Scoring LSP Tasks

Christian Krekeler
Konstanz University of Applied Sciences
krekeler@htwg-konstanz.de
ALIGN TEACHING AND ASSESSMENT
TBLT-Tasks

- **Complexity**: Complex tasks
- **Authenticity**: Real-world relationship
- **Meaning**: Focus on a communicative goal
- **Cognition**: Cognitively demanding tasks

Skehan 1998; Ellis 2003; Van den Branden 2006
TBLT and Assessment

TBLT

Complexity

Meaning

Authenticity

Cognition

Assessment

mismatch
TBLT and Assessment

<table>
<thead>
<tr>
<th>TBLT</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity</td>
<td>Complexity</td>
</tr>
<tr>
<td>Meaning</td>
<td>Meaning</td>
</tr>
<tr>
<td>Authenticity</td>
<td>Authenticity</td>
</tr>
<tr>
<td>Cognition</td>
<td>Cognition</td>
</tr>
</tbody>
</table>
Task-Based Language Assessment

TBLA is a direct performance assessment which uses authentic, complex and cognitively demanding tasks. Such "holistic" assessments are scored with regard to the desired communicative outcome of the task.

Ellis 2003; Wigglesworth 2008
ALIGN TEACHING AND ASSESSMENT
ADAPT TASKS TO ASSESSMENT PURPOSES
LSP-Task

You are working as an economics editor and newsreader for the "tagesschau", the flagship news programme of the German broadcaster ARD.

- Write a short news item to inform the general public on the development of incomes.
- Record the presentation of your news item.
LSP-Tasks and Language Assessment

- teaching and learning
- feedback

- assessment
- scoring
LSP-Tasks and Language Assessment

Content and Language
Task Difficulty
Generalizability
Scoring

Tensions
LSP-Tasks and Language Assessment

Content and Language: Control Input

Task Difficulty: Limit complexity

Generalizability: Several Assessments

Scoring
Revised Task for Assessment

You are working as an economics editor and newsreader for the "Deutschlandradio", a radio station with a focus on news of the German broadcaster ARD.

• Write a short news item to inform the general public on the development of incomes.
• You have 90 minutes to write the text, to read and record it.
Revised Task for Assessment

- Use the table below on income distribution.
- Structure your text into: lead, evidence, explanation.
- Mind the word stress when reading the news item.
- [...]
Revised Task for Assessment

- limit range of performances
- easier to compare
- easier to describe task difficulty
- easier to derive and communicate rating criteria
ADAPT TASKS TO ASSESSMENT PURPOSES
SCORING REFLECTS COMMUNICATIVE GOAL
Scoring TBLA

- Quantitative Scoring
- General Impression Scoring
- Global/Holistic Scoring
- Systematic/Analytic Scoring
Quantitative Scoring

Counting errors
Error ratios

→ objective, safe
→ not suitable for scoring complex LSP tasks

Focus on meaning and a communicative goal
General Impression Scoring

No fixed criteria

→ unsystematic, difficult to legitimate
→ improve accuracy by comparisons
→ flexible, open mind

Focus on meaning and a communicative goal
Global/Holistic Scoring

- Rating scale with global descriptors
- Global score for performance

Systematic/Analytic Scoring

- Sub-scales for different features
- Separate scores
- Weighted overall score

Focus on meaning and a communicative goal
Global/Holistic Scoring

- Weighting possible?

Systematic/Analytic Scoring

- 40% Content
- 40% Language
- 20% Intelligibility

Focus on meaning and a communicative goal
<table>
<thead>
<tr>
<th>Global/Holistic Scoring</th>
<th>Systematic/Analytic Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task-specific scales</td>
<td>Focus on meaning and a communicative goal</td>
</tr>
</tbody>
</table>

Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely [...]. (CEFR)
Scoring TBLA

Quantitative Scoring

General Impression Scoring

Global/Holistic Scoring

Systematic/Analytic Scoring

It's the rater, stupid.
SCORING REFLECTS COMMUNICATIVE GOAL
CONSIDER ACTUAL PERFORMANCE FOR THE DESIGN OF RATING SCALES
# Rating scale development

<table>
<thead>
<tr>
<th><strong>Measurement-driven approach</strong></th>
<th><strong>Performance data-driven approach</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• rating scale developed by experts</td>
<td>• descriptors derived from observations of performance</td>
</tr>
<tr>
<td>• a priori method</td>
<td>• related to a context</td>
</tr>
<tr>
<td>• not sensitive to context</td>
<td>• native speaker models?</td>
</tr>
</tbody>
</table>

Fulcher/Davisdon/Kemp 2011
Rating scale development

Learner model

Native speaker model
<table>
<thead>
<tr>
<th>English Description</th>
<th>German Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introducing news</td>
<td>Ich begrüße Sie zur Tagesschau.</td>
</tr>
<tr>
<td>Providing sources</td>
<td>Dies hat eine Studie ... ergeben.</td>
</tr>
<tr>
<td>Giving numbers</td>
<td>lagen überwiegend zwischen zwei und vier Prozent</td>
</tr>
<tr>
<td>Making comparisons</td>
<td>unterscheidet sich</td>
</tr>
<tr>
<td>Qualifying information</td>
<td>und damit deutlich über ...</td>
</tr>
<tr>
<td>Giving reasons</td>
<td>Der Studie zufolge lassen sich ... durch ... erklären.</td>
</tr>
<tr>
<td>Introducing news</td>
<td>Ich begrüße Sie zu CCTV news</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Providing sources</td>
<td>Dies kann man sehen in einer Statistik aus eurostat</td>
</tr>
<tr>
<td>Giving numbers</td>
<td>das durchschnittliche Gehalt ... betrug 2014 18.462 Euro.</td>
</tr>
<tr>
<td>Making comparisons</td>
<td>während das Einkommen ... war.</td>
</tr>
<tr>
<td>Qualifying information</td>
<td>... also etwa ein Viertel des Entgelts in ....</td>
</tr>
<tr>
<td>Giving reasons</td>
<td>... werden vielleicht die Hauptgründe für ...</td>
</tr>
</tbody>
</table>
Rating scale development

Specific language ability
Writing a news item

Focus on meaning and a communicative goal
Rating scale development

**Specific language ability**
- Writing a news item
  - ...

**Content ability**
- Development of incomes

**Intelligibility**
- Reading news
  - ...

**Journalistic ability**
- Creating/presenting news
  - ...

Focus on meaning and a communicative goal
Rating scale development

Specific language ability
Writing a news item
25%

Content ability
Development of incomes
25%

Intelligibility
Reading news
25%

Journalistic ability
Creating/presenting news
25%
CONSIDER PERFORMANCE FOR DEVELOPMENT OF RATING SCALES
Scoring LSP Tasks

ALIGN TEACHING AND ASSESSMENT

ADAPT TASKS TO ASSESSMENT PURPOSES

SCORING REFLECTS COMMUNICATIVE GOAL

CONSIDER PERFORMANCE
hat aber Spaß gemacht!
Scoring LSP Tasks

van den Branden 2006 *Task-based language education*. Cambridge: CUP.

Ellis 2003 *Task-based language learning and teaching*. Oxford: OUP.


Christian Krekeler
Konstanz University of Applied Sciences
krekeler@htwg-konstanz.de