EAP reading-to-write test design considerations

Amanda Smith, INTO Manchester

amanda.smith@into.uk.com
1. Authenticity and washback
2. My research and findings
3. Task design details specific to this project
4. Issues to consider...
Authenticity and washback

Assessments for EAP

Coursework includes

• Annotated bibliography
• Research proposal (including 1000-word literature review)
• Dissertation (5000 words)

Exams

• Reading
• Listening
• Speaking
• Writing (including a discursive essay)
Authenticity and washback

Write an essay to discuss the following opinion: “Fast food provides cheap and convenient meals for our busy modern lifestyles.”

*(NCUK Writing Examination, 2009)*

‘rhetorically formulaic’

*(Cumming et al, 2005, p5)*

Other literature confirms this view...
Authenticity and washback

• Focus on form over content
  *(Leki and Carson, 1997; Weigle, 2002)*

• Misconceptions about the academic writing process
  *(Cottrell, 2001; Weigle, 2004; Moore and Morton, 2005; Elander et al, 2006; Cumming, 2013; Wolfersberger, 2013)*

• De-prioritisation of critical thinking skills
  *(Alderson and Hamp-Lyons, 1996; Read and Hayes, 2003; Moore and Morton, 2005)*
Authenticity and washback

What’s the alternative?

‘...integrated tasks... have the potential to promote positive washback, ... encouraging teachers preparing students for large-scale tests to focus on academic skills such as discourse synthesis.’

(Weigle and Parker, 2012, p118-119)
My research and findings

Students:
• short term goal: get to university
• gatekeeping exams (eg: IELTS / NCUK discursive essay)

Teachers:
• long term goal: succeed at university
• real academic writing
My research and findings

• What are student perceptions of reading-to-write essays and independent exam essays and the process of preparing for each one?

• How do students’ grades and writing quality vary between the two essay types?
My research and findings

Procedure

• 31 Ss (5.5-7.5 IELTS)
• Discursive essay + reading-to-write essay exams
• Questionnaires (31)
• 2 focus groups (6 people in total)
• 1 teacher interview

Research categories for each essay type

• Usefulness
• Ease of preparation
• Interest of preparation
• The exam experience
• Suitability as selection tool
My research and findings

My expectations:

1. Simpler format, familiarity > prefer discursive / integrated: useful

Majority of Ss preferred r-to-w (qu’aires & FG) on all 5 criteria

2. Writing quality would suffer due to increased cognitive load

Majority performed the same or better on r-to-w

3. This loss would be most severe at lower levels

Low B- and C-grade students experienced largest and most consistent gains on the r-to-w
Task design

Text length / reading time? | Take home evening before
Whole or extract? | Whole but highlighted
How many? | Two
Topic? | PRP in education
Secondary referencing? | Editorial type articles with opposing stances
Evaluative element? | Which do you agree with?
Input / output volume? | Pilot
Task design

Please see:

• Reading task
• Writing task
• Reading texts
Task design

Criteria and marking:
Double marked using NCUK discursive essay marking criteria:

<table>
<thead>
<tr>
<th>Task fulfilment</th>
<th>Coherence / cohesion</th>
<th>Vocabulary</th>
<th>Grammar</th>
</tr>
</thead>
</table>

- Plagiarism, citation conventions ↔ task fulfilment
- Source integration ↔ coherence / cohesion
Issues to consider...

Choice of text

Influences on course design (please see sub-skills doc):
• Source integration, citation conventions, synthesis...
• Secondary referencing?
• Interpretation of author’s stance?

Large scale testing:
Different disciplines: structure and citation conventions?
Are short, factual texts necessary? Compromise?
Issues to consider...

Evaluative element

• ‘Which do you agree with?’ Feasible to come up with wide variety of evaluative qus for large scale testing?

• Possible to mark fairly?

• Limited to classroom assessment?
Issues to consider...

Reliability and fairness

• Texts in advance: possibility for pre-writing (2 focus group participants)

• Some Ss preferred r-to-w because they had time to prepare topic vocabulary in advance (even though no-stakes!)

• Reading with dicts before writing?

• Didn’t notice plagiarism, but... Turnitin??
Issues to consider...

Criteria

Adequate for purposes of this small study.

Effect of language proficiency on performance of r-to-w sub-skills???

(Cumming et al, 2005; Plakans, 2009a; Plakans and Gebril, 2012; Weigle and Parker, 2012; Cumming, 2013)

Interestingly, a pattern seemed to emerge...
Issues to consider...

Criteria

• Do criteria need to be somehow calibrated to eliminate this pattern?

• Can we be sure that discursive essay criteria are that reliable anyway?

• Authenticity of textual support > clearer picture of potential performance in real academic context???
Issues to consider...

Marker training

• Familiarity with the texts not an issue: both markers familiar – spotting plagiarism was relatively easy

• Large scale testing? Precludes use of long texts?

• Does input from texts (particularly with a stance) increase range of acceptable ways to answer? (T interview)
Conclusion

Ss might have

• greater intrinsic motivation
• a greater awareness of what they really need for university
• more willingness to engage in complex tasks

than we give them credit for...

...but many questions remain.
List of references


List of references


List of references


