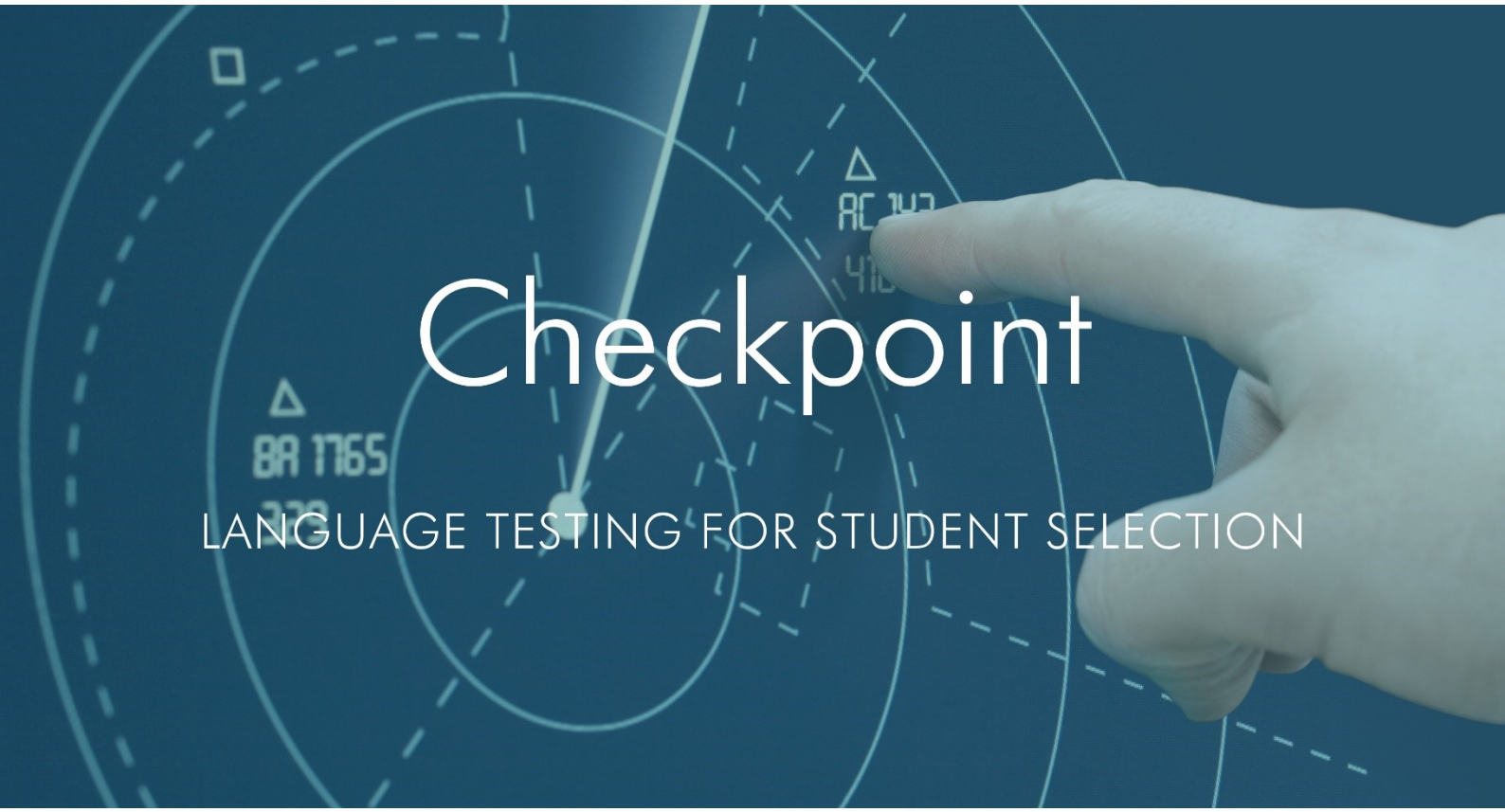
Latitude
Aviation English Services



Checkpoint

LANGUAGE TESTING FOR STUDENT SELECTION

Test development and maintenance

May 2017

Contents

1. Introduction

Checkpoint is an English Language Proficiency (ELP) test designed to measure the English language skills required for successful English-medium aviation training. Checkpoint is owned and operated by Latitude Aviation English Services Limited (UK).

At Latitude, we are committed to providing quality language training and testing products and services, and to helping our stakeholders to understand and use them. This document describes the development and maintenance of Checkpoint. The document is intended primarily to help aviation training decision-makers and admissions officers decide if Checkpoint meets their language testing requirements, but it may be of interest to other stakeholders in aviation training such students, student sponsors, English language instructors, aviation assessors and training managers.

2. Development rationale

Rapid growth in the aviation industry has created a demand for trained pilots and Air Traffic Control Officers (ATCOs). The majority of students entering aviation training around the world do not have English as a first language and yet much aviation training is conducted in the medium of English. As aviation training is a costly undertaking, students, student sponsors and ATOs that provide aviation training in the medium of English need to be confident that inadequate language proficiency will not impede training success. Therefore, language testing for student selection is a high-stakes activity that requires valid and reliable language assessment in order to reduce risk for all stakeholders.

During conversations with our customers and potential customers, we became aware of dissatisfaction with existing general, academic and aviation language tests used for the purposes of student selection. Customers reported that such tests:

- May produce scores which do not equate with the levels of ELP presented by students on arrival for training;
- May not measure the language skills that students require for successful aviation training;
- May not reflect the subject matter and language use settings that students encounter in aviation training; and/or
- May be expensive, inflexible and/or inconvenient.

3. Customer needs analysis

In early 2013 we conducted a survey of some of the general, academic and aviation language tests available and concluded that:

- ICAO language tests for personnel licensing are inappropriate for student selection because:
  - They are designed for licensed professionals, not for students;
  - They do not address reading which is a skill crucial for successful aviation training; and
  - They assume professional knowledge of RT and aviation operations that students do not have.
- General and academic English tests are also inappropriate for student selection because:
  - They do not test in the context of aviation training and so results are less valid; and
  - They test content and skills which are irrelevant to the needs of the students.

Following these conclusions, we presented a questionnaire on language testing for student selection to 16 training managers and language instructors working for ATOs in Bahrain, Cameroon, Canada, Germany, Italy, Jordan, Kazakhstan, Qatar, Russia, Serbia, Turkey and USA. Below are the responses to the questionnaire.

- All respondents reported that their ATO conducted aviation training in the medium of English.
- When asked to rate the importance of the four skills (reading, writing, listening and speaking) in aviation training, respondents answered as follows:

	No so important	Important	Very important	Essential
Reading	6.25%	43.75%	25.00%	25.00%
Writing	37.50%	50.00%	12.50%	0.00%
Listening	0.00%	0.00%	18.75%	81.25%
Speaking	0.00%	6.25%	31.25%	62.50%

- When asked what information they would find useful from test results, respondents answered as follows:

	Not so useful	Useful	Very useful	Essential
Whether a student is ready (or not) for aviation training	12.50%	18.75%	43.75%	25.00%
A student's strength and weakness	6.67%	20.00%	46.67%	26.67%
How much language training a student needs (if any)	12.50%	25.00%	43.75%	18.75%

- When asked whether a professionally-produced online English language test for student selection would be useful for their organisation, respondents answered as follows:

Yes	60.00%
No	6.67%
Maybe	33.33%

Based on the results of the questionnaire, we decided to develop a specific-purpose aviation language test for student selection. Work began in May 2013 and the test was launched for operational use in November 2014.

4. Test development team

4.1 Latitude team

Name	Role	Experience and expertise
Henry Emery	Project manager, lead item writer	➢ Project manager: English Test for Aviation – first ICAO endorsed language test ➢ Project manager: *ICAO Samples Training Aid* (RSSTA) developed by ICAEA with the universities of Lancaster, Cambridge and Melbourne ➢ Co-author, British Council award-winning *Aviation English* (Macmillan, 2008) and *Check Your Aviation English* (Macmillan, 2010)
Captain Dennis Gliddon	Content writer / editor	➢ KLM/Air France, Captain - Fokker 50/100, Embraer 170/190 (35 years' flying) ➢ Cambridge English Language Teaching to Adults (CELTA)
Dr Peter Bradon	Systems / IT / test analysis	➢ Ph.D Psychology of human reasoning ➢ Statistical analyst on item difficulty for the Educational Testing Service ➢ Co-Author, Item Generation for Test Development Routledge, 2002

4.2 External development consultants

Name	Role	Experience and expertise
Stephanie Whitebread	Content writer / editor	➢ Former Heathrow APP ATCO (25 years) ➢ Cambridge English Language Teaching to Adults (CELTA)
Neil Bullock	Content reviewer / editor	➢ Former assistant ATCO (Newcastle) and airside driver trainer ➢ Trinity Certificate in TEFL, Institute of Linguists PGDip, MA modern languages, Ma applied linguistics ➢ Authorised Skyguide ELPAC Examiner ➢ Swiss Federal Office of Civil Aviation authorised language assessor and assessor trainer
Magdalena Večeřová	Content reviewer / editor	➢ MA in TEFL, MA in Language Testing (Lancaster University) ➢ Testing expert: EUROCONTROL, IANS, Luxembourg – ICAO endorsed English Language Proficiency for Aeronautical Communications (ELPAC) - (Classical test theory and item response theory analysis)
Nia Franks	Content reviewer	➢ Captain/line pilot for Queensland Government - Police Air Wing. 35 years aviation experience as PIC in UK, France, Poland, Canada, U.S., South Africa, and Australia.

4.3 External test reviewers

Name	Role	Experience and expertise
Mike Langley	Test reviewer	➢ Royal Air Force pilot ➢ Former Head of Training, Oxford Aviation Academy
Richard Taylor	Test reviewer	➢ Royal Navy Air Traffic Control Officer ➢ Former Head of International Training, National Air Traffic Services (UK)
Nick Mylne	Test reviewer	➢ Former helicopter pilot ➢ Current ground school instructor (Human Performance)

5. Development process

1. Test specification

Specification of:

- Intended test users and test-taker population
- Target listening, reading and speaking abilities (test construct)
- e-assessment platform

2. Skill test and task specifications

- Listening test specification: Contexts an discourse/features of speech/number and order of texts/text content and length/task (type and number of items)
- Reading test specification: Discourse and text sources/number and order of texts/text content and length/task (type and number of items)
- Speaking test specification: Task type/number and order of tasks/task content

3. Drafting of test 'blueprint'

Drafting of:

- A complete set of test content (blueprint) and marking keys for listening and reading texts and items
- Speaking tasks and rating scale

4. 1st moderation committee review

- Review of 'blueprint': Reading and listening texts and items and speaking tasks and rating scale
- Consultants complete test tasks, commenting on abilities required and item quality and difficulty
- Consultants meet to discuss comments and revise tasks

5. Platform

- Research and procurement of e-assessment platform

6. Population of platform

- Population of platform with test content
- Configuration of test structure, timing and sequencing
- Development of specific technical functionalities for Checkpoint in collaboration with platform owner

7. Drafting of version 2 of test content

- Listening and reading texts, items and marking keys
- Speaking tasks

8. 2nd moderation committee review

- Review of version 2: Reading and listening texts and items and speaking tasks and rating scale
- Consultants complete test tasks, commenting on abilities required and item quality and difficulty
- Consultants meet to discuss comments and revise tasks

9. Audio production

- Recording and production of test rubrics
- Recording and production of listening texts using professional voice actors with a variety of native and non-native speaker accents

10. Candidate test familiarisation

- Production of sample of listening, reading and speaking tasks
- Production of 9 familiarisation videos targeted at non-native speakers with low language proficiency
- Placement of video material on the Latitude website

11. Test trials

- Tests are administered to a trial population consisting of:
  - Educated, adult native and non-native speakers including aviation and non-aviation specialists; and
  - Native and non-native speaking pre-sessional and undergraduate students at the Faculties of Business and Science and Engineering at the University of Plymouth
- Item response data analysed for item facility values and item reliability
- Qualitative feedback collected

12. External aviation SME review

- SMEs sit the test as candidates
- SMEs and project manager meet for guided review of all test content, tasks and items, making judgement on task quality and difficulty and thresholds for the traffic light scoring system

13. Rater standardisation materials

- Selection of 24 trial candidate speech samples representative of a wide range of language proficiency
- Creation of rater guidance documentation
- Creation of secure web pages for raters to access speaking task content and speech samples for training and standardisation purposes

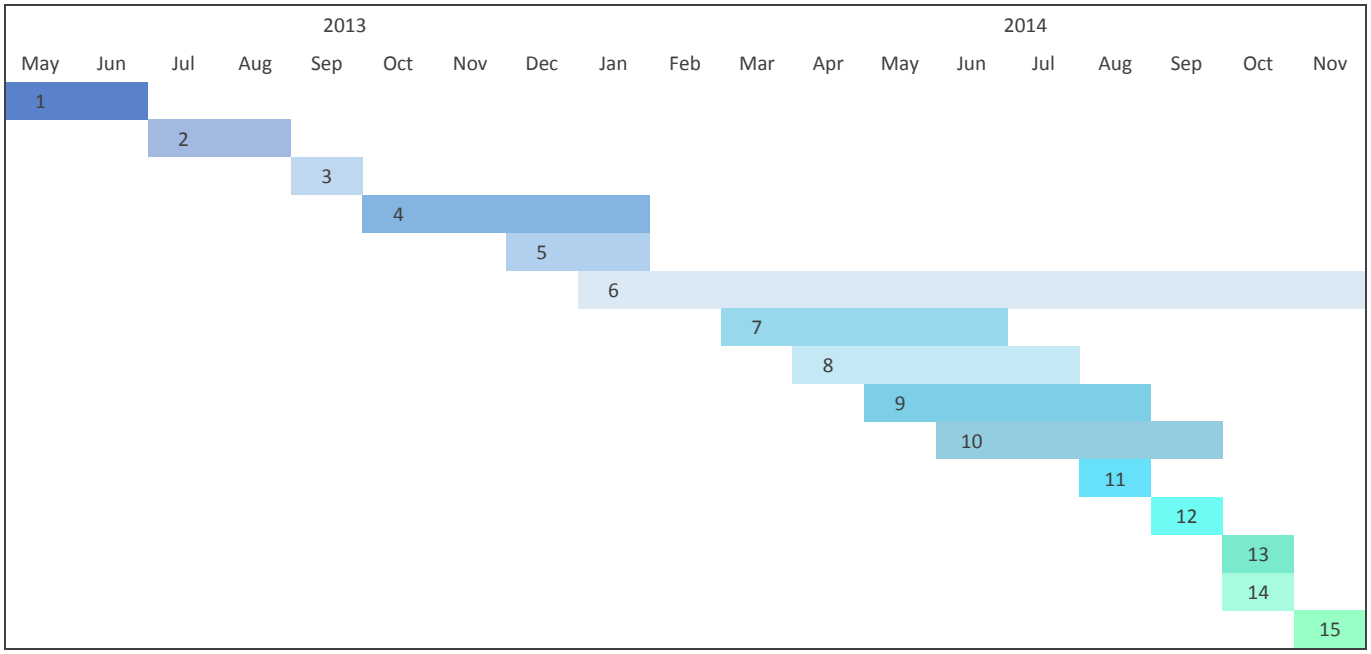
14. Standard setting

- English language expert and aviation subject matter expert raters meet to listen to and discuss a selection of with reference to the rating scale
- After the meeting, each rater rates a specific set of samples at red, yellow and green and returns scores
- Rating data analysed using Many Facet Rasch Measurement and:
  - Reliable exemplar performances selected;
  - Feedback returned to individual raters on severity and consistency

Test is launched for operational use

6. Development phases

The numbered phases in the table below correspond to the numbered stages in the test development process above.



7. Test battery and test maintenance

In the live test battery, there is a bank of two parallel versions of each task known as the ‘A set’ and ‘B set’. Tests are compiled from selections from the bank resulting in the potential for thousands of different test forms. As candidature grows, old content is retired and new content is developed, trialled and introduced into the test battery according to the procedures outlined in sections 7, 8, 9, 11 and 12 of the development process described above.

The test management team meet regularly to review test and task specifications and customer and candidate feedback, to set the research and development agenda and to identify possible improvements to the test and supporting documentation.