

Scoring LSP Tasks

Christian Krekeler

Konstanz University of Applied Sciences

krekeler@htwg-konstanz.de

1

ALIGN TEACHING AND ASSESSMENT

TBLT-Tasks

Complexity

Complex
tasks

Meaning

Focus on a communicative
goal

Authenticity

Real-world
relationship

Cognition

Cognitively demanding
tasks

Skehan 1998; Ellis 2003; Van den Branden 2006

TBLT and Assessment

TBLT

Assessment

Complexity

Meaning

Authenticity

Cognition

mismatch



TBLT and Assessment

TBLT

Assessment

Complexity

Complexity

Meaning

Meaning

Authenticity

Authenticity

Cognition

Cognition

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.

1

ALIGN TEACHING AND ASSESSMENT

2

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



Tensions

Task Difficulty

Generalizability

Scoring

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.
- [...]

	2009	2010	2011	2012	2013	2014	2015
Rumänien	2.162	2.037	2.116	2.116	2.066	2.195	:
Slowenien	11.864	11.736	11.999	12.122	11.852	11.909	12.332
Slowakei	5.671	6.117	6.306	6.927	6.737	6.809	:
Finnland	20.962	21.349	21.826	22.699	23.272	23.702	23.763
Schweden	21.248	19.709	22.506	24.719	26.414	27.120	:
Vereinigtes Königreich	16.262	17.106	17.136	19.166 ^(b)	18.694	20.584	:

Revised Task for Assessment

- limit range of performances
- easier to compare
- easier to describe task difficulty
- easier to derive and communicate rating criteria

2

ADAPT TASKS TO ASSESSMENT PURPOSES

3

**SCORING REFLECTS COMMUNICATIVE
GOAL**

Scoring TBLA

Quantitative Scoring

General Impression Scoring

Global/Holistic Scoring
Systematic/Analytic Scoring

Quantitative Scoring

Counting errors
Error ratios

Focus on meaning and
a communicative goal

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

General Impression Scoring

No fixed criteria

Focus on meaning and
a communicative goal

- unsystematic, difficult to legitimize
- improve accuracy by comparisons
- flexible, open mind

Global/Holistic Scoring

- Rating scale with global descriptors
- Global score for performance

Systematic/Analytic Scoring

Focus on meaning and a communicative goal

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

Global/Holistic Scoring

Systematic/Analytic Scoring

Focus on meaning and
a communicative goal

Weighting possible?

40% Content

40% Language

20% Intelligibility

Global/Holistic Scoring

Systematic/Analytic Scoring

Task-specific scales

Focus on meaning and
a communicative goal

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/Headline Scoring

Systematic/Analytic Scoring

It's the rater, stupid.

3

**SCORING REFLECTS COMMUNICATIVE
GOAL**

4

**CONSIDER ACTUAL PERFORMANCE FOR
THE DESIGN OF RATING SCALES**

Rating scale development

Measurement-driven approach

- rating scale developed by experts
- a priori method
- not sensitive to context

Performance data-driven approach

- descriptors derived from observations of performance
- related to a context
- native speaker models?

Rating scale development

Learner model



Native speaker model



Rating scale development



Language

Introducing news	Ich begrüße Sie zur Tagesschau.
Providing sources	Dies hat eine Studie ... ergeben.
Giving numbers	lagen überwiegend zwischen zwei und vier Prozent
Making comparisons	unterscheidet sich
Qualifying information	und damit deutlich über ...
Giving reasons	Der Studie zufolge lassen sich ... durch ... erklären.

Rating scale development



Language

Introducing news	Ich begrüße Sie zu CCTV news
Providing sources	Dies kann man sehen in einer Statistik aus eurostat
Giving numbers	das durchschnittliche Gehalt ... betrug 2014 18.462 Euro.
Making comparisons	während das Einkommen ... war.
Qualifying information	... also etwa ein Viertel des Entgelts in
Giving reasons	... werden vielleicht die Hauptgründe für ...

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

Focus on meaning and
a communicative goal

Intelligibility

Reading news

- ...

Journalistic ability

Creating/presenting news

- ...

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%

4

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

A photograph of a classroom. In the foreground, several rows of wooden desks are visible, each with a small black label on the front edge showing numbers like '05', '06', and '07'. A student in a blue long-sleeved shirt is leaning over a desk, resting their head on their hand. Another student in a blue shirt is sitting at a desk, looking down. In the background, a male teacher in a dark suit is standing and facing the class. A student in a black jacket is sitting at a desk, looking up and gesturing with his hands. The room has large windows on the right side, letting in natural light. An orange speech bubble is overlaid on the left side of the image, containing the text 'hat aber Spaß gemacht!'.

hat aber Spaß
gemacht!

Scoring LSP Tasks

Christian Krekeler

Konstanz University of Applied
Sciences

krekeler@htwg-konstanz.de

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

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

Fulcher/Davidson/Kemp 2011 Effective rating scale development for speaking tests. *Language Testing* 28, 5-29.

Skehan 1998 Task-based Instruction. *Annual Rev of Applied Ling* 18, 268–286.

Wigglesworth 2008 Task and performance based assessment. In Shohamy/Hornberger (Eds) *Language testing and assessment*. Boston: Springer, 111–122.