding game element definition and uses, several researchers have developed theoretical frameworks for gamification [43]. Toda et al. [44] have created a taxonomy of 21 elements classified in 5 dimensions: performance, ecological, social, personal and fictional, which has been used in our design process.

The game elements chosen for this gamification are all included in the Performance dimension (see Table 2), since they are the ones used to give feedback to the learner, in line with our main hypothesis, as mentioned above. In this first approach to the problem, game elements related to the ecological dimension have been discarded, since the format is focused on the technical requirements of the simulation set by the SBT standards in EBT, defined in Section 1. Regarding the social dimension, it has been avoided in order not to introduce elements oriented towards competitiveness. The elements of the personal and fictional dimensions have also been avoided, in order to isolate the effect produced by the performance elements, although their use in future works along this line cannot be ruled out.

The elements finally selected in the simulation are the following:

Table 2
Game elements belonging to the Performance dimension, according to Toda's taxonomy [44]. These are elements related to the environmental response, which can be used to provide feedback to the learner.

Game Element	Description
Acknowledgement	• Also known as badges, medals, trophies and achievements • It is a kind of extrinsic feedback that praises the players' specific set of actions
Level	• Also known as skill level, character level, etc. • This is related to an extrinsic hierarchical layer that provides the user with new advantages as they advance in the environment
Progression	• Also known as progress bars, steps, maps, etc. • Provides an extrinsic guidance to users of their progress in the environment, allowing these users to locate themselves
Point	• Also known as scores, experience points, skill points, etc. • It is a simple way to provide extrinsic feedback for users' actions • Point is the most basic concept found in almost all gamified applications
Stats	• Also known as information, Head Up Display (HUD) and data • It is related to the visual information provided to the learner by the environment (extrinsic)

- **Level**: The simulation has been structured in levels to graduate the difficulty of the task according to the degree of familiarity with the concepts to be learned.
- **Progression**: The simulation gives feedback about the flight phase taking place at any moment during the scenario, depending on the nature of the mission, so users can locate themselves by the degree of completion of the session.
- **Points**: In the assessment stage, a grade is given to the user based on the decisions made during the scenario.

Details on how these game elements were implemented in the proof of concept are the subject of the next section.

6. Proof of Concept

Based on the modeling presented in Section 3 and the design considerations of Section 5, a proof of concept was outlined in order to represent the simulator's main functions.

Once initialized, the tool will have three stages: session setup, simulation, and assessment (see Fig. 1).

6.1. Session Setup

The simulator includes two possible levels: Beginner and Expert. The difference between the two is that at the Beginner Level the main events and conditions in the scenario (see Section 3.1) are known beforehand. This fact allows the user to foresee the possible problems that will arise in the second stage of the simulation and to prepare for them.

On the other hand, at the Expert Level, the main events in the generated scenario will not be made explicit and therefore the behavior of the pilots must be evaluated without having been able to anticipate the various problems that these events will pose.

Next, the pilot will select four parameters for the scenario to be generated by the program:

- **Aircraft Type**: A drop-down list will show the different aircraft models available for the simulation. For the time being, only the A320 series has been implemented for this proof of concept.
- **Callsign**: This is name chosen for the flight. In normal commercial operations, the callsign consists of the name of the airline followed by the flight number assigned by air traffic control.
- **Weather Profile**: The meteorological conditions of the flight will be configured according to three possible weather profiles: winter, summer, or CAVOK (acronym for *Clouds and Visibility OK*). This last profile means that the weather does not contain adverse visibility phenomena such as cloudiness, haze or fog.
- **Malfunction Type**: The various malfunctions that can be recreated in a flight simulation are grouped into aircraft systems. The name of these systems is normally preceded by a code that is universal for all aircraft (ATA code). The user will have to select one of these systems for scenario generation. This allows for working on specific system malfunctions depending on the training objectives defined by the company's EBT program at a specific time.

Once the scenario design parameters have been accepted, the tool will display the specific settings for the scenario. The scenarios contain five fields of information:

- **Pilot Crew**: The program creates a crew made up of two virtual pilots whose name, flight experience, role and competency profile are displayed. The grading of these competencies is done on a scale of 1 to 5, as proposed by the EBT training guides [1].
- **Flight Plan**: The aircraft type, callsign, as well as departure and destination airports, and departure time chosen by the program will be displayed.
- **Meteo**: The meteorological conditions of departure and destination airports will be represented in the standard aeronautical codes, in the same way that these data are presented to the pilots in real aeronautical meteorological reports.
- **Mass and Balance**: The state of load and balance of the aircraft, resulting from the total load and fuel, influence the operation of the aircraft, affecting the rates of climb and descent, speed, take-off and landing distances, etc. This information is therefore essential for the recreation of a credible flight scenario. The numerical data that define this state of load will be represented, and are the ones that will be entered in the flight equations of the virtual plane that the virtual pilots will fly.
- **Events**: Finally, on the same screen, the events chosen for the scenario will be shown as distributed by flight phase. This information is only shown at the Beginner Level.

A screenshot of an example scenario that the tool could create is shown in Fig. 2.

6.2. Simulation

The information presented on the Simulation Screen for this proof of concept consists of three main parts. First, we have the events of the scenario that will be activated in the different phases of flight. As already stated, this information will not be available at the Expert Level. Second, we have a representation of the cockpit in which we can monitor the status of the aircraft systems at all times, as well as all the navigation parameters, in the same way as if the virtual pilots were playing a conventional flight simulation game. Third, there are the pilots. For this proof of concept, the pilots have been represented with only an image of their face, and their actions and communications have been represented in text. However, in a further developed version of the tool, they could be characters with movement that would perform the actions by interacting with the cockpit and would communicate by voice, so that the non-verbal aspect could also be judged. Finally, there are the evaluation boxes available to the instructor, a fundamental function of the simulator that is explained below.

For each competency, the instructor can assign a positive or negative point to the conduct of each of the pilots. The task consists of observing the pilots' actions and communications, and evaluating said behaviors according to whether they correspond with positive or negative OBs for each of the competencies. The way to evaluate the behaviors is by clicking on the smiley- and sad-face buttons for each of the competencies. Each of the assessments that is made, will be recorded and the

Fig. 2. Flight Scenario Screen. The values of the scenario that will be implemented during the simulation session are represented on this screen. The system chooses a crew with a specific skill profile. This profile will guide the behavior of the virtual pilots during the session. The flight plan and weather data are chosen while taking into account the parameters selected by the user in the previous menu. Likewise, the plane's load data and the events chosen to be activated in the different flight phases are generated. At the Expert Level this last information will not be presented.

system will evaluate them at the end of the simulation, indicating which assessments were more or less correct according to the criteria in which the system is programmed. (See Fig. 3).

In Fig. 6.2, an example of an active event and pilot reaction is shown, as well as a possible assessment given by the user.

6.3. *Assessment*

Once the flight simulation is finished, an Assessment Screen appears. It presents a summary of all the evaluations that the user has made of the pilots. The grade is represented as a percentage of success on the total of

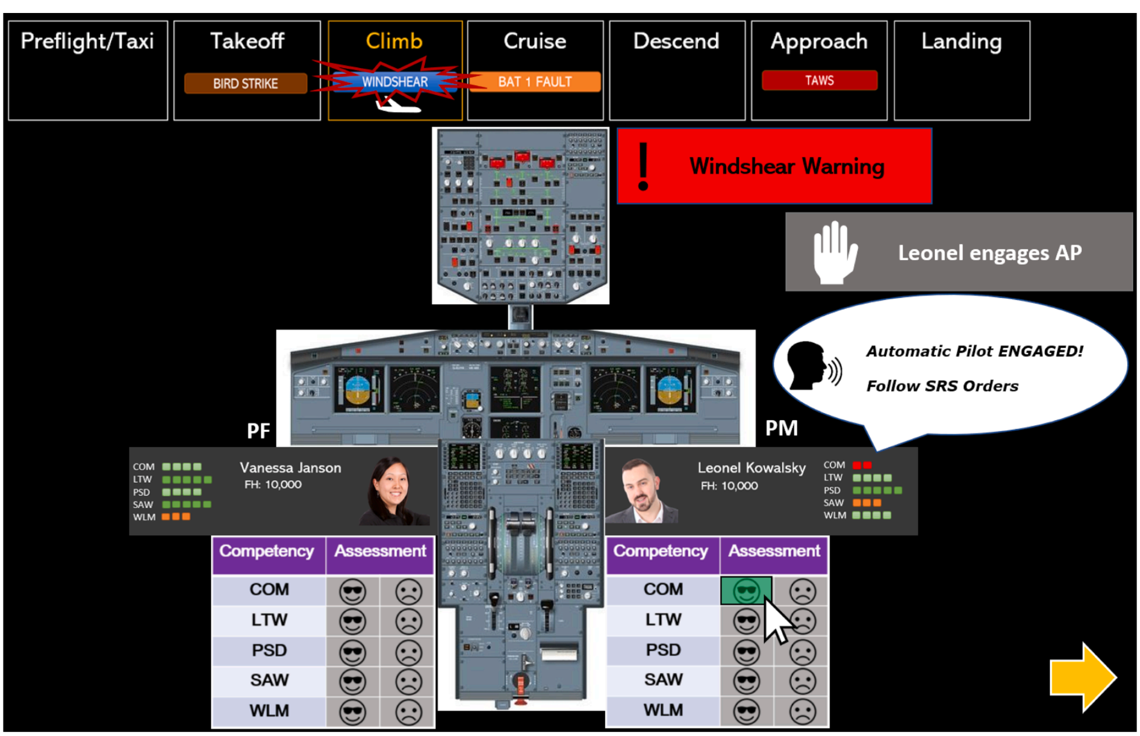

Fig. 3. Simulation Screen. The image shows a specific moment during the simulation, in the climb phase. An alarm has been triggered (windshear) and the pilot reacts with an action (represented by the text next to the icon of a hand) and a communication (represented by the text next to the speak icon). The user has observed the pilot's behavior as an example of good communication, so the smiley face next to the corresponding competition is clicked. The button lights green (a negative assessment would be red) and the specific evaluation is recorded.

gradable behaviors according to the system. For each pilot, a table is presented in which each gradable behavior is represented in a column, and possible positive or negative evaluations given to the pilot in each competency are represented in rows (see Fig. 4). A green cell in this table means that the behavior corresponding to that column was evaluated as correct according to the competency of the row in which it is located. Each of these evaluations carried out by the user is also evaluated by the system. The way to give feedback to the user about his performance in the evaluation is, on the one hand, through the total grade given, and, on the other hand, in a detailed way, through check marks for hits and X marks for errors or omissions (behaviors that were not evaluated as positive or negative in some skill) in the cells of the assessment tables. To review both the correct and incorrect evaluations, the user must click on one of the cells and the corresponding moment of the simulation will be displayed again as well as the justification for the assessment given by the system. This justification will be based on the relationship between the behavior assessed and the OBs contemplated by the EBT guidelines for each competency (see Section 3.2).

7. Experimental Set-up

A detailed video was recorded explaining each of the stages in the simulator and the functionalities that the tool will have [5]. To make the sample video, a scenario was generated and several examples of events from that scenario were simulated in different phases of flight to show how they would be depicted in the tool. Various pilot behaviors in these events were designed and evaluated in the video, thus explaining the mechanics of the simulator. Finally, the Assessment Screen is shown, depicting an evaluation of the assessments made during the simulation, as well as the feedback (justification) for each of the evaluations according to the corresponding OBs.

A group of 29 aeronautical experts were contacted and this proof of concept was sent to them, along with a questionnaire with three parts: demographics, user experience and usability.

7.1. Demographics

A demographic questionnaire was added, including several relevant questions about topics such as English proficiency and the expert's aeronautical training experience.

7.2. User Experience

As the researchers in this project are searching for a general description of the perception that the experts find in our proof of concept, we based our evaluation items on the shortest version of the standard Game User Experience Satisfaction Scale, also called GUESS-18 [45]. Some of the questions were grammatically reformulated for consistency. The questions that do not apply were directly discarded, specifically questions 9–12 relative to the Creative Freedom and Audio Aesthetics subscales. The items of the questionnaire were left as follows:

1. I find the controls of the game to be straightforward.
2. I find the game's interface to be easy to navigate.
3. I was captivated by the game concept from the beginning.
4. I enjoyed the example scenario provided by the game.
5. I felt detached from the outside world while watching the simulation.
6. I did not care to check events that were happening in the real world during this proof of concept.
7. I think the game is fun.
8. I would feel bored while playing the game.
9. Not applicable.
10. Not applicable.
11. Not applicable.
12. Not applicable.
13. I would be very focused on my own performance while playing the game.
14. I would want to do as well as possible during the game.
15. I found the game supports social interaction (e.g., chat) between players.
16. I would like to play this game with other players.
17. I enjoy the game's graphics.
18. I think the game is visually appealing.

7.3. Usability

With the objective of Searching for an indication as to whether the system appears to be usable in the future by flight instructors, we take the standard System Usability Scale (SUS) questionnaire [46] as our main reference. Some of the questions were grammatically reformulated for consistency. The specific evaluation items presented to our experts were the following:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I imagine that most flight instructors would learn to use this system very quickly.
8. I found the system very complicated to use.
9. I would feel very confident using the system.
10. I would need to learn a lot of things before I could get going with this system.

8. Results and Discussion

In this section, the results of our questionnaire are presented as discussed, following the same structure in three parts.

8.1. Demographics

Fig. 5 shows some demographic data for the sample of experts who evaluated the proof of concept. Of this group, only one person was a woman, representing 3.4 percent of the total, which is consistent with the current proportion of women in the group of airline pilots [47]. Regarding educational level, 86.2 percent have a university degree. Questionnaires in English being the standard, one of the questions was related to the experience in the use of the English language. As expected in international aviation personnel, all of them had more than 10 years of professional English speaking experience.

Fig. 6 shows a graph with the distribution of flight experience levels. 27.3 percent individually accumulate more than 5,000 flight hours, which is considered a high level of experience, and 40.9 percent had surpassed 10,000 flight hours, which is considered the highest level of aviation expertise [48]. The vast majority of the respondents had extensive experience in flight simulation, most of them, 44.8 percent, because they were captains, which implies significant flight simulator experience in the form of recurrent training, followed by 37.9 percent of instructors and examiners, who use it daily. (See Fig. 7).

8.2. User Experience

The GUESS-18 scale comprises 9 subscales. Seven of them were taken into account for this study: Usability, Narratives, Play Engrossment, Enjoyment, Personal Gratification, Social Connectivity and Visual

[5] https://youtu.be/Eb9s8TJ75eY

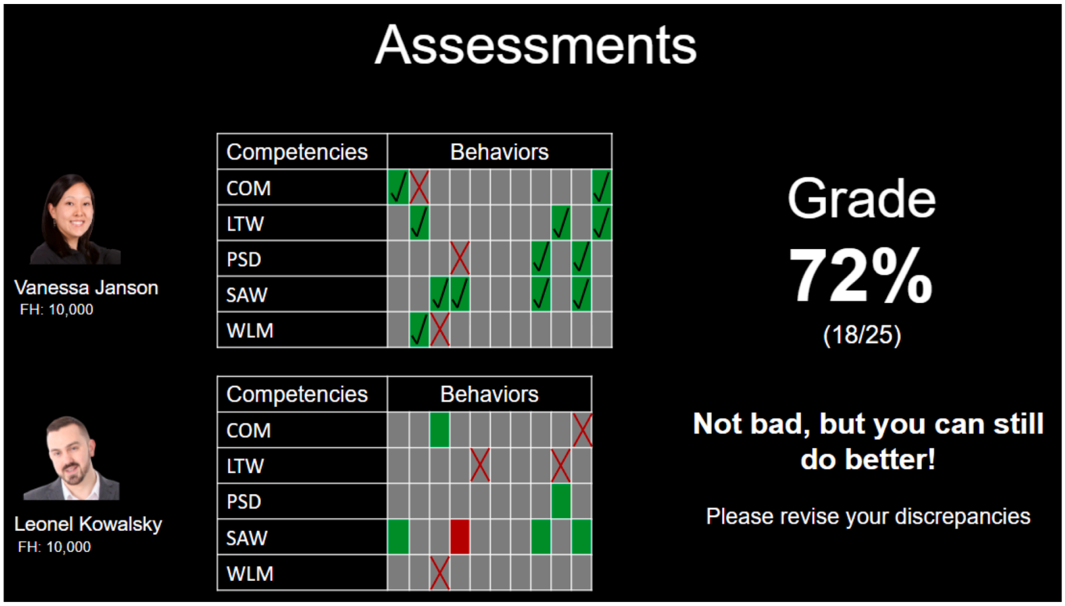

Fig. 4. Assessment Screen. Read the text for explanation.

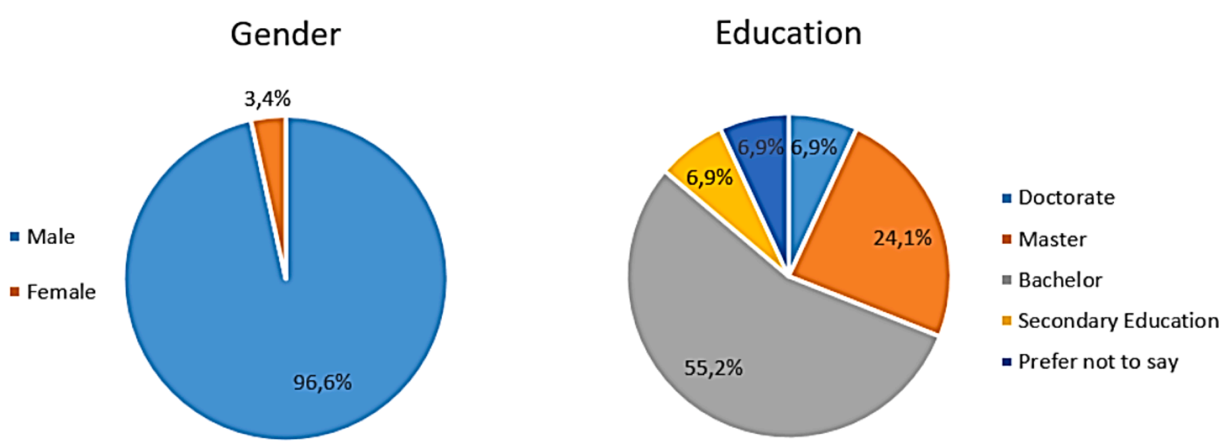

Fig. 5. Gender and education distribution of the expert sample.

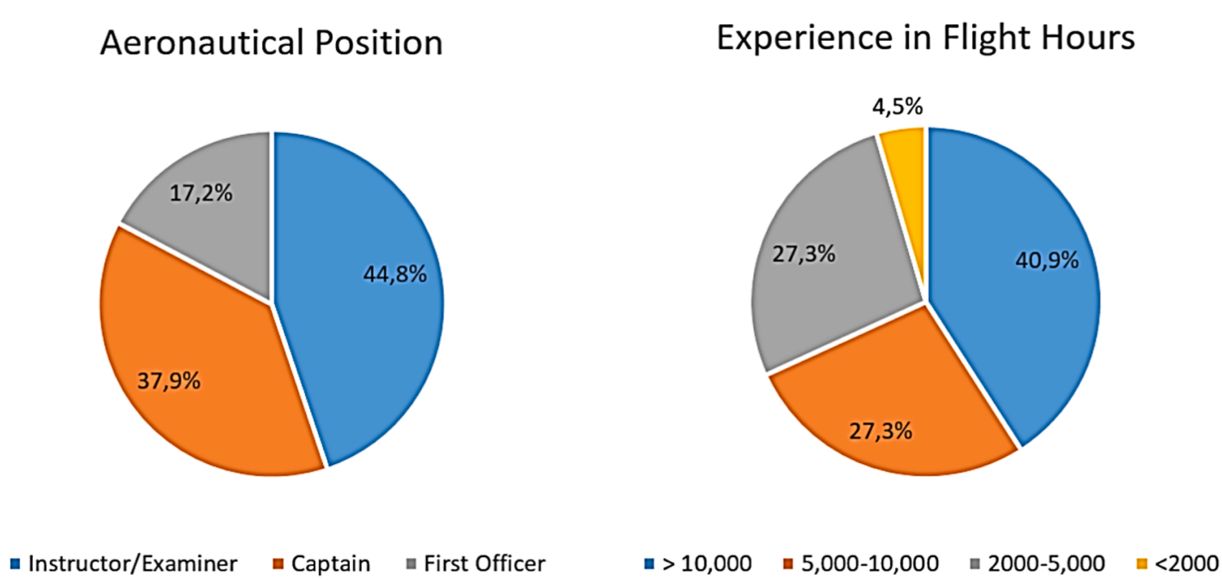

Fig. 6. Aeronautical experience of the evaluators.

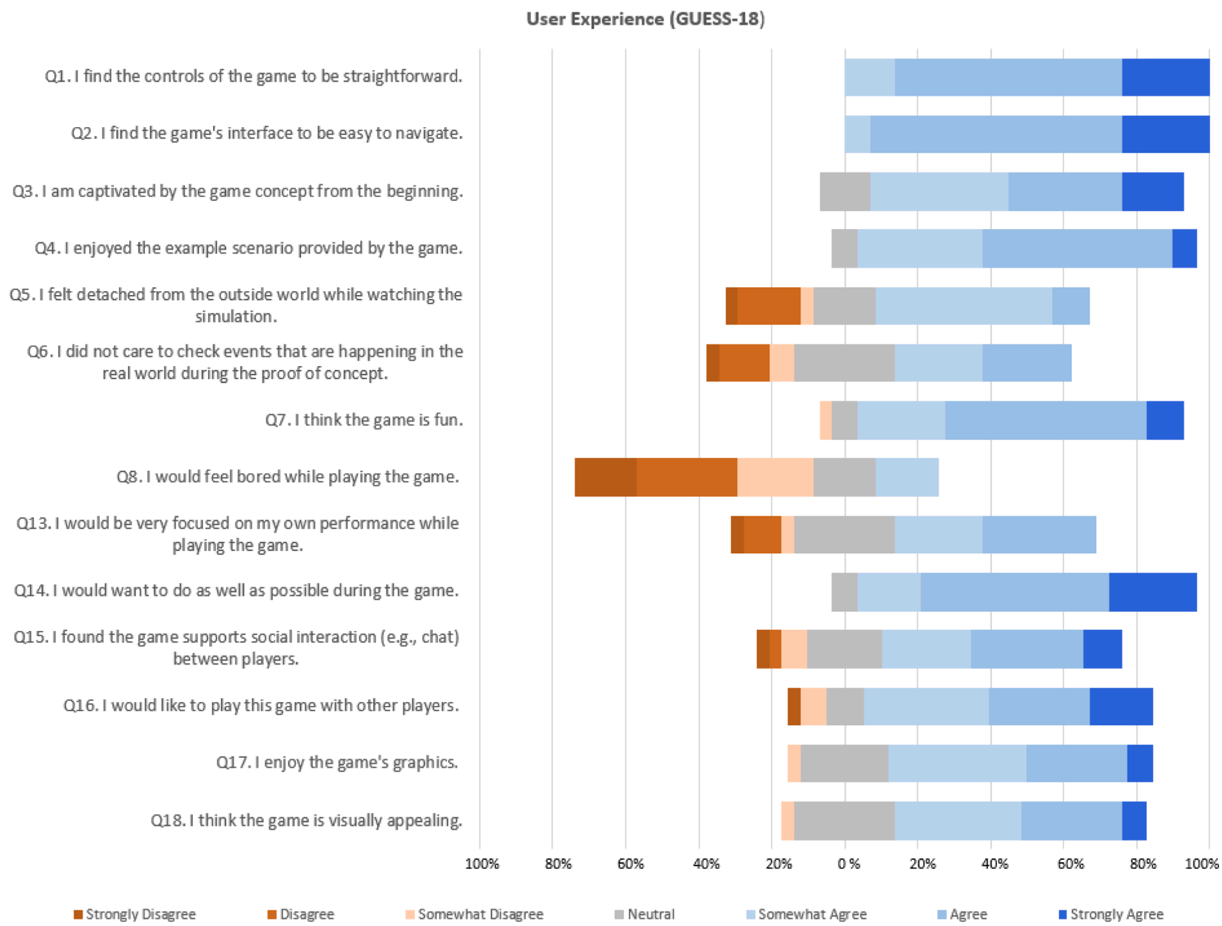

Fig. 7. Results of User Experience Questionnaire (GUESS-18). Questions 9 to 12 were discarded.

Aesthetics. Therefore, the standard Overall GUESS Score (min = 9; max = 63) should be corrected to reflect the evaluation of this proof of concept. The score for each subscale is obtained from the mean of the response means of two of the questions in the questionnaire, using a 7-point Likert scale. Therefore, the maximum value in each area is 7 points. The maximum value of the GUESS Score, having ruled out two

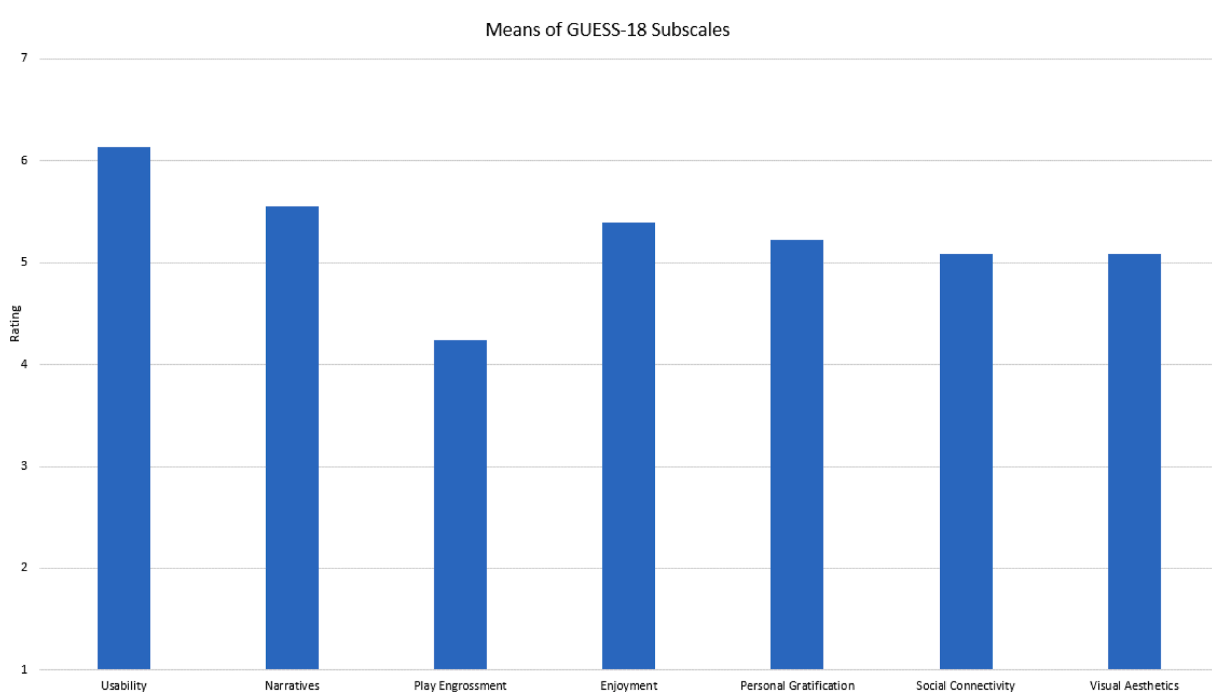

Fig. 8. Means of GUESS-18 Questionnaire. The subscales Creative Freedom and Audio Aesthetics were not evaluated.

areas, is 49. According to this corrected maximum overall score obtained by this proof of concept, it can be seen that experts perceive it as *Acceptable* in terms of usability (Mean = 38.72, Std. Dev. = 3.46).

With the exception of Play Engrossment, all the areas obtained an average of more than 5 points (see Fig. 8). The usability subscale stands out (Mean = 6.14, Std. Dev. = 0.49) with an average value above the *Agree* rating. This result is consistent with that obtained in the general usability questionnaire (see next subsection).

According to the two questions on this subscale, experts find the controls simple and the interface easy to navigate. These two issues are among the most important for the technical viability of this kind of applications. Flight scenarios are composed of operational situations whose management constitutes a challenge for the instructor, and that require the application of knowledge, skills and attitudes beyond the mere application of standard procedures. Evaluating pilot management of threats and errors in real time, determining whether the necessary competencies have been acquired and used correctly, is a highly demanding task in terms of cognitive resources. A simple assessment scheme such as the one presented in this simulator avoids unnecessarily increasing the user's workload, which can focus on the behavior of the pilots.

However, an excess of simplicity in the interaction can be problematic. Objections were raised by some of the experts to the evaluation system being reduced to positive or negative points. One of the experts expressed it as follows: "All in all, there is still some subjectivity in every appreciation of the actions to be judged, so this is not black or white, and the different responses from different users could be right even while being different". Along the same line, another expert commented: "I think that the ratings or grades given to the students are not just correct or incorrect straight away. I understand that it's quite difficult to simulate real time flight instruction, but I believe that there should be some more options". A possible improvement to this aspect of the assessment could be to refine grading, perhaps with a Likert-type scale or similar.

The subscale with the lowest value is Play Engrossment (Mean = 4.24, Std. Dev. = 1.33). However, the average is above the *Neutral* value, so in general the experts do not evaluate it negatively. In part, this effect is explainable in terms of reduced engagement due to the fact that the experts did not have real interaction with the tool. Nevertheless, predictably, this area seems the most difficult to scale up, for several reasons. The questions associated with this subscale are aimed at expressing the degree of separation from reality experienced by the user. As it is a gamified tool targeted at professionals, it must lack all kinds of narrative elements that would upstage the realistic situations and to which this type of professionals is quite accustomed. On the other hand, this simulator inevitably simplifies the experience in terms of immersion (2D, lack of movement, etc.), compared to what users of professional flight simulators are exposed to. These reasons may justify the development of the tool to incorporate the use of virtual reality, but this solution is difficult to make compatible with its execution on mobile devices, which can have a negative impact on the frequency of use, as we will analyze later.

On the Narratives subscale, experts evaluated the concept well above 5 points (Mean = 5.55, Std. Dev. = 0.71). The narrative aspect in flight simulation scenarios based on NTSs is very relevant, so it is very interesting that the experts valued it as *Acceptable*, despite the fact that they could only see one example of a scenario represented and therefore could not fathom all the narrative possibilities of the tool.

Another interesting result is a positive score in the Enjoyment subscale (Mean = 5.40, Std. Dev. = 0.98), especially considering that it is a gamified simulator oriented to a professional field in which there is a high level of stress [49]. Based on feedback from various experts, a clear line of improvement was detected in this regard, related to increased interactivity with pilots. In this first phase of development, only the passive evaluation of behaviors has been contemplated, but in a real training scenario the instructor truly interacts with the pilots, providing feedback when necessary. One expert put it this way: "I'd like to feel a little more involvement in it if I were to use it.".

Another area of great interest in this scale to assess the validity of the tool from a technical point of view is Personal Gratification (Mean = 5.22, Std. Dev. = 0.85). The ability of the tool to promote the desire to obtain a high level of performance in the task is essential for the tool to be successful in reaching its objective. In the same way, as in the Play Engrossment subscale, part of the expected improvement in this area can be explained by being a proof of concept and because the experts did not have access to the program directly, but to a video explanation, which could reduce their level of engagement. Two possible lines of improvement have been drawn from the comments of the experts. The first is related to connectivity. Being able to play with other users simultaneously can increase competition and provide an incentive to try to do better. Another possibility is to associate the events in the scenarios with real incidents. This would be possible by accessing public databases. This type of data could increase the degree of staff involvement by facing more realistic situations on the one hand, and by increasing curiosity about the real outcome compared to the simulated one.

The most striking result of the evaluation on this scale was undoubtedly Social Connectivity (Mean = 5.09, Std. Dev. = 1.27). Professional flight simulator training can be considered a very closed environment, in which usually only the participants (pilots and instructor) have knowledge of the details of the training session. The low level of sharing of what happens in training sessions has many reasons, among which are confidentiality issues. It therefore seems very interesting that the experts show interest in the connectivity of the tool (social dimension), a fact that a priori was not expected.

The final evaluation area of the GUESS scale is Visual Aesthetics. The experts also evaluated the concept as *Acceptable* with respect to this characteristic (Mean = 5.09, Std. Dev. = 0.93), although the margin for improvement is obviously great. One of the experts put it clearly: "The game shows a very useful but basic interface; that is, it would gain a lot with more elaborate graphics". Although in the current development phase, graphics was not a central objective, but rather focusing on the conceptual design of the simulator, it could be suggested that this factor is not too relevant in an application where the important factor is the cognitive aspect (linguistic and behavioral). However, users consider it to be a positive element, even in a simulation tool of this nature.

Finally, regarding the data collected in this questionnaire, it is worth highlighting an interesting comment regarding the compatibility of the tool with mobile devices. The expert put it this way: "The application would make a lot of sense if in the end it can be run from a tablet-type device, since most pilots cannot imagine their lives without a tablet…". In fact, the implementation of these devices to support all activities related to piloting (checklists, operating manuals, navigation charts, meteorological information, etc.) has completely replaced the use of paper-based flight bags in commercial aviation. Therefore, allowing the simulator to run on these devices would predictably increase their availability and frequency of use.

8.3. Usability

Unlike the GUESS Scale, SUS is not a diagnostic test. Consequently, it is not pertinent to analyze the responses in detail since they do not predict the aspects of the system that can be improved in terms of usability. Nevertheless, it is a highly validated test that gives very reliable results on the general usability of software. The SUS test uses a set of 10 questions that are answered on a 5-point Likert scale that allows a score from 0 to 100 to be calculated. This value does not constitute a percentage, and therefore the acceptability thresholds must be adjusted based on percentiles associated with the data accumulated in previous works compared to other scales [50]. Above 52 the *OK* can be considered, and the *Good* threshold is around the value 73 [51]. The SUS Score obtained by our proof of concept allows the perceived usability of the system to be considered as *Acceptable* (Mean = 67,93, Std. Dev. = 9.28), consistent with usability GUESS-18 subscale result.

Given the results, the answer to the two research questions posed in this work can be considered positive. This provides very promising perspectives for a tool based on the concept defined in this paper, since it was obtained with a proof of concept, that is, with an unfinished and unpolished tool, whose functionality has been simplified and exposed schematically for the purposes of evaluating the feasibility of its development. It is therefore reasonable to expect that once developed, the margin for improving the usability of the system is considerable, as has been analyzed in the previous subsection.

9. Conclusions and future work

This work presents a proof of concept of a gamified simulator for flight instructors aimed at training in the assessment of non-technical pilots' competencies. The tool was designed in accordance with the principles of the EBT paradigm, a paradigm currently advocated by international aeronautical authorities, which places the emphasis on the development of non-technical competencies. The concept of such a tool was tested by 29 experts directly involved in flight instruction, using questionnaires with three parts, one for demographics, another for user experience and a final one for usability.

The results of the validation allow the researchers to conclude that the development possibilities of a tool of this type are very promising. Although it is only a proof of concept, the experts have provided acceptable usability values on both the SUS scale and the corresponding subscale in GUESS-18. Likewise, in this last scale users positively valued the concept in the 7 applicable subscales, highlighting the usability area with an average value higher than 6 was been shown in relation to Play Engrossment, although the average value obtained was slightly higher than neutral. However, a high valuation in this area is hardly expected in a non-interactive proof of concept.

Based on the results and the feedback obtained with our study, we discuss below the lines of future work that we consider most interesting to continue the development of this project:

- Simple interface and controls that allow easy assessment have a positive impact on usability in a simulation tool. However, when dealing with non-technical behavior, their evaluation is rarely binary (good or bad), so it would be advisable to reformulate the evaluation system into one based on grades, similar to those used in the EBT paradigm.
- One of the competencies not considered in our design phase is the ability of the instructor to give feedback to the students at the appropriate moments, since the focus was placed on the passive evaluation of the students' behavior. Experts emphasize the need for more frequent and complex interaction to make the tool more interesting. The implementation of this functionality would positively impact not only the technical viability of the gamified simulator, but also its enjoyment. In fact, we have started working on a prototype that allows a mission to be aborted when the instructor considers that the pilots have lost the focus of the training.
- Once the tool is built, its validation, measuring its transfer of learning, would be essential. This would imply designing controlled experiments that would allow us not only to measure how much the tool contributes to improving the performance of instructors and reducing errors in real training situations, but also to verify whether or not the tool induces any negative training. This means to determine whether developing the skills that the tool stimulates is detrimental to others that are also necessary in real training, such as those related to technical skills like the specific application of procedures or the manual or automatic handling of the aircraft.
- Another interesting area for improvement is social connectivity. This is an unexpected result of this work, and it opens an interesting avenue for the purpose of sharing experience among experts in an environment where specific knowledge is normally highly confidential. This functionality can also have a positive impact on user gratification.
- Finally, because of the nature of their profession (pilots in our case), many students could spend large amounts of time away from home, and access to a PC is often not possible. A serious consideration is therefore to guide the development of this type of applications towards mobile devices (touch-screen tablets) that accompany students throughout their habitual activity. The interface is currently being redesigned with this deployment need in mind.

Despite the limitations of this work, the viability of a gamified simulator such as the one presented in this study has been proven. It allowed us to get in contact with specialists in this industry and we are currently collaborating with them on the development of a fully functional prototype of a gamified simulator. Furthermore, we think that we have unraveled some important keys for the gamified design of future environments for specific training based on non-technical competencies, an area of increasing interest in our scientific community.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors would like to thank the help of EASA Approved Training Organization *Global Training and Aviation, S.L.*

References

[1] ICAO, Manual of evidence-based training, Montreal: International Civil Aviation Organization, 2013.
[2] IATA, Evidence-based training implementation guide, 1st edition, Montreal: International Air Transport Association, 2013.
[3] IATA, Data report for evidence-based training, 1st edition, Montreal: International Air Transport Association, 2014.
[4] M. Henri, M.D. Johnson, B. Nepal, A review of competency-based learning: Tools, assessments, and recommendations, J. Eng. Educat. 106 (2017) 607–638.
[5] Z. Zainuddin, S.K.W. Chu, M. Shujahat, C.J. Perera, The impact of gamification on learning and instruction: A systematic review of empirical evidence, Educat. Res. Rev. 30 (2020) 100326.
[6] C.D. Wickens, J.G. Hollands, S. Banbury, R. Parasuraman, Engineering psychology and human performance, Psychology Press, 2015.
[7] L. Allen, Evolution of flight simulation, in: Flight Simulation and Technologies, 1993, p. 3545.
[8] R.L. Page, Brief history of flight simulation, SimTecT 2000 Proceedings (2000) 11–17.
[9] S. Popular Mechanics, p. 87, Airline pilots fly anywhere in world whithout leaving the ground (1954).
[10] P. Iwaniuk, Microsoft flight simulator review. taking 'open world' very literally, PCGamer (accessed September 14 2020). URL: https://www.pcgamer.com/microsoft-flight-simulator-review/.
[11] C. Massoth, H. Röder, H. Ohlenburg, M. Hessler, A. Zarbock, D.M. Pöpping, M. Wenk, High-fidelity is not superior to low-fidelity simulation but leads to overconfidence in medical students, BMC Med. Educ. 19 (2019) 29.
[12] E. Salas, K.A. Wilson, E.H. Lazzara, H.B. King, J.S. Augenstein, D.W. Robinson, D. J. Birnbach, Simulation-based training for patient safety: 10 principles that matter, J. Patient Saf. 4 (2008) 3–8.
[13] E. Salas, H.A. Priest, K.A. Wilson, C.S. Burke, Scenario-based training: improving military mission performance and adaptability, Military Life: Psychol. Serv. Peace Combat 2 (2006) 32–53.
[14] I. Munene, An application of the hfacs method to aviation accidents in africa, Aviat. Psychol. Appl. Human Factors (2016).
[15] N. Bienefeld, G. Grote, Silence that may kill, Aviation Psychology and Applied Human Factors (2012).
[16] J. Ford, D. O'Hare, R. Henderson, Putting the "we" into teamwork: Effects of priming personal or social identity on flight attendants' perceptions of teamwork and communication, Human factors 55 (2013) 499–508.
[17] P. Farago, M.L. Shuffler, E. Salas, The design, delivery, and evaluation of crew resource management training, in: Crew Resource Management, Elsevier, 2019, pp. 251–282.
[18] F. Jentsch, C. Bowers, D. Berry, W. Dougherty, J.M. Hitt, Generating line-oriented flight simulation scenarios with the RRLOE computerized tool set, Proc. Human Factors Ergon. Soc. Annu. Meet. 45 (2001), 749–749.

[19] M. Curtis, F. Jentsch, Line operations simulation development tools, in: Crew Resource Management, Elsevier, 2019, pp. 323–341.
[20] IATA, Guidance material for instructor and evaluator training, 2nd edition, Montreal: International Air Transport Association (2021).
[21] S.A. Bennett, Aviation crew resource management – a critical appraisal, in the tradition of reflective practice, informed by flight and cabin crew feedback, J. Risk Res. 22 (2019) 1357–1373, https://doi.org/10.1080/13669877.2018.1459800, arXiv:https://doi.org/10.1080/13669877.2018.1459800.
[22] J. Pirker, C. Gütl, Educational gamified science simulations, in: Gamification in education and business, Springer, 2015, pp. 253–275.
[23] H. Al Fatta, Z. Maksom, M.H. Zakaria, Game-based learning and gamification: Searching for definitions, Int. J. Simul. Syst. Sci. Technol. (2018).
[24] S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: defining gamification, in: Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments, 2011, pp. 9–15.
[25] A. Ahmed, M.J. Sutton, Gamification, serious games, simulations, and immersive learning environments in knowledge management initiatives, World J. Sci., Technol. Sustain. Develop. (2017).
[26] M.T. Bonde, G. Makransky, J. Wandall, M.V. Larsen, M. Morsing, H. Jarmer, M. O. Sommer, Improving biotech education through gamified laboratory simulations, Nat. Biotechnol. 32 (2014) 694–697.
[27] L. Gatti, M. Ulrich, P. Seele, Education for sustainable development through business simulation games: An exploratory study of sustainability gamification and its effects on students' learning outcomes, J. Clean. Prod. 207 (2019) 667–678, https://doi.org/10.1016/j.jclepro.2018.09.130. URL: https://www.sciencedirect.com/science/article/pii/S0959652618328531.
[28] C. Badea, Gamification: Improving Supervisory Control Performance in Highly Automated Air Traffic Control, Master's thesis, Delft University of Technology, 2021.
[29] C. Peng, L. Cao, S. Timalsena, Gamification of apollo lunar exploration missions for learning engagement, Entertain. Comput. 19 (2017) 53–64.
[30] F. Cornelissen, Gamification for astronaut training, in: SpaceOps 2012, 2012, p. 1275405.
[31] D. Noh, The Gamification Framework of Military Flight Simulator for Effective Learning and Training Environment, Master's thesis, College of Engineering and Computer Science, University of Central Florida, 2020.
[32] E. Kuindersma, J. Field, J. Pal, Game-based training for airline pilots, in: Simulation-Based Training For The Digital Generation Conference, National Aerospace Laboratory NLR, 2015, p. Electronic Proceedings.
[33] J.S. Brown, A. Collins, P. Duguid, Situated cognition and the culture of learning, Educ. Res. 18 (1989) 32–42.
[34] M. Ney, V. Emin, J. Earp, Paving the way to game based learning: a question matrix for teacher reflection, Proc. Comput. Sci. 15 (2012) 17–24.
[35] R.L. Helmreich, A.C. Merritt, J.A. Wilhelm, The evolution of crew resource management training in commercial aviation, Int. J. Aviat. Psychol. 9 (1999) 19–32.
[36] A. Mora, P. Zaharias, C. González, J. Arnedo-Moreno, Fraggle: A framework for agile gamification of learning experiences, in: Revised Selected Papers of the 4th International Conference on Games and Learning Alliance-Volume 9599, 2015, pp. 530–539.
[37] P. Caserman, K. Hoffmann, P. Müller, M. Schaub, K. Straßburg, J. Wiemeyer, R. Bruder, S. Göbel, Quality criteria for serious games: serious part, game part, and balance, JMIR serious games 8 (2020) e19037.
[38] R.N. Landers, Developing a theory of gamified learning: Linking serious games and gamification of learning, Simul. Gaming 45 (2015) 752–768.
[39] S. Bai, K.F. Hew, B. Huang, Does gamification improve student learning outcome? evidence from a meta-analysis and synthesis of qualitative data in educational contexts, Educ. Res. Rev. 30 (2020) 100322.
[40] J.R. Chapman, P.J. Rich, Does educational gamification improve students' motivation? if so, which game elements work best? J. Educ. Bus. 93 (2018) 315–322.
[41] R. Leitão, M. Maguire, S. Turner, L. Guimarães, A systematic evaluation of game elements effects on students' motivation, Educ. Inform. Technol. (2021) 1–23.
[42] A.M. Toda, P.H.D. Valle, S. Isotani, The dark side of gamification: An overview of negative effects of gamification in education, in: A.I. Cristea, I.I. Bittencourt, F. Lima (Eds.), Higher Education for All. From Challenges to Novel Technology-Enhanced Solutions, Springer International Publishing, Cham, 2018, pp. 143–156.
[43] O. Azouz, Y. Lefdaoui, Gamification design frameworks: a systematic mapping study, in: 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), IEEE, 2018, pp. 1–9.
[44] A.M. Toda, A.C. Klock, W. Oliveira, P.T. Palomino, L. Rodrigues, L. Shi, I. Bittencourt, I. Gasparini, S. Isotani, A.I. Cristea, Analysing gamification elements in educational environments using an existing gamification taxonomy, Smart Learn. Environ. 6 (2019) 1–14.
[45] J.R. Keebler, W.J. Shelstad, D.C. Smith, B.S. Chaparro, M.H. Phan, Validation of the guess-18: A short version of the game user experience satisfaction scale (guess), J. Usabil. Stud. 16 (2020).
[46] J. Brooke, Sus: a quick and dirty usability scale, Usability evaluation in industry (1996) 189–194.
[47] F. McCarthy, L. Budd, S. Ison, Gender on the flightdeck: Experiences of women commercial airline pilots in the uk, J. Air Transp. Manage. 47 (2015) 32–38.
[48] G. Li, S.P. Baker, J.G. Grabowski, Y. Qiang, M.L. McCarthy, G.W. Rebok, Age, flight experience, and risk of crash involvement in a cohort of professional pilots, Am. J. Epidemiol. 157 (2003) 874–880.
[49] E. Demerouti, W. Veldhuis, C. Coombes, R. Hunter, Burnout among pilots: psychosocial factors related to happiness and performance at simulator training, Ergonomics 62 (2019) 233–245.
[50] J. Brooke, Sus: a retrospective, J. Usabil. Stud. 8 (2013) 29–40.
[51] J. Sauro, A practical guide to the system usability scale: Background, benchmarks & best practices, Meas. Usabil. LLC (2011).