BULATS – A TOOL FOR MEASURING AND COMPARING PROGRESS IN LANGUAGE LEARNING?

KATHRIN HAHN, GOETHE-INSTITUT (ZENTRALE, BEREICH 44)

DR. WINFRIED KOCH, BDS KOCH
• Pilotproject for the implementation of tablets (iPads) in the foreign language classroom,

• Multimedial equipment for the Goethe-Institutes London, Dublin and Amsterdam,

• Preparation of special sequences and materials using tablets in class,

• Special training für the „iPad-Teachers“ at the test Goethe-Institutes,

• Monitoring throughout the whole project: BULATS test for listening comprehension as one part of the monitoring.
HOW DO TABLETS CHANGE LANGUAGE TEACHING AND LEARNING?
Mobility in the classroom
Encouragement of the learner's initiative
Multimedia products can be easily produced
Working with authentical material
RESEARCH DESIGN

Test

Bulats-Test in Test (iPad) and Control

Language Course

• B1.1
• 3 months
• 3 UE/week

Re-Test

Bulats-Test in Test (iPad) and Control
<table>
<thead>
<tr>
<th>Goethe-Institute</th>
<th>Test</th>
<th>Control</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amsterdam</td>
<td>12</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>Dublin</td>
<td>6</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>London</td>
<td>9</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>All</td>
<td>27</td>
<td>21</td>
<td>48</td>
</tr>
</tbody>
</table>
✓ since 1997, the Goethe-Institut is a BULATS-Partner from Cambridge ESOL, der Universidad Salamanca, der Alliance française
✓ BULATS ist orientated on the Common European Framework of Reference for Languages
✓ available for German, English, Spanish and French
✓ since 2008 adaptive online-version of BULATS:
  all participants start with an introductory questions at the same level. The program chooses each test-item on the basis of the previous answer. The level of difficulty is adapted to the learner's level of language.
✓ Scaled (A1-C2)
✓ Adaptive
✓ Online
✓ Reading, **Listening Comprehension**, Vocabulary und Grammar

The test is adaptive to the candidate's level of language.
BULATS – POINTS PER LEVEL

<table>
<thead>
<tr>
<th>Level</th>
<th>0-A1*</th>
<th>A1</th>
<th>A2</th>
<th>B1</th>
<th>B2</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points</td>
<td>0-9</td>
<td>10-19</td>
<td>20-39</td>
<td>40-59</td>
<td>60-74</td>
<td>75-89</td>
<td>90-100</td>
</tr>
</tbody>
</table>

✓ the points scored in the test are represented in a scale from 0-100

✓ the points in the scale represent the different levels of the Common European Framework of Reference for Languages as shown in the table

* 0-A1 includes those learners who are on their way to level A1
OBJECTIVES OF STATISTICAL ANALYSIS

• Identifying Change/Progress in Listening Comprehension (LC)

• Null Hypothesis: No difference in change/progress in LC between Test (iPad) and Control Group

• Influence of the LC-results in the first test and of the Goethe-Institute tested

• Learn our lesson for future studies
STATISTICAL TESTING OF THE NULL-HYPOTHESIS

- Mean Difference = 3.7
  95%-C.I. [-3.8, 11.2]
  n=48 (27 Test, 21 Control)

- Two-sided p values:
  - t-test $p = 0.326$
  - Wilcoxon test $p = 0.298$

$\Rightarrow$ Null-Hypothesis not rejected
Dependency of LC Change from pre-test level

Listening Comprehension Change vs. Listening Comprehension pre-test

$R^2$ (Test): 0.16
$R^2$ (Control): 0.02

Class
- Test
- Control
- Test
- Control
Progress in LC decreases with increasing pre-test level,

Above pre-test level of 80: not enough lee-way for improvement

Up to a LC pre-test level of up to 70 the 95%-confidence bands of the linear fit of the Test Group are completely above 0 → enough leeway for differences to develop
TESTING OF THE NULL-HYPOTHESIS FOR LEARNERS WITH PRE-TEST $\leq 70$

- Mean Difference = 8.0
  95%-C.I. [-2.7, 18.7]
  n=29 (17 Test, 12 Control)

- t-test $p = 0.137$
  Wilcoxon test $p = 0.144$

- Relative Effect Size = 0.578

$\Rightarrow$ Null-Hypothesis not rejected but $p<0.2$,
$\Rightarrow$ Mean difference of 8.0 relevant,
$\Rightarrow$ Signal of a relevant effect detected
Listening Comprehension mean Changes and mean pre-test Values by Institute

Mean(LC Change) & Mean(LC pre) vs. Goethe-Institute & Class

- Test
- Control

Goethe-Institute
- Amsterdam
- Dublin
- London
LISTENING COMPREHENSION MEAN CHANGES AND MEAN PRE-TEST VALUES BY INSTITUTE

- Baseline LC results are well comparable in Amsterdam (n=23) and London (n=15) less well in Dublin (n=10),

- At all three Goethe-Institutes the mean increase of LC is slightly higher in the Test Group compared to the Control Group,

- Amsterdam shows a good mean increase despite a high LC mean baseline level.
SUMMARY AND CONCLUSIONS

- BULATS is a suitable test to study factors that may influence learning of German,

- Significant effects of pre-test LC (p=0.011) and Institute (p=0.061) could be discovered,

- Use of IPad had a positive effect (p<0.2 on the subset of pre-LC <=70),

- Include subjects up to a BULATS pre-test level of 70 or at most 80 into future trials,

- To discover a relative effect size of 0.5 with 80% power, 64 learners per group would be needed.
THANK YOU FOR YOUR KIND ATTENTION.

KATHRIN HAHN:
kathrin.hahn@goethe.de

DR. WINFRIED KOCH:
winfried.koch@bdskoch.de