TEACHERS’ INVOLVEMENT IN HIGH STAKES TESTING:
DESIGNING AND IMPLEMENTING AN ENGLISH SPEAKING EXAMINATION

Daniel Xerri & Patricia Vella Briffa
University of Malta
Concerns with high stakes testing

- Affects the quality and nature of instruction and makes teachers feel relatively powerless (Costigan, 2002)
- Lack of consistency or coherence between teachers’ visions of desirable education and those articulated in high-stakes examinations (Atkin, 2007)
- The difficulty to act autonomously in the face of testing pressures forces teachers to reinvent their professional identity so that it’s in harmony with the new testing mentality (Assaf, 2008)
Teachers as examiners

- Educators must be able to engage with an assessment system so as to contribute to policy making and prevent a centralization of power (Gregory & Clark, 2003)

- Through involvement teachers develop ownership of the procedures and criteria and understand the process of assessment (Harlen, 2005)

- The teacher, not the test, is the primary change agent so we must bring teacher judgement to centre stage (Klenowski & Wyatt-Smith, 2012)
Developing assessment literacy

- Teachers will be better able to respond to the demands of the avalanche of high stakes testing (Gulek, 2003)
- Unless the quality of teacher judgements is addressed they will continue being excluded from high stakes testing (Brookhart, 2013)
- Hands on involvement in high stakes test development is a means of developing assessment literacy (Runté, 1998; Black et al., 2011)
Context

- MATSEC Examinations Board
- Advanced English Examination
- 600 candidates (prospective university students)
- New syllabus in 2010 with new speaking component
- First sitting in May 2013
- “a measure of the candidates’ ability to speak and converse in English” (MATSEC, 2010)
- Teachers’ initiative
- Five teachers at a postsecondary school
PART 1

- Guided examiner-to-candidate conversation
- About 3 minutes
- 4 marks
- An informal interview intended as a conversation starter
- The examiner asks questions about topics such as Work, Study, Leisure, and Career Plans
Part 2

- Guided examiner-to-candidate conversation
- About 4 minutes
- 6 marks

Consists of two stages:

a) a description of the visual prompt
b) a two-way exchange where topics and issues related to the visual prompt are developed
Part 3

- Candidate-to-examiner long turn
- About 3 minutes
- 8 marks
- A presentation expressed as a long turn by the candidate based on a topic
- Candidate selects topic from a list of 5: lifestyle, music, sport, religion, relationships, international news, environment, war, education, entertainment
- Candidate is allowed some minutes to prepare presentation before entering the examination room
Exam specifications

- The greater the detail in the specification of content, the more valid the test is likely to be (Hughes, 2003)
- Syllabus constraints
- Aims and content
- Structure, timing and procedures
- Rubrics and instructions
Writing items, moderation, trialling

- Specimen materials in line with specifications
- Avoid problems with test use
- Knowledge of student cohort
- Internal and external moderation
- Trialling, observation and feedback
It would be useful if the criteria employed in the assessment of language production on tasks could be related in a principled way to the criteria for the teaching of a skill (Weir, 2005)

- Fluency & coherence
- Pronunciation
- Vocabulary
- Grammar
# Rating scale

<table>
<thead>
<tr>
<th>Part 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency &amp; coherence</td>
</tr>
<tr>
<td>1-2 descriptor</td>
</tr>
<tr>
<td>3 descriptor</td>
</tr>
<tr>
<td>4 descriptor</td>
</tr>
</tbody>
</table>

**Calibration:**
- Video recorded 50 mock tests
- Samples of performance covering the entire range of the scale
Test manual

- Accountability via a test manual (McNamara, 2000)
- The degree of a test taker’s familiarity with the demands of a particular test may affect the way the task is dealt with (Weir, 2005)
- Dynamic and evolving document rather than monolithic (Fulcher, 2003)
- Concrete connections between theory and practice and thus educational (Luoma, 2004)
- Dissemination
Examiner training

- Socialise raters into a common understanding of the scale descriptors, and train them to apply these consistently in operational speaking tests (Fulcher, 2003)
- Trainee examiners selected from current secondary/postsecondary teachers
- Test content, procedures and assessment criteria
- Rating standardisation
- FAQs
- Cyclical and iterative (Taylor & Galaczi, 2011)
Benefits

- Empowerment via ownership
- Enhanced assessment literacy
- Beliefs, attitudes and practices
- Confidence in teachers’ judgement
- Equitable examinations
References


References

Contact

- daniel.xerri@um.edu.mt
- patricia.vella-briffa@um.edu.mt