

Standards for test development in aviation: from placement to proficiency

*Dr Jeremy Mell
Head of Language Studies
ENAC, Toulouse, France
jeremy.mell@enac.fr*

Test design specifications (ICAO)

- *“As of 5 March 2008, aeroplane and helicopter pilots, air traffic controllers and aeronautical station operators shall demonstrate the ability to speak and understand the language used for radiotelephony communications to the level specified in the language proficiency requirements in the Appendix.” (Annex 1, 1.2.9.4)*
- *“An individual must demonstrate proficiency at level 4 in all categories in order to receive a level 4 score.” (Manual, 2.8.4)*
- *“any test at all” is NOT better than “no test” (E Mathews, PRICE Study Group)*

Institutional context

- current lack of independent international oversight to ensure:
 - validity of testing procedures used
 - comparability of testing procedures and outcomes
- testing outcomes are linked with personnel licensing
 - testing may not be in the sole hands of the ELT community
 - partnerships with CAA, service provider, airline,...
- system cannot tolerate a high failure rate
 - waste of expensive training investment
 - staffing levels must ensure continuity of service
- high priority for “wash-back” effects on training
 - trainee motivation and attitude
 - development of appropriate training systems

Delivery considerations

- time/space constraints
 - dispersed locations (recurrent)
 - irregular schedules and limited availability of test-takers (recurrent)
 - availability of technologies
 - test/re-test
- test components
 - single event, integrative?
 - different item designs to test separate skills?
- standardisation of results
 - extensive trialling
 - comparable conditions of test administration
 - examiner training/auditing
 - interlocutors
 - raters
- test security
 - secure storage and transfer of test materials
 - multiple parallel versions

Points of view within the aviation English profession

- *« We need to find workable and practical solutions to the training and testing challenges of the next few years. Those solutions may not be able to wait for the « perfect » aviation English proficiency test. »*
 - ICAEA_WORLD discussion group, 29th May 2004.
- *“I have some considerable concerns about... an individualistic approach being adopted by airlines / ATC administrations / civil aviation authorities around the world. ... my view is that it is essential that testing be subject to ‘external standards’ and be ‘standardized and universal to some degree’.*
 - ICAEA_WORLD discussion group, 14th April 2004

Current options

- Short-medium term
 - General purpose tests:
 - Standardised
 - OPI, IELTS, ...
 - “jump-off” points to ICAO scale
 - Specific tests:
 - Non-standardised
 - may be local adaptations of general tests
 - TOEFA, TNP, ...
 - Standardised
 - pilots: FCL 1.200, ...
 - controllers: PELA, ...
- Longer term
 - Extend availability of existing specific tests
 - regional variants
 - translation of mother-tongue components
 - Develop new test according to:
 - testing purpose
 - test-takers
 - technologies
 - test-givers

Testing purposes

- entry/placement
- diagnostic
- progress/achievement
- proficiency

Entry/placement tests

• Purposes

- ab initio training:
 - recruitment
 - form level groups
- recurrent training:
 - benchmarking a population
 - form level groups

• Characteristics

- items chosen to cover a broad range of levels of general language
- formats reflect previous education
- scores tend to form a « bell curve »
- stakes
 - high (recruitment)
 - mid (benchmarking)

Diagnostic tests

- Purposes

- identify specific areas of skill/knowledge for improvement in subsequent training
- individualisation of training programmes

- Characteristics

- each item chosen to represent a single significant area of knowledge/skill
- overall score less important than analysis of right/wrong responses - focus on errors
- stakes: low

Progress/achievement tests

- Purposes

- to measure effectiveness of a phase of learning
- to allow access by learner to next phase of learning

- Characteristics

- items chosen to closely reflect content and methodology of preceding training phase
- scores will often be interpreted with regard to average or pre-established norm
- stakes: low to medium

Proficiency tests

- Purposes

- to establish the competence of candidate to exercise language skills in operational conditions

- Characteristics

- items chosen to resemble real-world tasks
- overall scores are holistic: YES/NO
- stakes: very high

Tests for ICAO level 4

- Placement (entry)

- must relate to ALL 6 skills in ICAO scale
- must measure distances above and below level 4
- may have diagnostic capability
- tasks/topics/input materials focus on general or job-related skills
- development requires some statistical validation

- Proficiency (exit)

- must relate to ALL 6 skills in ICAO scale
- must characterise test-takers as BELOW or AT levels 4, 5 or 6
- tasks/topics/input material must reflect radiotelephony language competencies (including standardized phraseology)
- development requires strong statistical validation (high stakes)

Detailed specifications (proficiency)

- *provide a representative range of intelligible international accents as input for comprehension;*
- *provide a professionally relevant format for candidates to display comprehension;*
- *elicit an adequate continuous speech sample to test fluency/pronunciation;*
- *provide a voice-only setting for “diadic” (2-person) interactions;*
- *provide examples of routine and unexpected events in a work-related context;*
- *allow the candidate to use basic grammatical structures creatively;*
- *allow the candidate to demonstrate ability to paraphrase;*
- *allow the candidate to change between rehearsed/formulaic speech and spontaneous interaction;*
- *simulate unexpected events to create opportunities for misunderstanding.*
 - EUROCONTROL/ENAC preliminary feasibility study 2004

Existing specific tests

• Controllers

– ab initio

- placement: EPT (Eurocontrol)
- proficiency: PELA (Eurocontrol)

– recurrent

- placement: TNP (France)

• Pilots

– ab initio

- proficiency: FCL 1.200 (France)

– recurrent

- proficiency: TOEFA (Peru)

TNP Test national de positionnement (DGAC, France)

- developed by language experts and subject-matter experts (controllers and ATC management) as a placement/diagnostic tool
- used since 1999 with qualified air traffic controllers:
 - benchmarking to establish national training needs
 - formulating periodic individualised language development and maintenance programmes in order to meet national qualification renewal requirements
- 3 papers cover professional language content:
 - a written multiple choice test of language knowledge
 - a tape-mediated written test of listening comprehension (RT messages)
 - an individual oral interview using paper-mediated graphic and text input
- expresses results in terms of the ICAO level chart

FCL 1.200 (DGAC, France)

- developed as a proficiency testing tool in conformity with European JAR FCL (Flight Crew Licensing) requirements;
- used nationally with ab initio pilot trainees as part of the basic theoretical certification;
- 3 papers cover professional language content:
 - a written multiple choice test of reading ability (flight documentation)
 - a tape-mediated written test of listening comprehension (RT exchanges)
 - a paired “fictional flight” interaction with trained interlocutor including emergency phase and using paper-mediated text input
- currently under revision by language and subject-matter experts to conform with ICAO language proficiency requirements

Commonality of test-takers

- pilots and controllers are partners in R/T communication
- but they are set apart by:
 - different interactive roles
 - complementary passive/active competencies
 - opportunity to use other job-related language uses to extend speech sample:
 - controller: telephone co-ordinations, report to supervisor, ...
 - pilot: pre-flight, intra-cockpit, cabin announcements, ground staff,
 - e.g; PELA Paper 3, FCL 1.200
- common core test + job-specific components?

Appropriate technologies

- voice-only
 - telephone, 2-way radio, language laboratories, training simulators
- input delivery
 - analogue players, computer screen (multimedia)
- performance storage and access
 - analogue recorders, computer sound files (rapid access)
- speech recognition
 - standardise perception of intelligibility, save rating costs
 - BUT needs to filter all possible speech variations
- but « interactions » need to be human-human
 - initiate and maintain exchanges
 - deal with misunderstandings

Rater/interlocutor qualifications

- 3 basic qualifications
 - language level
 - aviation background
 - principles of language proficiency and language testing
- suitable candidates
 - aviation personnel
 - language trainers
- L1 background
 - native speakers(NS)/non-native speakers (NNS)
 - risk of familiarity with a given form of NNS spoken English