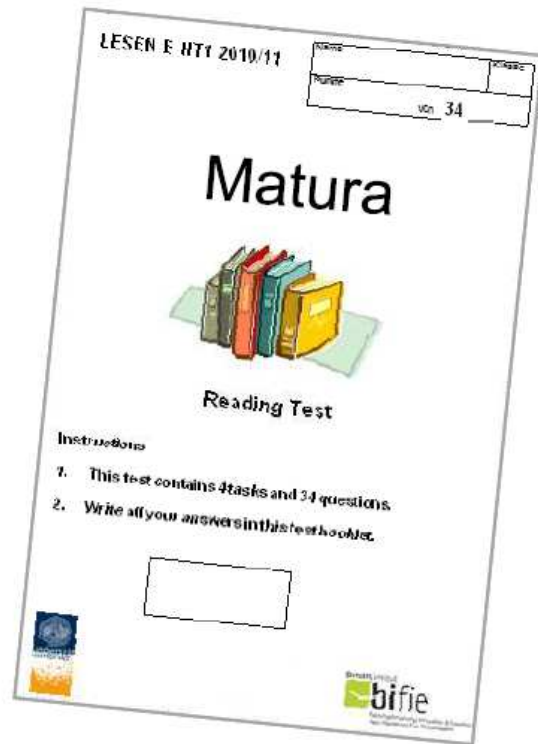# How classroom tests are shaped by a high-stakes examination: Teachers' voices

Doris Froetscher
Lancaster University

# Research Question



washback?

# 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

Important to **research** the ways in which **external tests affect assessment in the classroom**

Wall and Alderson, 1993
Wall and Horak, 2006
Watanabe, 2000

# Research into washback of standardized tests on classroom-based assessment

- ## language testing
  - limited washback (Wesdorp, 1982)
  - skills and item types of the exam (Tsagari, 2009; Wall and Alderson, 1993)
  - past papers (Wall and Alderson, 1993)
  - textbook as mediating factor (Tsagari, 2009)

- ## general education
  - more time spent on CBA (Stecher et al., 1998)
  - task types of standardized assessment (Abrams et al., 2003; Mabry et al, 2003, McMillan et al., 1999; Mertler, 2010; Stecher et al., 1998)
  - assessment literacy as mediating factor (Tierney, 2006)

# Context: Austrian school-leaving exam

- Grade 12
- high-stakes
- standardized
- national
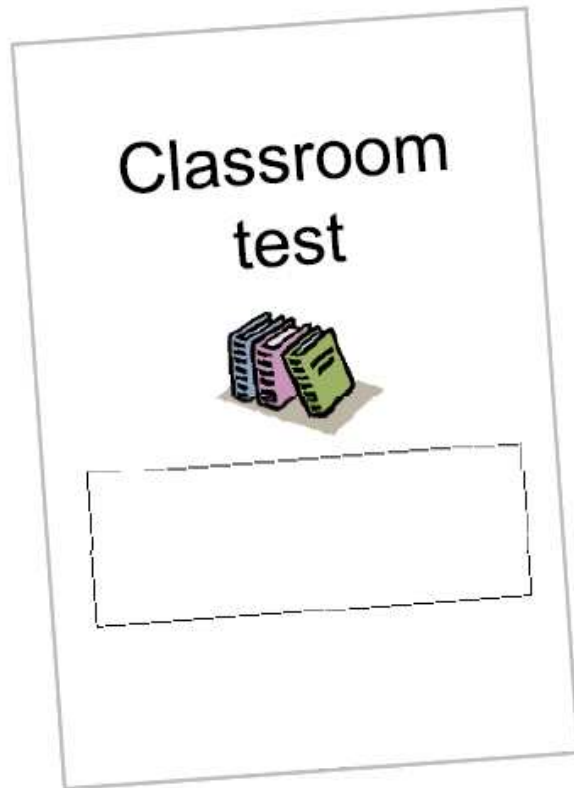  - 2015
  - 2008
- before: teacher-made

# Context: Austrian school-leaving exam
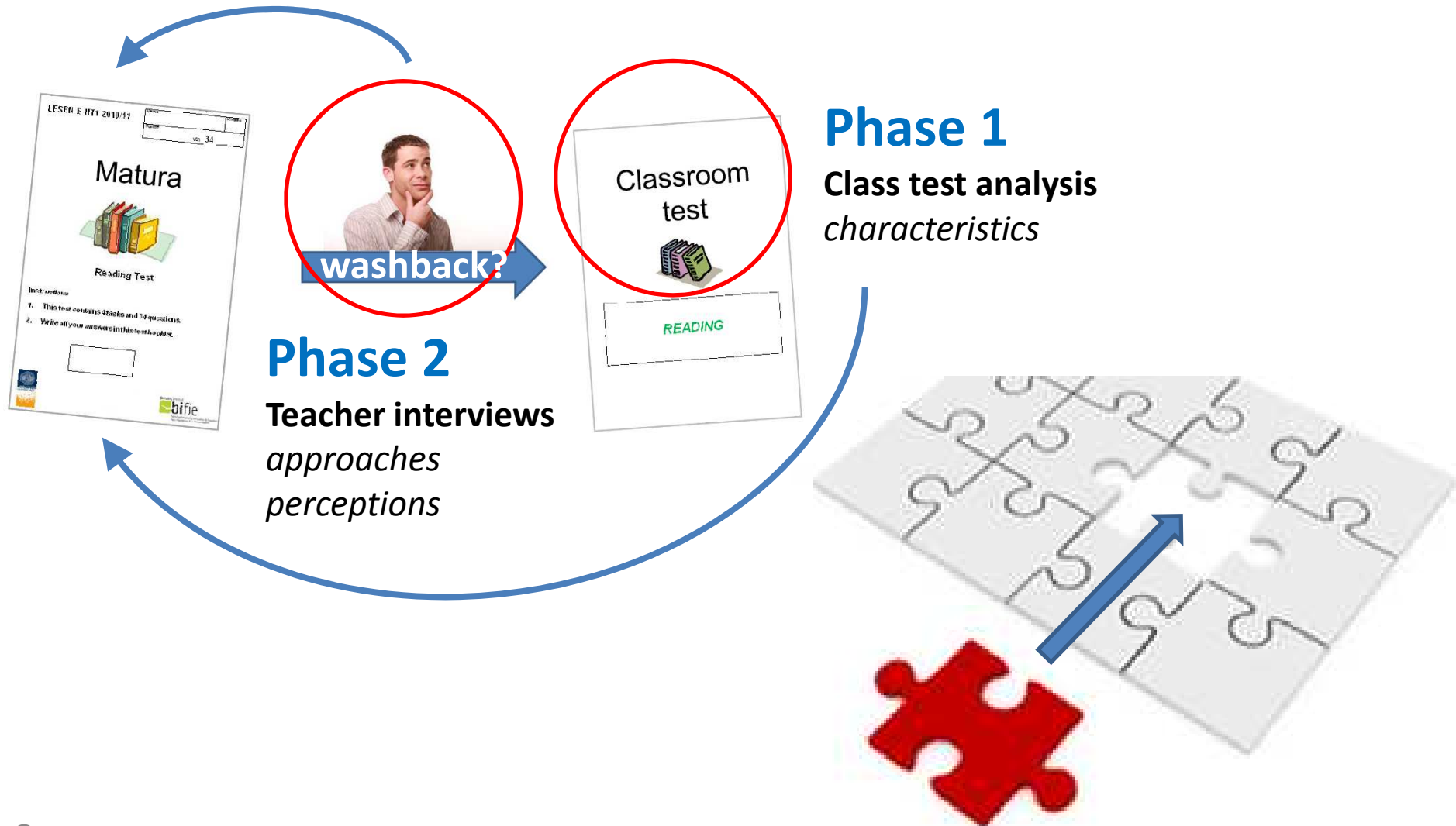


- English: B2 level
  - Reading, Listening, Language in Use, Writing

- Reading
  - 4 independent tasks
  - 6-10 items each

# Class tests

Classroom test

- set written tests
- 2-3x/school year
- 50% of final grade
- Grade 12 → Matura

- Oct12: regulations
  Grades 11 and 12
  - HAVE to include all 4 skills
  - SHOULD use Matura methods

# This study



**Phase 1**
**Class test analysis**
*characteristics*

washback?

**Phase 2**
**Teacher interviews**
*approaches*
*perceptions*

# Phase 1 methodology

22 teachers

173 reading tasks

1075 items

*Instrument for Analysis of Reading Tasks*

- text, task and item characteristics
  - pre versus post 2008
  - past papers
  - teacher variables

# Selected results: pre → post

**Text**
- different distribution of CEFR domains
- more topic variety
- less potentially distressing topics

**Task**
- fewer & different test methods (→Matura)
- more items per task
- more and clearer instructions
- more example items
- more tasks with item sequence as in text

# Factors possibly enforcing washback

pilot phase
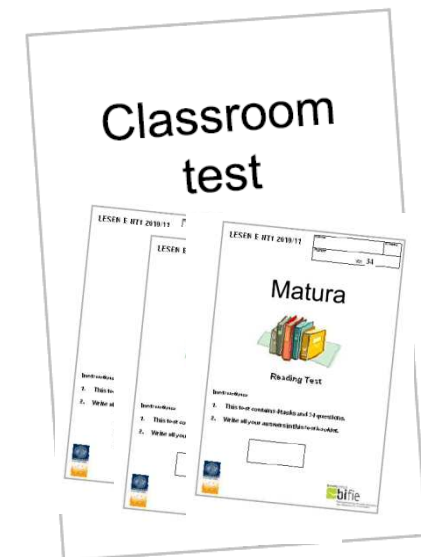- test methods
- example item

training in TEA
- presence of instructions

past papers in post 2008 class tests
- example items

# Conclusions Phase 1

- positive washback
  - students' familiarity with test methods
  - "usefulness"of class tests (Bachman & Palmer, 1996)
    - unproblematic topics, clear instructions, example item, meaningful item sequence
    - use of past papers = professionally developed tasks
- negative washback
  - narrowing down of test methods

# Phase 2 methodology

# Today: 4 teachers

- Teaching experience
  - 10-20 – 31-40 years

- Involvement in the new Matura project
  - 1 no, 3 yes – (were) item writers for writing tasks

- Training in TEA
  - mostly Matura-related
  - 1 specifically Reading
  - 1 MA in Language Testing

# Results

| | Old Matura Grade 12 | New Matura Grades 9-10 | New Matura Grade 12 |
|---|---|---|---|
| New Matura test methods | NO! | Yes, but also others | ONLY |
| Topics linked with teaching | YES! | YES! | Yes/No |
| Use of BIFIE tasks | --- | No | Yes/No |
| Use of other material | Yes | Yes | Yes/No |
| Self-created tasks | Only | Rarely | Yes/No |

# Grade 12: variety WITHIN Matura methods

*You try to prepare them in a way that they have got it all, that they have done it all, that they know what to expect, that they can't be surprised by any new formats.*

*Interviewer: Now BIFIE has a set number of methods that they use, but there are others out there, I suppose.*

*T23: But I don't do that, I don't see the point actually. No problem to do it as long as they are smaller, but in Grade 11 and 12 we really try to cover all the formats that they will be confronted with. Why bother them with something else, basically?*

16

# Results

| | Old Matura Grade 12 | New Matura Grades 9-10 | New Matura Grade 12 |
|---|---|---|---|
| New Matura test methods | NO! | Yes, but also others | ONLY |
| Topics linked with teaching | YES! | YES! | Yes/No |
| Use of BIFIE tasks | --- | No | Yes/No |
| Use of other material | Yes | Yes | Yes/No |
| Self-created tasks | Only | Rarely | Yes/No |

# Topic

- Linked with teaching

  *T15: This depends very much on the topic. I'd always try to fit the topics of the test into the topics I deal with in class.*

- Not linked with teaching

  *T23: For Grade 11 and 12 I don't take the topic so much in consideration.*

# Results

| | Old Matura Grade 12 | New Matura Grades 9-10 | New Matura Grade 12 |
|---|---|---|---|
| New Matura test methods | NO! | Yes, but also others | ONLY |
| Topics linked with teaching | YES! | YES! | Yes/No |
| Use of BIFIE tasks | --- | No | Yes/No |
| Use of other material | Yes | Yes | Yes/No |
| Self-created tasks | Only | Rarely | Yes/No |

# Use of past papers from BIFIE

*We have all the reading comprehension stuff, [...] all the gone, the past papers. And we make use of them, especially for Matura year and also Grade 11. So basically we use them for the tests, the past papers.*

*Interviewer: Yeah. Do you use other material as well in the 8th grade, other than Bifie tasks?*

*T23: No. Because it's the best thing you can have, I mean it's closest to what they will have to do then.*

20

# Considerations when using BIFIE tasks

**Trust in quality**

*T02:      But usually I was and I still am quite confident as far as the BIFIE tasks are concerned because I also saw for myself how tasks are made and how much work is behind it and that it's not something that is done very quickly, so I'm quite, yes, confident that they do work.*

**Public availability**

*T02:      And of course you have to pay attention to the tasks that have been published [...] so that you don't choose tasks that students might know from former years [...] also from teachers outside school and so on.*

# Results

| | Old Matura Grade 12 | New Matura Grades 9-10 | New Matura Grade 12 |
|---|---|---|---|
| New Matura test methods | NO! | Yes, but also others | ONLY |
| Topics linked with teaching | YES! | YES! | Yes/No |
| Use of BIFIE tasks | --- | No | Yes/No |
| Use of other material | Yes | Yes | Yes/No |
| Self-created tasks | Only | Rarely | Yes/No |

# Use of other material

*T15: I also take tasks from books, sometimes from teachers'
handbooks, but sometimes also from other course books that I
find. [...] Textbooks, there are a  lot of now Matura preparation
books, kind of, there is Going for Finals, there is Primetime
Matura Training, I think in Make your Way, someone told me, I
haven't used it yet, there are also good tasks for Matura, or for
year 8.*

# Results

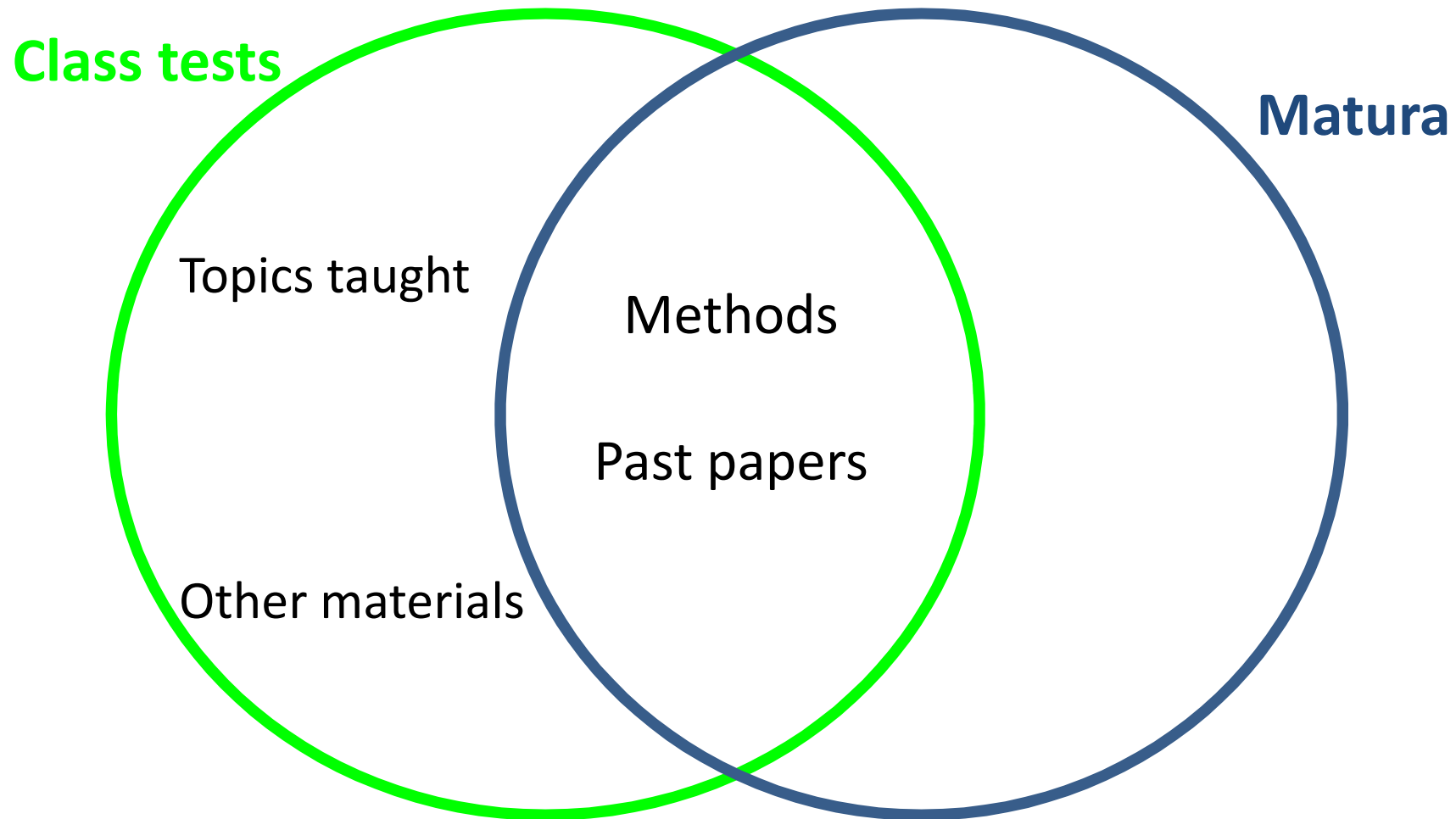| | Old Matura Grade 12 | New Matura Grades 9-10 | New Matura Grade 12 |
|---|---|---|---|
| New Matura test methods | NO! | Yes, but also others | ONLY |
| Topics linked with teaching | YES! | YES! | Yes/No |
| Use of BIFIE tasks | --- | No | Yes/No |
| Use of other material | Yes | Yes | Yes/No |
| Self-created tasks | Only | Rarely | Yes/No |

# Use of self-created tasks (T15)

*And sometimes I can't find any [tasks fitting the topics I dealt with in class], then I try to find newspaper articles and then I design my own questions. That's kind of more an open format then. [...] I have never designed a multi-choice on my own, I think that's just too much work and too difficult for one person, but the four-word-answers, that's something I did design. And then, yes actually, that's the only format I've actually designed on my own. I have no experience with multi-matching, for example, at all.*
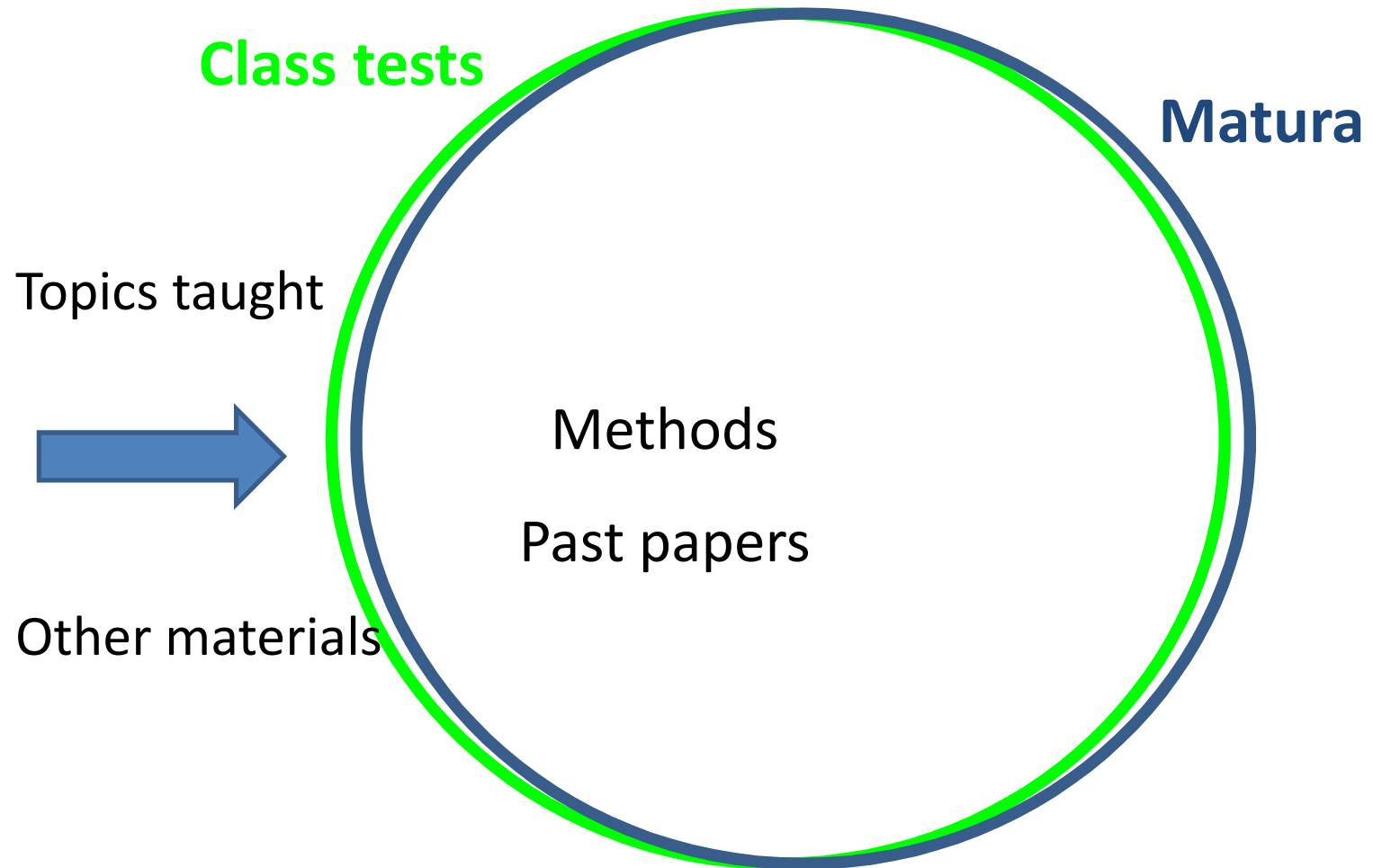
# Influence of the Matura

*And of course, maybe I should add this, we are kind of inspired by the Matura system, because we know, there will be multi-matching, there will be short answers, there will be multi-choice, and that's why I actually put them into the test situation as well. So there is this background which has to be considered.*

# Preliminary conclusion Phase 2



Class tests

Matura

Topics taught

Methods

Past papers

Other materials

# Preliminary conclusion Phase 2

**Class tests**

**Matura**

Topics taught

Methods

Past papers

Other materials

# Broader conclusions



**Washback!**

**complexity**

**Phase 2**
**Teacher interviews**
*approaches*
*perceptions*

**Phase 1**
**Class test analysis**
*characteristics*

# Implications

Teachers
→ teaching and assessing reading

Course-/testbook writers
→ language assessment

# Limitations

- Biased sample?
- Reliable coding?

# Next steps



**RELIABILITY**

# 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.

Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice (Vol. 377): Oxford University Press.

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), 241-256.

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, Los Angeles, CA.

Tierney, R. (2006). Changing practices: influences on classroom assessment. *Assessment in Education: Principles, Policy & Practice, 13*(3), 239-264.

Tsagari, D. (Ed.). (2009). The complexity of test washback (Vol. 15). Frankfurt am Main: Peter Lang.

Wall, D., & Alderson, J. C. (1993). Examining washback: the Sri Lankan impact study. Language Testing, 10(1), 41.

Wall, D., & Horák, T. (2006). The impact of changes in the TOEFL ® examination on teaching and learning in Central and Eastern Europe: Phase 1, the baseline study. ETS *RESEARCH REPORT 6, Princeton, NJ.*

*W*atanabe, Y. (2000). Washback effects of the English section of the Japanese university entrance examinations on instruction in pre-college level EFL. *Language Testing Update, 27*, 42-47.

# Thank you!

# Questions?
# Recommendations?

doris.froetscher@gmx.at