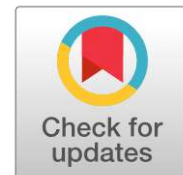


THE RECURRING ISSUE OF AVIATION ENGLISH TEST VALIDITY: ECHOES FROM TEST-TAKERS AND ASSESSORS OF THE ENGLISH FOR AVIATION LANGUAGE TESTING SYSTEM IN ALGERIA

Tarek Assassi¹, Tarek Ghodbane²

¹Mohamed Khider University, Biskra, Algeria

²University of Ghardaia, Ghardaia, Algeria



Abstract

The English for Aviation Language Testing System (EALTS) is one of the international tests for pilots wishing to operate in international airspace. This test presents a wide range of difficulties for Algerian pilots, but little research has been conducted to estimate its validity. This paper seeks to answer two questions: (1) What are test-takers perspectives on the EALTS test construct? and (2) What is the assessors' perception of the testing procedure and test validity? A descriptive study was conducted using qualitative data from a semi-structured questionnaire for ten pilots and a semi-structured interview for three certified assessors and four university researchers. The authors used 'first-hand' data from the targeted sample to cross-check results' validity through triangulation. The results show specific difficulties from an affective dimension, such as stress and anxiety caused by inefficient preparation and unfamiliarity with test tasks. A different interpretation of ICAO descriptions using the rating scale is another issue noted by assessors. Additionally, technical issues with the computer-based listening test and non-compliant features of the test contents with features of the target situation language use are among the main issues noted by both test-takers and assessors.

Keywords: ESP, Assessment, Aviation English, test validity, test design, aviation phraseology, language proficiency

Article history:

Received: 06 September 2023

Reviewed: 18 September 2023

Accepted: 22 September 2023

Published: 20 December 2023

Contributor roles:

Conceptualization: T.A. (lead); Data curation: T.A, T.G. (equal);

Investigation: T.A. (lead), T.B. (supporting); Methodology: T.A. (lead),

T.G. (supporting); Writing – original draft: T.A. (lead);

Writing – review and editing: T.A. (lead), T.G. (supporting)

Copyright © 2023 Tarek Assassi and Tarek Ghodbane




This open access article is published and distributed under a [CC BY-NC 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited. Permissions beyond the scope of this license may be available at t.assassi@univ-biskra.dz. If you want to use the work commercially, you must first get the authors' permission.

Citation: Assassi, T, & Ghodbane, T. (2023). The Recurring Issue of Aviation English Test Validity: Echoes from Test-Takers and Assessors of the English for Aviation Language Testing System in Algeria. *English Studies at NBU*, 9(2), 239-269. <https://doi.org/10.33919/esnbu.23.2.6>

Correspondence: Tarek Assassi is a senior lecturer, researcher and pedagogical coordinator in the English Language and Literature Department at Biskra University, Algeria. He is a certified assessor of Aviation English using EALTS UK. His main research interests are ESP course design and language testing.

E-mail: t.assassi@univ-biskra.dz

 <https://orcid.org/0000-0002-3514-1941>

Tarek Ghodbane is an associate professor in the English Language Department at Ghardaia University, Algeria. He is a certified assessor of aviation English using EALTS UK. His main research interests are language testing and teaching English as a foreign language.

E-mail: tarkghodbane@gmail.com

 <https://orcid.org/0000-0002-4517-7438>

Following the adoption of the Convention on International Civil Aviation (also known as Chicago Convention) in 1944, the ICAO recommended the use of the English language as the international language of aviation (Annex 10, Vol. I, 5.2.1.1.2). As a result, any radiotelephony conversation held by aeronautics professionals in international airspace must be in English. Therefore, it is highly required for all aviation personnel, most notably pilots and air traffic controllers (ATCs), in international airspace, to communicate effectively using the English language. Pilots must demonstrate their communicative proficiency by taking a standardized language test recognized by the International Civil Aviation Organization (ICAO). A shortlist of recognized tests can be found online, and test providers must rely on one of them. The EALTS (English for Aviation Language Testing System) is not one of these tests (see. Figure 1), yet it has been administered as an internationally recognized test since 2008, particularly in Algeria since 2015.

Figure 1.

Aviation English Tests Recognized by the ICAO

Test	Test Provider	Recognition
English Language Proficiency for Aeronautical Communication (ELPAC) An English language test to assess air traffic controllers as required by the ICAO LPR	Eurocontrol www.elpac.info	<input checked="" type="checkbox"/>

Note. Screen Capture by T. Assassi, December 8th, 2022. <https://www4.icao.int/aelts/Home/RecognizedTests>

The Aures Aviation Academy is the exclusive test provider in the country. Nearly three hundred (300) candidates sat for the test in two years (president of the academy). The significance of the study stems from the significance of the test and its effect on candidates' performance and licensure to be operational pilots. Communication is one

of the major safety concerns in the fastest-growing industry nowadays. Communication-related issues have been identified in the literature as a main contributing factor in several incidents and accidents, as summarized in Table 1, Cookson (2009), Cushing (1994).

In March 2013, ICAO held a Language Proficiency Requirements (LPRs) Technical Seminar designed to assist States and the industry with the implementation of the safety-critical language provisions (ICAO, 2013). At the seminar, Nicole Barrette, Technical Specialist (Training and Licensing standards) explains that the implication of this is that all pilots and controllers involved in international operations that do not share a common language must have stated on their license their level of English language proficiency. Flanagan (ICAO, 2013), AELTS Manager, also adds that language is a component of communication and the ability to speak at a certain level of proficiency to communicate with each other if something unexpected occurs is essential. As a result, test providers and assessors must eliminate any discomfort or confusion before tests and ensure candidates share their real language level with minimal impeding factors. Since “clear communication is critical because of the potential safety repercussions of misunderstandings” (Cox & Karimi, 2022, p. 183), and most of the tests are not recognized by the ICAO (Alderson, 2008), what constitutes aviation English test validity? And who validates these internationally recognized tests? ICAO did not design or develop any aviation English tests; however, their description of proficiency requirements can be used for designing different types of tests, namely, proficiency tests.

After checking the ICAO website to see the officially recognized aviation English tests, we noticed that only one test is recognized, as shown above in Figure.1. To be sure of the data we received, we checked the website repeatedly during the first 6 months of 2022. After, we confirmed the information with the head of the local aviation academy that provides the test, to receive the same answer stating that the test is recognized in Algeria and several other countries around the world even if the official website does not confirm the information. Accordingly, the EALTS follows the description of proficiency requirements set by the ICAO EALTS Handbook, (2012). Many researchers specializing in applied linguistics have investigated assessment in aviation English (Fowler et al., 2021; Garcia & Fox, 2020; Knoch, 2014; Dusenbury & Bjerke, 2013; Moder

& Halleck, 2012; Alderson, 2010; Douglas 2001, 2000). By the same token, Friginal et al., (2020) claim that investigating communication between pilots and air traffic controllers to elicit pedagogical implications for aviation English educational programs has been an important aspect for applied linguists. However, there is still “little confidence in the meaningfulness, reliability, and validity of several of the aviation language tests currently available for licensure” (Alderson, 2010, p. 1). Kim and Elder (2014) confirm the importance of eliciting information from different stakeholders who are well-placed to assess communication in the aviation field. From this point, there is a growing need for test validity given the immense importance of communication in aviation operations and eliciting information from the concerned parties, mainly test-takers and assessors. This study addresses the key issue of test validity through information from the target situation. We draw the data from test-takers' performance in the test, in addition to certified assessors' and aviation English teachers' views on the testing process and its compliance with the ICAO linguistic requirements. Therefore, the study addresses the following two fundamental questions: (1) what is the test-takers' perspective on the EALTS test construct? And (2) what is the assessors' perception of the testing procedure and test validity? The research paper begins with a general overview of the literature on aviation English tests. Next, a description of the research procedure and the case under investigation is provided. Finally, a discussion of the main findings concerning the main issues faced by test-takers of the EALTS and assessors'/professors' perspectives on the test are debated. Elicited suggestions and recommendations were also elaborated. The study sees if the test complies with the Language Proficiency Requirements of the ICAO, if it mirrors Target Language use, and if it manifests real-life situations using test-takers and assessors' perspectives.

Literature Review

In this section, an extensive literature review has been conducted to explore what was already discussed concerning this research problem, that is to say, the issue of testing validity in aviation English assessment procedures as a high-stakes test and the compliance of the internationally recognized tests with the ICAO Linguistic Proficiency Requirements (LPRs). To follow the chain of events, there is a substantial increase in the demand for air travel around the globe. This has led to a significantly high demand for flight training. Subsequently, the training programs for novice pilots and ATCs must

follow very strict standards considering the safety issues, most notably communication-related incidents and accidents. “Inadequate English language proficiency is a significant safety issue that causes delays in progress and could even prevent international student pilots from completing their flight training” (Fowler et al., 2021, p. 27). As the de facto language for international aviation talk, communicating in English among the aviation community has become a major concern given the percentage of non-native English speakers NNES compared to native English speakers NES. Miscommunication on radiotelephony might occur due to problems related to the speaker, problems related to the channel, and/or problems related to the listener (ICAO, 2010). First, the problems related to the speaker include propositional failure (*e.g., inaccurate assumptions about shared background knowledge with the listener*), encoding failure (*e.g., wrong choice of vocabulary or grammar mistakes*), or delivery failure (*e.g., pronunciation problems or inappropriate speech rate*). The problems related to the listener can be due to decoding failure (*e.g., language, attention, memory problems*), interpretation failure, which can be a consequence of the speaker encoding problems, and feedback failure.

Research has also shown that the testing policy lack of fit with pilots’ and ATCOs’ real-life communicative needs might lead to construct underrepresentation (Douglas, 2014; Kim, 2012; Kim & Elder, 2015), which may in turn: i) threaten the validity of implications drawn from test scores; ii) impact individuals, on teaching and learning activities and testing policies and practices; and iii) bring about potentially deadly unintended consequences (Messick, 1989).

Safety concerns due to the breakdown of communication in aviation have led many researchers to pinpoint the ins and outs of communicative issues and the application of ICAO’s linguistic requirements. Fowler et al., (2021) found that the number of reported incidents due to inadequate English language proficiency did not decrease after the 2003 ICAO LPRs’ strengthening procedure. They reported data drawn from the National Aeronautics and Space Administration (NASA) and the Aviation Safety Reporting System (ASRS) between the years 2009 and 2019.

The flight crew involved in the famous Tenerife accident were highly experienced airline pilots. KLM 4805 Captain Jacob Van Zanten had approximately 11,700 total flight hours, with 1,545 hours on the Boeing 747 (NASB, 1978). This led us to consider that if

highly experienced pilots with that many flight hours can fall into such a breakdown of communication, novice pilots who must communicate in English as a second or foreign language are more likely to commit more serious communication errors. To back up this claim, a more recent accident was reported by the Transport Safety Board of Canada (2018). On March 17, 2017, two Cessna 152 aircrafts dedicated to training were involved in a fatal mid-air collision 1.7 nautical miles east southwest of the Montreal St. Hubert Airport (CYHU). As Fowler et al., (2021) indicated, student pilots had 135.8 and 39.5 total flying hours. They added that the accident was a result of a breakdown of communication, as both international student pilots were neither native speakers of English nor French, the main languages used in Canadian airspace.

Serious safety concerns in the aviation industry, especially those caused by a breakdown of communication, have been discussed in the literature. Yet, much more open discussions are needed. The teaching, learning, and assessment of aviation communication that led to licensure for pilots to be operational require much more effort and openness. Regarding testing for licensure, which is the last and most serious step to effective communication in the target situation, much more research is needed to shed light on test performance issues, test constructs, and test validation. Unfortunately, there are few aviation English assessment programs available to evaluate NNES flight students for aviation English proficiency. There are also very few aviation English training programs are available for those who are unable to demonstrate proficiency (Fowler et al., 2021)

The test design process and validation have been very hot and sensitive topics given the seriousness of the damages they cause if administered poorly, and the effect the criticism may cause on one of the most flourishing fields of business, i.e., the aviation industry. According to previous studies (Alderson, 2008; Moere et al., (2009); Alderson, 2010; Dusenbury & Bjerke, 2013), test construct and validation is a highly technical and stern procedure. The studies concluded that test validation is a must since most of the available tests nowadays are not validated by any organism, yet recognized by national authorities and are currently operational. One of the most discussed causes of the validation issue is elaborated by Moere et al., (2009), stating that although the ICAO manual is informative, little information is provided about the development and theoretical rationale for the rating criteria, they added that this can present a challenge not only to test designers but will lead to poorly operationalized criteria then to different interpretations

of requirements. By the same token, different tests, as we can notice currently, designed using the same criteria, could exhibit variation in task requirements and scoring, they added. Eventually, the test validity and reliability may be put in question.

Test-takers' feedback and perception as domain experts with relevant work experience play a major role in the validation process of high-stakes tests. As a sign of test validity, although test-takers' perceptions are extensively discussed and argued that in Language for Specific Purposes (LSP), their feedback as domain experts can provide insightful details as far as effective communication in the target situation is concerned (Katsarska, 2021; Borowska, 2018; Elder, 2007; Douglas, 2005, 2000; Douglas & Myers, 2000; Jacoby & McNamara, 1999; Lumley & Brown, 1996; Brown, 1993); however, face validity and appearances are the only evidence for eliciting test-takers' perception (Davies et al., 1999).

Brown (1993) studied test-takers' feedback on a tape-mediated oral proficiency test of Japanese in the tourism and hospitality industry. She found that test takers' feedback was of paramount importance in eliciting the type of language needed in the industry. She concluded that feedback from domain experts can serve as a source of validity for tests and must be considered by policymakers to determine the relevance of the test construct to the target language use; therefore, deciding the test appropriacy for determining its validity for the target situation's communication. Elder (2007) compared the Occupational English Test (OET), a test for measuring health professionals' communicative competence, and the International English Language Testing System (IELTS) designed for measuring academic proficiency. The two tests are used interchangeably for screening purposes of health professionals' registration process. The fifty-three health professionals who took both tests shared their feedback on the nature of the tests and their compliance with the communicative requirements in their fields. They believed that even if the IELTS is a valid test for measuring general English proficiency; the OET is a more efficient test to measure their specific professional communicative competence.

On a different note, Bassette (2005) studied test-takers' objections to mandated proficiency testing policy aiming at ensuring public servants' proficiency in English and French in bilingual regions under the Official Language Acts in Canada. The test-takers'

feedback as domain experts revealed negative perceptions of the test, which they claimed was unfair, subjective, and irrelevant to their target situation. This shows that feedback from domain experts can be negative and create a challenge for test validation. In the Australian state of New South Wales, Murray, Riazi, and Cross (2012) consulted domain experts. Their study investigated 105 internationally qualified teacher test-takers' attitudes toward the Professional English Assessment for Teachers (PEAT). The professional screening test administered in the abovementioned state faced more objections from teachers who had experience teaching in Australian schools than those with no teaching experience. The test-takers' dissatisfaction might be rooted, according to the authors, in the teachers' newly acquired expertise, resulting in confidence in their skills to work more effectively in the target situation; thus, claiming that PEAT did not meet their needs.

Zhou and Yoshitomi (2019) investigated test-takers' perceptions of 64 Japanese university students who took the Test of English for International Communication (TOEIC). The researchers assumed that negative test-taker perception may influence students' test performance by decreasing test-taking motivation. The assumptions were not verified, but the authors revealed students' reservations about computer delivery.

Relatively closer to the field of aviation and the current study, Knoch (2014) revealed that Native English Speaker pilots presented a larger number of criteria encompassing the ICAO guidelines when asked to evaluate speech samples from several aviation English tests. The criteria include non-linguistic factors, such as purely technical awareness, professional experience, and differences in training levels. On the same train of thought, Kim and Elder (2014) analyzed 400 questionnaires and 22 interview sources from Korean airline pilots and air traffic controllers, investigating both the construct of the English language proficiency test for these aviation professionals and the ICAO proficiency testing policy. The test developed and administered in Korea, in domain experts' opinion, lacked a fit between the policy construct and the reality, in addition to the strong disapproval of the ICAO's advocated construct and the test itself from language users in the target field. The authors added that eliciting the views of such stakeholders who are well-placed to determine efficient communication in the target context is of paramount importance.

The ICAO Linguistic Requirements

Communicative competence is defined as the appropriate usage of language between participants in a specific social context or situation (Hymes, 1972). Communicative issues have always been and still are a concern to civil aviation policymakers. Breakdown of communication, which is more than just grammatical formations, but more of a working aspect of language use (Assassi & Benyelles, 2016, p. 167) is crucially important in aviation radiotelephony; in other words, it must be avoided at any cost given the catastrophic consequences to which they may lead. The following table shows three examples of aviation accidents that were caused by a breakdown of communication.

Table 1.

Three Major Crashes Caused by Breakdown of Communication (Assassi, 2016, 2020)

Dates	Aircrafts	Companies	Effects	Victims
1977	Two 747 Boeings	KLM (Royal Dutch Airlines) & Pan American Airlines	Collided on the runway in Tenerife	583
1990	Boeing 707	Avianca Flight	Crashed into a village due to fuel exhaustion near the JFK airport	73
1995	Boeing 757	American Airlines	Flew into a terrain in Cali	159

The reason behind this focus on the working aspect of language, English in this case, is to shed light on fluency and interaction as well as not only the knowledge of grammatical rules and a set of lexical items that marks out a competent language learner/user. Equally important, the ICAO linguistic requirements focus on fluency and interaction, giving the significance of the very limited talking time on radiotelephony that is counted in seconds only. This is indicated in the ICAO (2006) and the Civil Aviation Authority (2016), both manuals of radiotelephony communication, that encourage brief and straightforward messages that reduce the risk of errors and misunderstanding. There is a six holistic descriptors chart on which Aviation English Certified Assessors must rely to evaluate candidates' language based on six levels of proficiency, as it is shown in Appendix A.

The rating scale starts with level one labeled as (Pre-Elementary level 1) and ends with level six (Expert level 6). Every holistic descriptor (Pronunciation, Structure, Vocabulary, Fluency, Comprehension, and Interaction) has six levels of mastering the descriptor itself, as shown in Appendix A. Accordingly, every descriptor focuses on specific language aspects. Pronunciation considers the degree of effect of mother tongue interference, rhythm, stress, and intonation on the clarity of the message. Structure focuses more on the relationship between the grammatical and sentence patterns, and the precision of the delivered message; so, the clearer the message is in terms of structure, the better the level of the candidate will be.

As for vocabulary, the descriptor considers common, concrete, and work-related lexical items. However, covering a wide range of vocabulary will help the test takers' final scores and level. Fluency, as a descriptor in aviation English, refers to the appropriate use of tempo, discourse markers, and connectors with little to no interference with the quality of the information. Following vocabulary, comprehension requires the understanding and successful deciphering of common, concrete, and work-related topics; furthermore, it is highly required to be familiar with and comprehend a range of speech varieties and registers, particularly on the unexpected turn of events, i.e. non-routine situations such as incidents and accidents. Finally, interaction as a skill requires the pilot to maintain a clear and continuous flow of information exchange by checking, clarifying, and confirming data until the communication is satisfactory to both ends. To sum up, it is compulsory for aeronautics professionals while operating in international airspace to make sure they send and receive very clear messages, as it plays a major role in ensuring the flights' safety.

The English for Aviation Language Testing System (EALTS)

The EALTS is a language proficiency test. It is recognized and administered around the world to test native and non-native speakers' language proficiency to fulfill ICAO's linguistic requirements in aviation. The EALTS is:

A multi-level, English for Aviation language testing system designed to assess the language proficiency of commercial flight crew, recreational pilots, and air traffic controllers in the context of aviation and aeronautical communications for ICAO Language Proficiency Requirements compliance. The EALTS measures and reports proficiency in the skills of speaking and listening across all levels of the ICAO Language Proficiency Rating Scale from Pre-Elementary Level 1 to Expert Level 6. (CAA, 2012, p. 2)

It is important to note that the EALTS is not the only aviation English test administered in Algeria. However, after a closer look at different tests, the researchers (author 1) decided to consider the EALTS out of convenience; in other words, as a certified assessor of the test, author 1 had access to data and respondents.

The EALTS consists of a listening and speaking test. Two language skills are the most important in the target situation. The listening test takes up to forty minutes of testing time (individual test), it is a computer-based test in which candidates listen to aeronautical communication recordings and then listen to recorded questions. After that, they choose one of three main options: positive, negative, or not mentioned, depending on the availability of the information in the recordings (conversation/question), i.e., if it exists and it is correct (positive), if it exists and it is incorrect (negative), and if it does not exist at all (not mentioned). The candidates are allowed to take notes at all stages of the tests (both listening and speaking). The speaking test on the other hand (paired test) is divided into three main tasks. Each task must be limited to an average of seven minutes.

In the first task, the interlocutor assessor sits facing the two candidates, the assessor asks simple questions related to common, concrete, and work-related topics such as: What is your job? Where do you work? Where have you had your training? And how many flight hours do you have so far? The second task focuses more on asking for information and confirming existing data. The candidates sit back-to-back and then listen to an indistinct recording of aeronautical communication between aviation professionals who are operational at the time of the communication. The candidates listen to the recording, which is divided into short segments and exchange information, clarifying and confirming to one another what they have heard until they are satisfied with the general and detailed information on the recording. The last task is related to the professional reaction in non-routine situations. The assessor provides each candidate with an unusual scenario and then allows them one minute to prepare their responses. An example of a non-routine scenario is *you are flying from Algiers to London, and while on-cruise you experience a sudden depressurization*. Each candidate reads back the scenario and talks about what s/he should do. In the end, candidate 2 asks or comments about something her/his colleague has said. It is essential to note that the objective of the EALTS test is to assess candidates' language comprehension and performance only, without marking or judging their specialized knowledge in aeronautics or their operations.

In the Algerian context, very scarce attention was allotted to a thorough investigation in the field of aviation English. Apart from the work of Mekkaoui (2013), Mekkaoui and Mouhadjer (2019) who investigated the communicative issues faced by Algerian pilots in their jobs and language proficiency needs of ATCs in Algeria, in addition to Assassi (2017) who suggested a course based on formulaic expressions to reach communicative competence in aviation English courses, no other studies have focused on aviation English so far. As for aviation English tests, no research has been conducted even if, from 2014 until January 2023, three aviation English tests have been administered in the Algerian context, which is the *English Language Test for Aviation* (RELTA) designed by the Royal Melbourne Institute of Technology (RMIT), *English Language Proficiency for Aeronautical Communication* (ELPAC) by EUROCONTROL, and *English for Aviation Language Testing System* (EALTS) administered by Language Testing and Assessment Services (LTAS) and approved by the UK Civil Aviation Authority. These tests were recognized by local authorities in Algeria. Thus far, no research has been conducted to check the validity and reliability of the tests in Algeria. This study is conducted to elicit information from stakeholders and not to criticize the validity or reliability of the EALTS; however, the information we provide may or may not affect the test credibility or offer additional information to avoid malpractices in the testing process and avoid any issues that may jeopardize test takers scores and later communication performance in the target situation.

Research Methodology

As a response to the increasing concern over the validity of high-stakes tests, namely aviation English tests, in addition to the safety-related apprehensions in the aviation industry, the literature review sought to provide a firm foundation for the research design and a rationale for the research questions. The study aimed at providing additional contextual information by answering the following overarching research question First, what is the test-takers perspective on the EALTS test construct? And (2) what is the assessors' perception of the test's procedure and validity?

To answer the research questions, we opted for a qualitative approach within the framework of a case study. As a descriptive study design, we aim at eliciting data of a qualitative nature to understand respondents' perspectives and opinions to pinpoint problems they face before and after sitting for the pilots' English proficiency test. This

will help future research suggest solutions and recommend practices that may diminish the negative impact of these issues in different aviation English tests.

The seventeen participants in this investigation are (i) ten licensed pilots who work for the two national companies in Algeria (Air Algérie and Tassili Airlines), two of them are captains and eight first officers; more details provided below in the analysis section; (ii) three aviation English teachers who have been trained and certified as assessors of aviation English by EALTS experts (Civil Aviation Authority in the UK); (iii) and four university professors from Biskra University in Algeria specializing in applied linguistics, language teaching, and assessment. Participants for this study were chosen out of convenience. The pilots had a testing session together, and all the assessors were present on the day of the test. The pilots come from different companies and with different flight hours, i.e., flight experience, which gave more saturation to the data collected. The assessors have been operational for three years with the testing center and the aviation academy, which gave them much to share about the test construct, procedure, and objective. The professors'/researchers' perspective as experts in assessment, language teaching, and ESP is essential to put high-stakes tests such as the EALTS under investigation and provide feedback on the test's validity and reliability concerning the target situation.

The study outcomes are based on the responses collected through a semi-structured questionnaire designed for teachers and assessors to have a specialized and experienced view of the issues under investigation. A semi-structured interview is selected for candidates (pilots) to give them more freedom to respond comfortably and to ensure the quality of the obtained data. Finally, as the researcher is an assessor himself, a participant observation process based on rubrics designed in an observation grid is chosen for the sake of checking the validity of the responses and the reactions and facial expressions of the candidates as the latter plays a major role in the display of emotions and attitudes. The observation procedure is for confirmation purposes only; no new data was collected nor reported further through observation. The observation process was launched after the administration of the questionnaire and interview so we could cross-check obtained information based on codes and themes we already collected and analyzed. Qualitative data were analyzed using codes and then shaped into recurrent themes based on the preset codes from research questions.

The Study

Test-takers Perspectives

Interviews are seen as rich data and source provider. We can observe this data collection tool from two angles, the researchers' and instruments. angles First, it provides the researcher with in-depth information that might not be elicited by other research instruments. Second, the researcher, as an interviewer and most notably in a face-to-face situation, can notice much more and can provide a contextual and an emotional basis to interviewees' responses. Duff (2008, p. 134) clarified "Interviews are one of the richest sources of data in a case study and usually the most important type of data to be collected. Interviews provide the researcher with information from various perspectives." In other words, interviews provide resourceful information, especially in discussing controversial issues in education. The researchers designed a semi-structured interview for ten (10) pilots working for both national aviation companies. All these pilots must pass the test so they can remain operational in their positions. The interview, in this case, is chosen mainly because of its flexible and rich nature, particularly when dealing with a specialized sample (pilots). The questions are designed to check test-takers' background information, such as professional and linguistic backgrounds, in addition to their perspective on the EALTS test construct, process, and compliance with the ICAO linguistic requirements. Also, we tried eliciting information from the participants before and after the test to give them the chance to share ideas in a more organized manner. In a nutshell, we tried answering the following question: what is the test-takers perspective on the EALTS test construct?

Before the Test

According to the researcher's practices as an assessor, we have noticed that the majority of pilots in this country working for both national companies are males. Nine out of ten pilots were males. These numbers are closer to the ones shared by the FAA's Aeronautical Center (December 31st, 2020 data), indicating that only six percent of U.S. pilots are females.

Another question is designed to have a clearer idea of our respondents' age, rank, and flight hours since the beginning of their professional careers. It is worth noting that the candidates excluded flight hours during their training. The data are summarized below.

Table 2.

Candidates' Age, Rank, and Flight Hours (Hs)

Age Range	25 - 34	35 - 44	Older than 45
Number of Candidates	06/10	02/10	02/10
Rank	First Officer		Captain
Number of Candidates	08/10		02/10
Flight Hours Range	4000 - 5999 hours	6000 - 7999 hours	More than 8000 hours
Number of Candidates	05/10	02/10	03/10

The current English language level of respondents who are directly involved in the research process plays a major role in the identification and analysis of their linguistic needs, or as labeled in ESP, needs analysis. The following table shows the collected responses from pilots concerning their language levels.

Table 3.

Candidates' English Language Level

Level	Respondents /10
Beginner	02/10
Intermediate	06/10
Upper-Intermediate	02/10
Advanced	00/10

Most of our candidates chose "intermediate" as their current language level (six out of ten). on the other hand, only two pilots labeled their level in English as a beginner and two as upper intermediate, while none of them believe that they have advanced foreign language proficiency.

By questioning current English language courses taken by test-takers in our study, we wanted to have a general idea about the efforts our candidates make to learn the language given its importance for their professional careers. Five (5) pilots responded negatively, whereas 2 of them were taking private courses in local schools, but they were not satisfied yet they claimed they attend regularly. Three pilots, of a younger age, explained that they prefer to be self-taught using mobile phone applications and educational websites.

All the respondents are aware of the ICAO linguistic requirements and what is necessary to ensure a smooth transition of information and avoid any breakdown of communication. We noticed during the interview and after their responses that communication is crucial to ensure flight safety.

As far as standardized and high-stakes tests are concerned, the responses varied from one respondent to another. What is common between the pilots' careers are the main standardized tests taken in this country such as the middle and secondary school standardized tests. However, two pilots took other tests as they used to work with law enforcement (national gendarmerie) as helicopter pilots. Three respondents took standardized tests abroad during their training as they claimed. The researcher had to elaborate on the question to help respondents sort standardized tests from other types of tests.

Passing a specialized aviation English language test is imperative for their careers and mainly to be operational with their companies. All the pilots sat for tests before, RELTA (RMIT English Language Test for Aviation) is provided by the Royal Melbourne Institute of Technology. Experienced pilots sat for the test multiple times. Six pilots said that they had failed the test recently, and they decided to try with a different test provider. As for the main reasons behind failing the test, respondents mentioned that in addition to lack of preparation and anxiety, RELTA assessors considered specialized aviation knowledge besides language knowledge in scoring, which made it more challenging for test-takers. Respondents' answers showed that they were tested by subject specialists rather than language specialists.

Apart from the brief elaboration the assessors gave concerning the testing process, candidates had no previous knowledge about the test construct, its phases, or the testing methods. The candidates who sat for the RELTA test before had a closer idea of what was expected from them in the EALTS test. Consequently, the respondents answered negatively about taking the EALTS test before.

After the Test

The candidates had similar responses as far as the testing process and content are concerned. They believed they encountered some difficulties within the listening test, more precisely, with comprehension. Two of the candidates faced more difficulties as

they declared they felt very anxious, and it affected their performance in the speaking test. The two most experienced pilots (Table 3) said that they faced technical issues with the computers as they could not have a solid grip over the mouse and the listening devices and that impeded keeping up with the questions and suggestions in addition to the short time allotted to answering.

The candidates (10/10) declared that anxiety, shyness, and fear of committing mistakes, which are related to both language and specialty content mistakes, had a considerable negative impact on their language production and performance. Cushing (1994) classified pilot/ATCO communicative problems into language-based and problems not based on language. This indicated that not only language problems caused a breakdown of communication, but other issues as well may cause misunderstanding. This reflects the test as well because such ESP high-stakes tests for licensure are designed to test language in real-life situations (target situation). For him, the communication problems based on language that he identified were problems of language (e.g., ambiguity, homophony, intonation), problems of reference (e.g., uncertain reference, uncertain addressee), problems of inference (e.g., implicit inference, lexical inference, unfamiliar terminology, false assumptions), and problems involving repetition (e.g., partial readbacks). The communication problems not based on language that he lists are problems with numbers, problems with radio, problems of compliance, and other general problems. Therefore, the communication problems might be caused by confusion, stress, and anxiety, which may affect negatively language production and cause a breakdown of communication.

As for preparation, candidates agreed on the importance of preparation and three of them focused more on preparation and how sitting for such a high-stakes test requires both mental and linguistic preparation. As stated earlier in this section, more experienced candidates (Table 3) faced technical problems related to audiovisual aids during the listening test. One pilot declared “even if the assessors [helped us install] all devices, we faced some issues with these devices and [keeping up with] the fast pace of the recordings and the limited timing given for answering questions”.

The candidates also focused on the issue of the listening test. They stated that there is a lack of concordance between the test contents and the target situation requirements. As an example, none of the recordings were in any regional accent; also,

the recordings were all in standard phraseology (orders, requests, advice, permissions, approvals), which does not comply with real-life situations in aviation communication. In their day-to-day operations, candidates faced several instances in which plain English, defined by ICAO (2010, p. x) as "the spontaneous, creative and non-coded use of a given natural language" was the main form of aviation talk used during most of the conversation. Plain English, in this case, requires much more time and cognitive effort from the candidates to decode meaning, which is not recommended in aviation talk and may cause safety concerns, especially in busy airspace.

Another issue the candidates faced is the clarity of recordings. Although clear communication is not always guaranteed in radiotelephony, candidates believed that in addition to not being allowed to listen to recordings more than once, the recording quality, timing, and confusing answer options affected their answers negatively. Field (2019, p. 1) argued, "testing second language listening proficiency validly and reliably has always posed a challenge." The challenges faced by test developers when writing test specifications for listening tests include deciding whether the candidates will be allowed to listen to the text more than once (Taylor; Geranpayeh, 2011), issues related to task authenticity (Brindley, 1998; Lynch; Mendelsohn, 2010; Wagner, 2014) and to memory (Wu, 1998). The candidates declared that both native speakers and proficient non-natives usually use plain English in routine situations, which complicates communication for them. Subsequently, such scenarios must be a part of the listening test, given the importance of these situations in avoiding a breakdown of communication. Concerning the cognitive requirements of the listening test, the options provided for the test-takers after listening to the recordings "*positive, negative, and not mentioned*" were mentioned repeatedly in their responses. For them, it was quite confusing to choose *negative* or *not mentioned* to answer if the information in the question was incorrect (e.g. recording: phase of flight "take off"; the question: phase of flight "landing" = **negative**) or the information is not mentioned (e.g. recording: altitude 35,000 feet; the question: altitude not mentioned = **not mentioned**). The candidates still found it confusing even if the assessors clarified the options before the test and they were allowed to take notes given the very limited time they must answer at the end of each recording. Consequently, it is imperative to consider the complexities of cognitive processes required in listening so that "the cognitive processing activated in the test taker by a test task corresponds as closely as possible to what they would expect to do in the (...) listening context" (Taylor

and Geranpayeh, 2011, p. 96). Finally, on a question related to the assessors' performance, nine pilots added that there were no problems understanding or interacting with assessors, while one candidate said that he sometimes could not follow what was said due to the assessor's fast speaking pace.

As for test-takers' suggestions, three respondents shed more light on the mental discomfort they faced and proposed more preparation, more time, and a practical side of the preparation in the form of tryouts or a mock test. Four candidates out of ten objected to the random selection of the pairs who take the test together, and they said they prefer to sit for the test with someone they know to feel more at ease and comfortable. Three candidates showed their concern regarding the listening test and the unclear audio materials (recordings). Concerning the audiovisual aids used during the listening test, two candidates proposed a tentative test on the same computers as the official test so they can get used to the devices (mouse and headset). Finally, the second phase of the speaking test, in which candidates sit back-to-back, caused some comprehension issues for six candidates who did not quite understand the process, even if they answered positively during the test protocol elaboration by the assessors before the test.

As far as developing their communicative abilities on radiotelephony, candidates suggest practicing in English during all their flights. One issue we have noticed is that Algerian pilots tend to use French during their national flights or during flights to destinations where French is an operational language. ICAO recommends sorting to another shared language only when phraseology does not serve the communicative purpose or during the unexpected turn of events when the said language, French in this case, serves as a last resort to solve a communicative issue. However, the use of the English language during all flights, national and international, is highly recommended by not only pilots but teachers, assessors, and supervisors as well.

Our respondents shared their concerns related to the use of plain English, mostly by native and proficient non-native speakers of English. The use of standard phraseology in routine situations is a must, and respondents added that proficient language users must "*meet us halfway*" to avoid any breakdown of communication. This specific point has been overlooked in several cases. However, Estival et al. (2016, p. 199) claimed that "pilots who are native English speakers commit, in some cases, as many communication errors as English as a second language pilots". Read and Knoch (2009) argued that the

ICAO LPRs have given “native-speaking aviation personnel no incentive to develop their communicative competence in ELF [English as a lingua franca] terms” (p. 217). This is confirmation that even after these two studies as an example, the problem persists and there is still much to be discussed about this phenomenon.

Five candidates out of ten felt confident and reassured of their success, and the other five candidates felt skeptical, mostly because of the anxiety and hesitation they showed during their performance on the speaking test. Two out of the last five candidates added that their answers during the speaking test were not convincing or accurate, and parts of them were wrong.

Assessors’ and Teachers’ Perspectives

Given its reliability and how much time and effort it saves, the questionnaire is a frequently selected data collection tool in social and human sciences, more precisely teaching English as a foreign language. It is crucial to realize the importance of a well-designed questionnaire and its effect on the quality of the responses the researcher collects (Dornyei, 2003). Thus, we have chosen the questions very carefully, taking into account important aspects of our study, the research questions, and the coherence of the rubrics. The questionnaire was administered to three certified assessors (EALTS) and four EFL (English as A Foreign Language) teachers and researchers at Biskra University in Algeria. We believe that assessors and specialized teachers are central to the detection and mitigation of learning issues and malpractices. The questions investigate respondents' background and experience, then tackle very specific aspects related to the test and the respondents' perspectives on the test validity, compliance, and reliability concerning test-takers performance issues. Consequently, through these inquiries, we attempted to answer the following question: what is the assessors' perception of the test's procedure and validity?

The first question is designed to collect data related to respondents' educational background and current profession. The first question is associated with the second, which discusses respondents' experience. Both questions are highly important for the study to have a clear idea of teachers' and assessors' experience and practices vis-à-vis assessment and evaluation in general, and high-stakes tests in particular. The majority of respondents are university professors with different ranks from assistant professor “B” (either not yet enrolled in a doctoral project or currently in one for less than three years),

to assistant professor "A" (three years or more working on a doctorate project). Additionally, two of the university professors already have a doctorate (associate professor B). Both the Private Academy Assessors hold a bachelor's degree in language sciences.

The second question, as stated above, is related to the first one. The purpose is to identify the experience our respondents have in teaching English as a foreign language and the specialty they graduated studying or researching currently.

Table 4.

Respondents' Specialty and Experience

Respondent	Specialty	Experience
01 (assessor)	Didactics	15 years
02 (assessor)	Didactics	3 years
03	Applied Linguistics	5 years
04	Didactics	9 years
05	Linguistics / Phonetics	6 years
06	Applied Linguistics	6 years
07	Applied Linguistics and Assessment	19 years

The data in the table shows the similarity respondents have with their main area of specialty (Applied Linguistics - Didactics). As far as experience is concerned, there is a vast difference between the least and most experienced (03 years - 19 years).

All respondents sat for standardized and high-stakes tests. There are similar tests they sat for, such as the elementary, middle, and secondary school final exams; and there are different tests in the form of general English tests such as the IELTS and TOEFL (five respondents). Four teachers claimed that they have taken such types of tests when they were filing for a teaching position either in middle or secondary school. This showed that our respondents were in a good position to share experiences and answer our questions.

The two Aviation English assessors had similar responses to what they noticed before the test. They both noticed very stressed candidates before the test. Throughout the large number of candidates, they have tested in their careers, they claimed that they have witnessed different degrees of nervousness and anxiety. However, assessors said that the candidates felt less stressed and noticeably relieved after the test.

All respondents shared similar perspectives concerning high-stakes tests causing mental discomfort that is manifested through learners' behaviors, both verbal and physical reactions. One of the respondents asserted that it is quite rare to find confident and calm test takers even if the assessor knows that they are high achievers. Another respondent pointed out that even if the age factor is different between EGP and ESP test takers, the reactions and behaviors before and during the test seem similar.

The respondents indicated that the disrupted mental state of any test takers affects their performance, and it is inevitable to feel anxious and stressed, even minimally, particularly during high-stakes tests. The reasons behind taking the test and its value, the testing physical environment, the assessors' and invigilators' behavior, the content, and the preparation for the test, were all different issues that test takers faced and that affected their performances according to what the respondents claimed.

One issue that assessors mentioned is related to the grading process. The test requires two assessors, one as an interlocutor assessor and the other as an observer assessor, whose main duty is to grade the subskills according to the ICAO scales. The assessor observer keeps the scale in front of him/her in addition to a grading sheet. The assessors faced several issues related to timing and overlap in scales. As for timing, assessors believe that extra focus is required, especially when test-takers do not speak sufficiently in the speaking test, which prevents assessors from evaluating language produced fairly. Overlap between subskills was one of the issues faced by assessors. Usually, they find it very challenging to pinpoint the communicative problem and which of the six descriptors they should relate the problem to, as most of them at a specific level seem to refer to very similar language aspects. Moere et al., (2009) discussed a similar issue with the Versant Aviation English Test regarding ICAO descriptors in level 4 which include:

- Comprehension, '(when confronted with) an unexpected turn of events, comprehension may be slower or require clarification strategies';
- Interactions, 'maintains exchanges even when dealing with an unexpected turn of events. Deals adequately with apparent misunderstanding by checking, confirming, or clarifying';
- Vocabulary, 'can often paraphrase successfully when lacking vocabulary in unusual or unexpected circumstances.

This, according to our respondents, created a particular challenge in separating subskills for assessment. Accordingly, the highly required language functions in aviation talk: comprehending, clarifying and paraphrasing appear to be related to all three descriptors, which created an operational difficulty in the isolation of the six subskills in spoken language proficiency for the non-compensatory composite scoring according to the same authors above. As a result, the validity, reliability, and fairness of the scoring and the test as a whole might be compromised.

Another issue noted by the assessors and related to timing discussed earlier is the difficulty faced in assessing each one of the six measures (language descriptors) separately. Either distributing the measures on the three speaking tasks, allot specific timing for each descriptor, or jumping from one measure to the other across tasks and throughout all the speaking test tasks, none of these techniques seemed reliable as raters might conflate fluency with pronunciation or well-structured sentences (grammar) with accurate use of vocabulary. This might lead to at least one of the subskills not being assessed appropriately (McNamara, 1996; Orr, 2002).

The test is designed for licensure, which makes it a high-stakes test and puts raters on the spot to ensure fair assessment and even hold responsibility for flight safety. Much effort is invested by assessors to avoid any overlap in the scoring of the subskills because the final scoring level, according to the ICAO, is the one subskill in which the candidate is least proficient. The different understanding and interpretation of descriptors and their requirements at each level are added to the heavy cognitive load of the assessors. These exhibited variations may compromise the validity of the scoring system and the reliability of the test.

Finally, it is worth noting that the interlocutor assessor is also responsible for the rating at the end of the test. However, it is very challenging to keep track of ratings while interacting with test-takers, and to grade the subskills of both candidates at the end of the test does not seem feasible since the assessor may overlook the specificities of candidates' performance that affect their final scoring negatively.

Specialized professors stated that assessors' duties are not only related to questioning or invigilating, so they advised creating a friendly and comfortable testing

environment and taking into account the candidates' stressful nature during these educational events. However, respondents affirmed that this task is easier said than done, impediments such as chaotic settings, and lack of equipment in addition to many test takers cannot help them pay attention to comforting and stabilizing the test takers' mental state. The two ESP assessors confirmed what is stated above except for the environment, which as they claimed, "in ESP tests like these, we did not face any of the issues related to the environment, large numbers of candidates and equipment".

The Aviation English assessors focused more on interaction and clarification with their candidates to create a friendly and comfortable testing environment. Alternatively, specialized professors said that there is a difference between ESP and EGP tests, most notably language proficiency and high-stakes tests in ESP, as they must pay more attention to test designs, content, and testing methods. Afterward, centering the focus on the mental state of the candidates before and during the test. To sum up, the focus of our respondents was testing preparation. Assessors said that even candidates who took the test before still faced technical and comprehension issues. Thus, a preparation session with a mock test simulating the real test with all its tasks and procedures can help test-takers and enhance the test's practicality, impact, and validity. They believed that this would help test-takers manage time more efficiently and provide them with more practical techniques on how to manage the breakdown of communication, especially with time limitations in aeronautical talk. The test preparation can also help candidates be accustomed to the test construct and other details, such as the listening test recordings that lack regional accents. This ultimately puts the test itself in representativeness-related issues, which jeopardized the test's validity and compliance with the ICAO and target situation linguistic requirements.

Conclusions and Summary of the Main Results

The accuracy and naturalness of the collected data play a major role in ensuring the validity and reliability of the research outcomes. It is important to note that the research outcomes the researcher obtained are not to be generalized nor altered for the sake of studying a different sample. The researcher in this paper tried to be as objective as required by limiting the observation process outcomes he opted for as a third data collection tool to be only for cross-checking obtained information from questionnaires

and interviews. As he is a specialized professional in teaching English as a foreign language and an Aviation English assessor as well, he relied only on data collected from other colleagues both assessors and professors in addition to our main focus, which is the aeronautics professionals (pilots).

The results show an immense focus on the mental side of testing, most importantly before and during the test. Candidates feel mostly nervous and stressed before the test due to lack of preparation, and unfamiliarity with the test content or testing methods, in addition to the value and significance of the test, bearing in mind that succeeding in this test is a requirement to become or remain operational as a pilot. The candidates mentioned the prior preparation and sitting for the same version of the test as a tryout before sitting for the official one. It is important to know that the test provider of the EALTS does not allow such a technique and they provide documents and a video to help candidates have the clearest idea about the test, its stages, its components, and its requirements. As stated within the background of the study, the EALTS does not require the accuracy and correctness of the specialized information from candidates. It is a test designed for testing candidates' language comprehension and production only. This is explained to all candidates before every test.

As for specialized teachers, they are all for creating a comfortable testing environment and paying attention to the test takers' mental state, given its immense impact on their performance and test results. As a result, it is crucial to rethink the stage before sitting for the test and help learners be more familiarized with the EALTS construct through a mock test or a tryout so they can apply what they have taken as information from the assessors or see in the EALTS demonstration video. This way, both the assessors and the test takers will face fewer issues. Most importantly, the technical difficulties faced by more experienced pilots with computers, and other issues we noted such as comprehension hitches and breakdown of communication in the second task of the speaking test, are caused by unfamiliarity with test tasks. These concerns can be overcome through the application of a mock test. After all, successful communication is a vital matter for the safety of any flight, and a breakdown of communication is never to be ignored or underestimated, starting from the Aviation English test itself. Thus, intensive training and formal and informal individual or collective meetings, in addition to advice

after the test, will help candidates mentally, which positively affects their test performance, and language learning; and eventually ensures flight safety.

Reconsidering the contents of the test in terms of compliance with ICAO linguistic requirements and target situation language use took their fair share of our respondents' reflections. The main issue discussed is the relatively noticeable difference between what test-takers are assessed for and what they experience during their flight operations. As an example, real-life situation misunderstanding, and regional accents are two important occurrences in everyday flight operations; yet test-takers believed that the test excluded such real-life communicative situations.

Another issue discussed by assessors was related to testing usefulness and fairness. The scoring process during the test requires much more effort than expected from a criterion-referenced scoring procedure. Assessors claimed that objectivity seemed to be compromised at times. The main problem was discussed based on ICAO's description of the six holistic descriptors. The interpretation of the descriptions might be different, and there was an overlap between the descriptors themselves. Subsequently, assessment in ESP contexts is very demanding, and it calls for rigorous testing policies and processes for all stakeholders. Monteiro (2022) asserts that "a clearer definition of the aeronautical RT construct is of utmost importance, one that is aligned with current views of language use, with the multiple factors that impact RT communication, and with stakeholders' perspectives". (p. 225)

To sum up, our study calls for robust testing validation studies, not to state that the EALTS failed validity, but to elicit and exchange data from the main stakeholders, i.e., test-takers. After all, the test is designed for them and their perspective is of high importance to maintain why not enhance the communicative abilities of pilots and ATCs around the world, which leads to a contribution to flight safety. To improve test usefulness and credibility, test providers, researchers, and test administrators must react to test-takers' and assessors' feedback. The EALTS is an interesting test that is well-constructed and provides fair opportunities to candidates from different linguistic and cultural backgrounds. However, the points discussed in this study and the outcomes shared in the results section can provide informed decisions to test providers and our colleagues around the world. The test ensures reliability as far as the scoring system is

concerned through inter-rater reliability as an external consistency of scores. The interpretation of descriptors though can cause measurement errors and alter the scoring outcome, then jeopardize the scoring fairness. As for test validity, no communication-caused incidents or accidents by tested pilots have been recorded up until January 2023, which makes it clear that the testing and licensure process was fair and valid. Additionally, this shows that the knowledge, skills, and abilities tested are reliable measures of construct. However, an update of the test tasks and contents to mirror target language use is highly recommended, especially when it comes to non-routine situations and different accents. The test's impact on assessors and, more importantly, test-takers showed different perspectives. Viewpoints differed between satisfaction and high hopes for passing the test to the uncertainty of success and test fairness. Accordingly, test construct and administration must take into account test-takers' perspectives to enhance test quality.

On a different note, test practicality was somehow criticized in terms of timing and practicality of the computer-based listening test. Respondents faced issues with the listening test procedure as far as technical and operationalization issues are concerned. Besides, human, material, and time resources invested in the EALTS test show motivating procedures. Test authenticity was challenged by the absence of regional accents and real-life communicative concerns in the recordings, for instance; rather, the test draws an interesting correspondence between features of the test tasks and features of real-world language use tasks. Finally, the interactive aspect of the test is of high importance and the test under investigation showed interesting signs, as we have noticed through the test-takers' engagement. Subsequently, as motivating topical and language knowledge tasks are for candidates, the affective schemata of the test in the form of stress and sometimes severe anxiety noticed by assessors during the test is critical to creating a more comfortable testing context and an engaging test preparation phase. This idea will help test-takers interact more efficiently and comfortably with test tasks, which will reflect their real language proficiency.

We do encourage researchers to consider a deeper understanding and investigation of proficiency tests, especially high-stakes tests in ESP because of their economic value and more importantly, safety concerns in the workplace.

References

- Assassi, T. (2017). *The Effect of Formulaic Language on ESP Learners' Conversation Competence: The Case of Aviation Professionals in Algeria*. [Unpublished Doctoral Dissertation]. Tlemcen University, Algeria.
- Assassi, T. (2020). Algerian Aeronautics Professionals' Needs Taxonomy: A Target Situation Needs Analysis. *International Journal of English Linguistics*, 10, 90-103. <https://doi.org/10.5539/ijel.v10n6p90>
- Assassi, T., & Benyelles, R. (2016). Formulaic Language for Improving Communicative Competence. *Arab World English Journal*, 7, 163-176. <https://doi.org/10.24093/awej/vol7no2.11>
- Borowska, A. P. (2018). A Multilingual Speaker in Global Aviation Communication. *Applied Linguistics Papers*, 25(2), 11-19. <https://doi.org/10.32612/uw.25449354.2018.2.pp.11-19>
- Brindley, G. (1998). Assessing listening abilities. *Annual Review of Applied Linguistics*, 18(1), 171-191. <https://doi.org/10.1017/S0267190500003536>
- Brown, A. (1993). The role of test-taker feedback in the test development process: Test-takers' reactions to a tape-mediated test of proficiency in spoken Japanese. *Language Testing*, 10(3), 277-301. <https://doi.org/10.1177/026553229301000305>
- Civil Aviation Authority (CAA International). (2012). *English for Aviation Examinations. Handbook 2012*. <https://sciarium.com/file/283262>
- Civil Aviation Authority. (2016). CAP 413: Radiotelephony Manual. Intelligence, Strategy and Policy, Safety and Airspace Regulation Group. Civil Aviation Authority.
- Cookson, S. (2009). Zagreb and Tenerife: Airline accidents involving linguistic factors. *Australian Review of Applied Linguistics*, 32(3), 22.1-22.14. <https://doi.org/10.2104/arat0922>
- Cox, A., & Karimi, M. (2022). Creating a rubric for placement tests for Aviation English programs. In P. Tosqui-Lucks & J.C. Santana (Eds.), *Aviation English a Global Perspective: Analysis, teaching, assessment* (pp. 182-199). Bookerfield. <https://doi.org/10.53268/BKF22080510>
- Cushing, S. (1994). *Fatal words: Communication clashes and aircraft crashes*. University of Chicago Press.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of language testing*. Cambridge University Press.
- Denscombe, M. (2014). *The Good Research Guide: for small-scale social research projects*. McGraw-Hill Education.

- Dornyei, Z. (2003). *Questionnaires in second language research: Construction, Administration, and Processing*. Lawrence Erlbaum Associates.
- Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge University Press.
- Douglas, D. (2005). Testing languages for specific purposes. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 857-868). Lawrence Erlbaum.
- Douglas, D., & Myers, R. (2000). Assessing the communication skills of veterinary students: Whose criteria? In A. J. Kunnan (Ed.), *Fairness and validation in language assessment: Selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida*. Cambridge University Press.
- Duff, P. (2008). *Case Study Research in Applied Linguistics*. Taylor & Francis.
- Elder, C. (1993). How do subject specialists construe classroom language proficiency? *Language Testing*, 10(3), 235-254.
<https://doi.org/10.1177/026553229301000303>
- Federal Aviation Administration's Aeronautical Center (FAA). (2021, December 31). *U.S. Civil Airmen statistics*.
https://www.faa.gov/data_research/aviation_data_statistics/civil_airmen_statistics
- Field, J. (2019). *Rethinking the Second Language Listening Test - From Theory to Practice*. Equinox.
- Field, J. (2022, July 6). *Idle chatter: What really goes on in tests of interactive communication? CRELLA Symposium, 2020*. [Video]. YouTube.
<https://www.youtube.com/watch?v=ONM7xcJTPD0>
- Fowler, R., Matthews, E., Lynch, J., & Roberts, J., (2021). Aviation English Assessment and Training. *Collegiate Aviation Review International*, 39(2), 26-42.
<https://doi.org/10.22488/okstate.22.100231>
- Hymes, D. (1972). On Communicative Competence. In J.B. Pride and J. Holmes (Eds.), *Sociolinguistics. Selected Readings* (pp. 269-293). Penguin.
- International Civil Aviation Organization (ICAO). (2004). *Manual on the Implementation of ICAO Language Proficiency Requirements (Doc. 9835)*. International Civil Aviation Organization.
- International Civil Aviation Organization (ICAO). (2006). *ICAO Manual of Radiotelephony. Doc 9432-AN/925. Annex 10, Volume II, Chapter 5*. International Civil Aviation Organization.
- International Civil Aviation Organization (ICAO). (2010a). GBAS CAT II/III development baseline SARPs. International Civil Aviation Organization.

- International Civil Aviation Organization (ICAO). (2010b). *Manual on the Implementation of ICAO Language Proficiency Requirements* (2nd ed.) (Doc. 9835). International Civil Aviation Organization.
- International Civil Aviation Organization (ICAO). (2013). Language Proficiency Requirements: Critical to aviation safety. *ICAO Journal*, 5, 64-64.
- Jacoby, S., & McNamara, T. (1999). Locating competence. *English for Specific Purposes*, 18(3), 213-241. [https://doi.org/10.1016/S0889-4906\(97\)00053-7](https://doi.org/10.1016/S0889-4906(97)00053-7)
- Katsarska, V. (2021). Formulaic Language in Aeronautical English-Friend or Foe. *Journal of Teaching English for Specific and Academic Purposes*, 9(2), 239-245. <https://doi.org/10.22190/JTESAP2102239K>
- Lumley, T., & Brown, A. (1996). Specific-purpose language performance tests: Task and interaction. *Australian Review of Applied Linguistics, Series S*, 13, 105-136. <https://doi.org/10.1075/aralss.13.06lum>
- Lynch, T., Mendelshon, D. (2010). Listening. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd ed) (pp. 180-196). Hodder Arnold.
- Mekkaoui, G. (2013). *English for Aviation in the Algerian context: The case of Pilots and Air Traffic Controllers in Zenata-Messalli El Hadj Airport*. [Unpublished master's thesis]. Tlemcen University, Algeria.
- Mekkaoui, G., & Mouhadjer, N. (2019). Addressing air traffic controllers' English language proficiency needs: Case of Zenata Airport. *Global Journal of Foreign Language Teaching*, 9(3), 167-183. <https://doi.org/10.18844/giflt.v9i3.4245>
- Moere, A.V., Suzuki, M., Downey, R., & Cheng, J. (2009). Implementing ICAO Language Proficiency Requirements in the Versant Aviation English Test. *Australian Review of Applied Linguistics*, 32, 27.1-27.17. <https://doi.org/10.1075/aral.32.3.07van>
- Monteiro, A.L.T. (2022). The Assessment of English in Aeronautical Radiotelephony Communications: A mixed methods study. In P. Tosqui-Lucks & J.C. Santana (Eds.), *Aviation English a Global Perspective: Analysis, teaching, assessment* (pp. 224-245). <https://doi.org/10.53268/BKF22080512>
- Netherlands Aviation Safety Board (1978). *Final Report and Comments of the Netherlands Aviation Safety Board of the Investigation into the Accident with the Collision of KLM Flight 4805 and Pan American Flight 1736*. <https://www.faasafety.gov/files/gslac/courses/content/232/1081/finaldutchreport.pdf>
- Taylor, L., Geranpayeh, A. (2011). Assessing listening for academic purposes: Defining and operationalising the test construct. *Journal of English for Academic Purposes*, 10(2), 89-101. <https://doi.org/10.1016/j.jeap.2011.03.002>

- Wagner, E. (2014). Assessing listening. In A. Kunnan (Ed.), *Companion to language assessment Volume 1* (pp. 47-63). Wiley-Blackwell.
<https://doi.org/10.1002/9781118411360.wbcla094>
- Wu, Y. (1998). What do tests of listening comprehension test? - A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing*, 15, 21-44.
<https://doi.org/10.1177/026553229801500102>
- Zhou, Y., Yoshitomi, A. (2019). Test-taker perception of and test performance on computer-delivered speaking tests: the mediational role of test-taking motivation. *Lang Test Asia* 9(10), 1-19. <https://doi.org/10.1186/s40468-019-0086-7>

Reviewers:

1. Anonymous
2. Anonymous

Handling Editor:

Boris Naimushin, PhD,
New Bulgarian University