

# Investigating the washback of a standardized examination on classroom assessment

Doris Froetscher

University of Innsbruck, Austria

Lancaster University, UK



# Overview

- Background
- Research Question
- Washback
- Hypotheses
- Methodology
  - Participants
  - Instrument
- Next steps

# Background

- CEFR-linked syllabus introduced 2004
  - 4 language skills
- reformed school-leaving exam „Matura“
  - operating in pilot phase since 2008 (voluntary)
  - obligatory as of 2014 in general schools
  - professionally developed, standardized tasks
  - Listening, Reading, Language in Use, Writing
  - CEFR level B2 (English)

# Research Question

- What is the washback of the standardized Matura for English in Austria on how reading is assessed in Schularbeiten?

(Schularbeiten = summative classroom-based tests)

# Washback

“the extent to which the introduction and use of a test influences language teachers and learners to do things they would not otherwise do that promote or inhibit learning”

Messick, 1996, p. 241

# Washback

changes in educational matters such as:

- curriculum
- teaching methods
- teaching and learning strategies
- material and courseware
- assessment practices
- content of instruction

Shohamy, 1999, p. 711

# Washback research

- until 1990s, general education research
- since 1993 (Alderson & Wall) in the field of language testing
  - a number of evidence-based studies
  - wide range of countries and contexts
  - washback on:
    - teaching
    - learning
    - teaching materials

# Washback of standardized tests on classroom-based assessment

- few studies
- general education literature  
(eg. Maths, English)
- triggered by No Child Left Behind Act (USA)

# Washback of standardized tests on classroom-based assessment

## results

- more time spent on classroom-based assessment (Stecher et al., 1998)
- teachers mirrored task types of standardized assessment  
(Mabry et al, 2003, McMillan et al., 1999; Stecher et al., 1998; Abrams et al., 2003; Mertler, 2010)
- mediating factors:
  - teachers' assessment literacy (Tierney, 2006)
  - stakes of the standardized test (Abrams et al., 2003)

# Washback of standardized tests on classroom-based assessment

- methods used
  - surveys
  - Interviews

→ self-report data

# Washback of standardized tests on classroom-based assessment

language education and language testing

- limited influence (Wesdorp, 1982)
- both negative and positive washback (Wall and Alderson, 1993)
  - stopped to test skills which were not part of the exam
  - more attention to skills which were part of the exam
  - same item types as in the exam
  - copied from past papers

# Research Question

- What is the washback of the standardized Matura for English in Austria on how reading is assessed in Schularbeiten?

(Schularbeiten = summative classroom-based tests)

# Phase 1

- How do Grade 12 Schularbeiten tasks for reading administered before the introduction of the new Matura differ from tasks administered afterwards?
- Can any differences be attributed to the influence of the new Matura exam?

# Hypotheses

- reading tasks in Schularbeiten constructed after the introduction of the new Matura (2008) will:
  - differ from those constructed before
  - show influence from Matura tasks in terms of
    - text characteristics
    - Task/item characteristics
  - include tasks from the new Matura (past papers)

# Methodology

- analysis of reading tasks in Schularbeiten
  - pre 2008
  - post 2008
    - Grade 12 (year when students take the Matura)
- analysis of standardized Matura tasks
- specially designed task analysis instrument:  
*Reading Task Analysis Instrument*

# Participants

- 130 teachers → 20 so far
  - using the new Matura
  - not using the new Matura (yet)
  - different backgrounds (questionnaire)
- raters
  - me
  - two other persons
    - from within SRP project
    - external

# Development of the instrument

## sources

- SRP test specifications for reading
- CEFR
- Dutch CEFR Grid
- ALTE task analysis checklists
- IATM (Instrument for Analysis of Textbook Materials; Bonkowski, 1996)
- reading literature
- coursebook material evaluation literature

# Sections of the instrument

- General
- Text
- Task
- Item

# General

- ID
- year

# Text

- Domain
- Topic
- Writer's intention
- Authenticity
- Nature of content
- Register
- Grammar
- Number of words
- Vocabulary profile
- Coh-Metrix measures

# Task

- item type
- reading behaviour tested
- rubric
- example
- sequence of items
- layout

# Items

- clarity
- paraphrased
- quality of distracters
- expected response
  - formulation
  - location
- other linguistic skills tested
- non-linguistic skills tested
- specialist knowledge tested
- cultural knowledge tested

# Response format

## selected response

### Domain

- Personal
- Public
- Occupational
- Educational

## Likert scales

### Grammar

- only simple structures
- mainly simple structures
- limited range of complex structures
- wide range of complex structures

# Next steps Phase 1

## Pilot

- part of the data
- revision of the instrument

# Phase 2

- approaches / strategies teachers follow
- construct or select reading tasks for Schularbeiten
- link to the standardized Matura exam?

→interviews

→coding analysis (CAQDAS)

# Phase 3

- teachers' beliefs
- extent/ways the standardized Matura has (not) influenced
- the way they assess reading in Schularbeiten

→ questionnaire

→ statistical analysis

# Contribution

- washback of the Austrian exam reform
- on classroom assessment in a foreign language
- combine analysis of artefacts with self-report data

# Implications

- recommendations
  - reinforce positive washback
  - prevent negative washback
- teacher training
  - in-service
  - pre-service

# References

- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory Into Practice, 42*(1), 18-29.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics, 14*(2), 115.
- Bonkowski, F. (1996). Instrument for Analysis of Textbook Materials. Lancaster University.
- Mabry, L., Poole, J., Redmond, L., & Schultz, A. (2003). Local impact of state testing in southwest Washington. *Education Policy Analysis Archives, 11*, n22.
- McMillan, J., Myran, S., & Workman, D. (1999). *The impact of mandated statewide testing on teachers' classroom assessment and instructional practices*.
- Mertler, C. A. (2010). Teachers' perceptions of the influence of No Child Left Behind on classroom practices. *Current Issues in Education, 13*(3).
- Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*(3), 241.
- Shohamy, E. (1999). Language testing: Impact. In B. Spolsky & R. Asher (Eds.), *Concise encyclopedia of educational linguistics*: Pergamon Pr.
- Stecher, B., Barron, S., Kaganoff, T., & Goodwin, J. (1998). *The effects of standards-based assessment on classroom practices: Results of the 1996-1997 RAND survey of Kentucky teachers of mathematics and writing*. University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.
- Tierney, R. (2006). Changing practices: influences on classroom assessment. *Assessment in Education: Principles, Policy & Practice, 13*(3), 239-264. doi: 10.1080/09695940601035387

**doris.froetscher@uibk.ac.at**

