Eliciting the Constructs – Task Design

Claudia Harsch and Carol Spoettl

Universität Bremen   Universiteit Utrecht   Paragon

Overview

- Relevant factors for task design

- Task Types and Assessment Approaches

- Test Specifications

Relevant Factors for Task Design

Pedagogical Tasks ≠ Assessment Tasks

Factors impacting on Designing Tasks

- Construct (specific purpose)
- Task (types, demands and authenticity)
- Instructions, rubrics (level of input language)
- Topics (level of familiarity, abstractness, taboo topics)
- Purpose (communicative functions, goals, targeted audience)
- Assessment criteria
- Expected levels of performance

Thinking of... the Construct

- Definition of theoretical concept of the targeted ability in a given context
  - Writing / speaking processes
  - Communicative functions
  - Sub-skills (linguistic, pragmatic. etc.)
  - Targeted Output (productive, interactive, 'text' types, etc.)
  - ...

Assessment Task

- Authentic real-world situations (not course book or teacher based)
- Eliciting the targeted construct
- Purpose: solving a problem or task
- Consistency: aligned to teaching and learning goals of assessed institution
- Doable within the expected ability range
- Meaningful for test takers

Instructions

- Clear, short and concise
- Accessible for target group (ideally in a language below the target level)
- Situation, goal and audience clear
- Limitations re. content, length, time stated
- Enough information, but no redundancies
- N.B. all text you provide in the instructions and the prompt can be 'lifted' and used in the answer

Topics

Trial feedback questionnaires

Should be

- of potential interest for target group

- Consistent with curricular goals of institution, selected on basis of pedagogic principles, and consistent with frameworks like the CEFR

Avoid topics

- which could be potentially controversial or affect the candidates' performance (e.g. gender bias topics, or taboo topics; religion, health issues, death),

- or difficult to assess (e.g. creative tasks, poems)

Purpose and Assessment

- Purpose of the assessment determines the test construct:
  task and assessment criteria need to reflect the different facets of the targeted ability, appropriate for the assessment purpose

- Assessment Criteria (details tomorrow):
  - Criteria and descriptors need to be transparent, reflect construct
  - Levels need to reflect construct, expected performance levels
  - Objective, criterion-oriented

Difficulty-determining factors

Task dimensions and performance conditions influencing task difficulty

Task dimension	Performance conditions	
	Less difficult	More difficult
Number of participants or elements	few	many
Abstractness of information or task	concrete	abstract
Type of task information	immediate, here-and-now	remote, there-and-then
Nature of operation required on task information	retrieval	transformation
Familiarity of task information	familiar	unfamiliar

Note. This table is based on Skehan, 1998, p. 174, Table 7.2.

Difficulty-determining factors

- Familiarity with topic, content, task, text type, etc.
- Number of elements to address (e.g. content, characters, relationships)
- Cognitive operations / Processing demands (e.g. reasoning, interpretation, inferencing, organising, transforming, etc.)
- Task distance: here/now vs. there/then, personal vs. unknow addressees
- Abstractness of topic (concrete objects – abstract concepts)
- Degree of precision (in instructions, prompt, as well as in task demand to express an idea)
- Linguistic complexity (in prompt, as well as demanded in output)
- Length, time, stakes, amount of control
- Layout

Based on Norris et al. 2002, Prahbat 1987, Robinson and Gilabert 2007, Skehan and Foster 2001

Task Types

Speaking task typology (Fulcher, 2003: 57)

- Task Orientation
  - Open (outcomes dependent upon speakers)
  - Guided (by instructions)
  - Closed (outcomes dictated by input or instructions)
  - Goal orientation (none, convergent, divergent)
- Interactional Relationship
  - Non-interactional (monologue)
  - Two-way (interlocutor – student; student – student)
  - Multi-way (interlocutor – student – student)
- Interlocutor
  - face-to-face
  - online
  - no interlocutor

Writing task typology

Think about:
- The authenticity of the writing situation
- The audience and the impact on the register
- The appropriacy of the test takers role in the task
- The functions and the impact on the language
- The consistency in the instructions across tasks
- The amount of input given
- The amount of context given
- productive or interactional or integrated tasks

Task Characteristics

- Task demands
  - Information processing
  - Cognitive demands
- Expected output
  - Text type
  - Register, style – readership, audience
  - Content
  - Linguistic range and accuracy
- Reflected in assessment criteria
- Reflected in reporting (e.g. in reference to CEFR)

Brown et al. (1984), Prabhu (1987), Robinson (2001), Skehan (1998), Skehan & Foster (2001)

Different approaches: multi-level or level-specific?

Multi-level approach

- Tasks spanning several levels
- Performances assessed against criteria spanning several bands/levels

- Difficulty of tasks not necessarily known a-priori
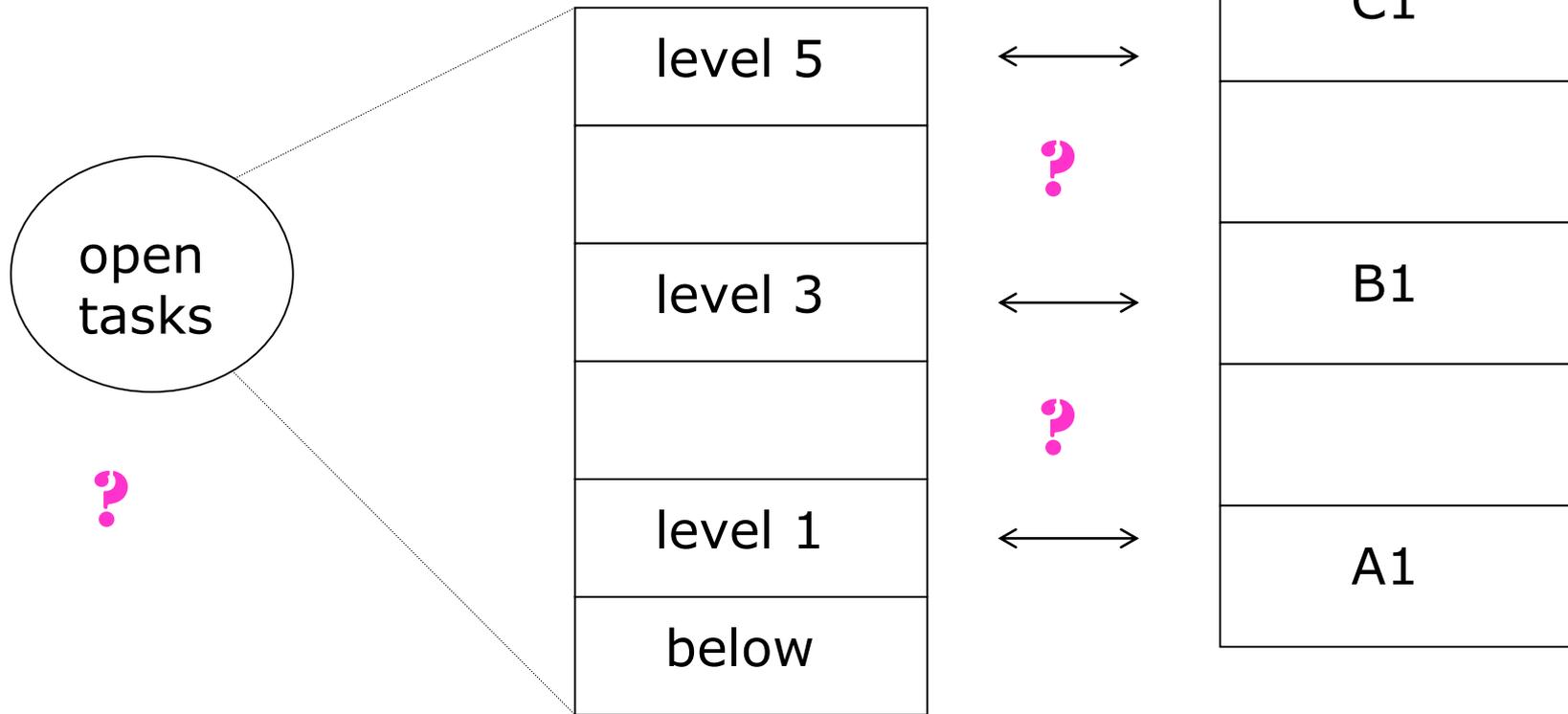- Relation task difficulty – rating scale levels – proficiency levels?

e.g. Hamp-Lyons, 1996; Hamp-Lyons & Kroll, 1996; Kroll, 1998; Weigle, 2002
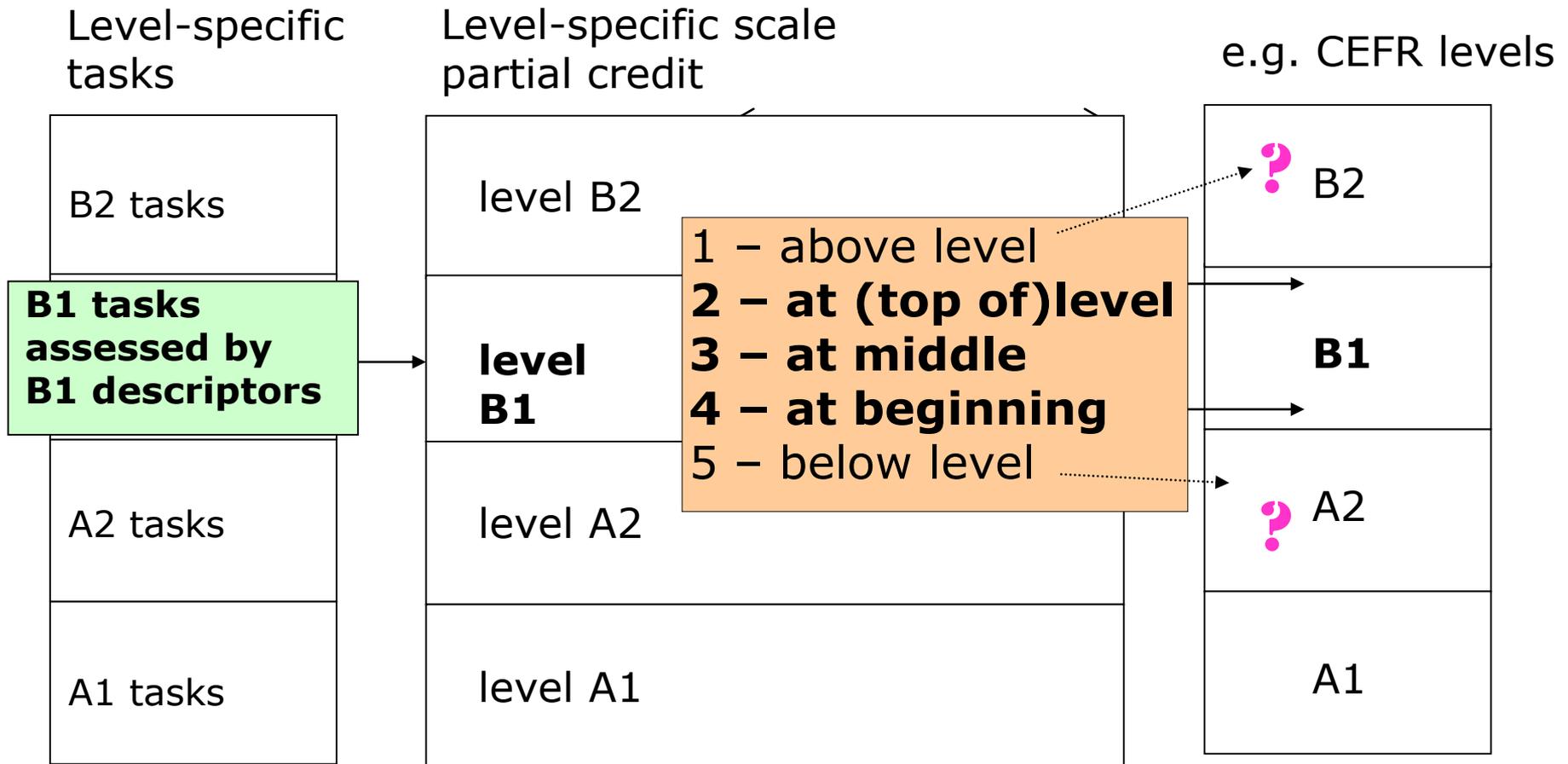
Illustration multi-level approach

Different approaches: multi-level or level-specific?

Level-specific approach

- Tasks targeting a specific level
- Performances assessed against
  - only the proficiency level targeted by the task
  - level further differentiated by several bands

- Difficulty of tasks a-priori targeted
- Relation rating scale levels – proficiency levels?

e.g. Harsch, 2010; Harsch & Rupp, 2011

Illustration level-specific approach

Level-specific tasks	Level-specific scale partial credit		e.g. CEFR levels

Level-specific tasks

- B2 tasks
- **B1 tasks assessed by B1 descriptors**
- A2 tasks
- A1 tasks

Level-specific scale partial credit

- level B2
- **level B1**
- level A2
- level A1

1 – above level
2 – at (top of)level
3 – at middle
4 – at beginning
5 – below level

e.g. CEFR levels

- ? B2
- **B1**
- ? A2
- A1

Test Specifications

Why test specifications?

- Test specifications describe and define what a test tests and how it sets out to test what it tests
- There are different test specifications for different audiences (e.g. test takers, school inspectors, test users etc.)
- They serve three main purposes:
  - guide the test development – 'blueprint'
  - form the basis for validation research
  - inform potential test users

What do test specifications look like?

- Main aspects to be considered:
  - the test's purpose, target group, the areas and content covered
  - the targeted skills, competences, abilities (construct)
  - the approaches, tasks and formats employed
  - aspects of administering and scoring
  - interpretation and reporting aspects, intended washback effects

- Possible form – a [table](#)

Construct and learning outcomes: why cover them in test specifications?

- Learning outcomes as assessment objectives
- Your tests need to validly elicit the processes, skills, knowledge etc. which are set out in your learning outcomes
- in 'test speak': *the learning outcomes need to be operationalized in the test* ☺
- also called 'constructive alignment'
- We need formats which validly elicit these processes and allow us to assess whether our students have reached the learning outcomes
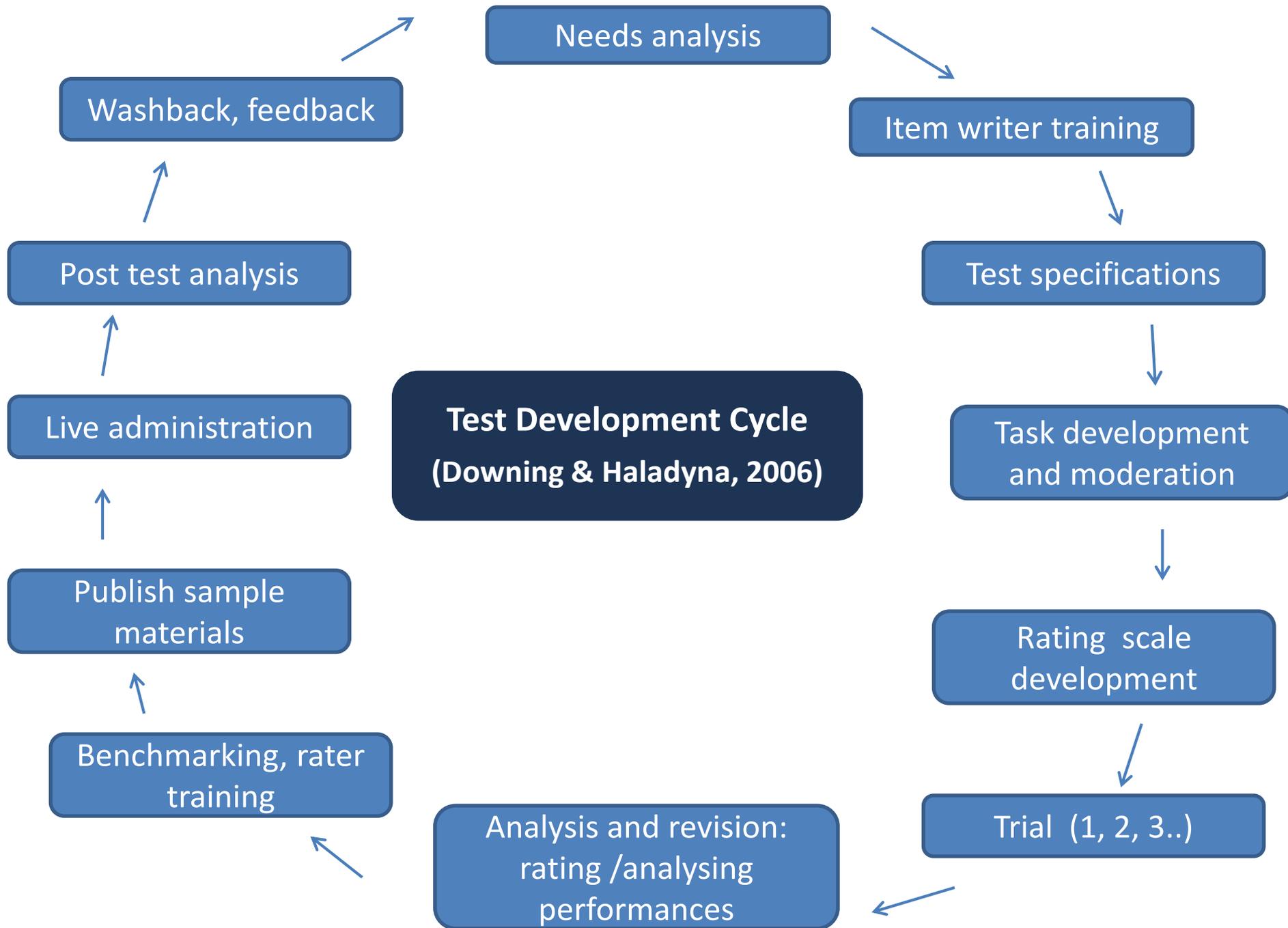
Task characteristics which worked in the context of evaluating German Educational Standards (9th graders, EFL) => reflected in test specifications

- Differentiate *prompt, topic, content* (based on SLA research)

- Define *specific purposes, text types, discourse types* (based on CEFR)

- Describe *expected output, textual, organisational and linguistic features* (based on CEFR and empirical analysis of performances) in the rating scale

Harsch & Rupp 2011

Task characteristics German Educational Standards (9th graders, EFL)

CEFR-Level	Task Characteristics	Specific Purpose and Text Types	Discourse Types	Output: Textual / Linguistic Expectations
A1	• very short prompts on concrete topics • concrete, very simple and familiar topics • all expected content points clearly stated	Specific Purpose: • ask for or pass on personal details Text Types: • write a list • fill in forms • write a postcard	• Narrative • Descriptive • Phatic • Simple expository	• very short and simple output • groups of words • simple isolated phrases • short sentences • highly frequent, elementary vocabulary
B1	• short prompts on concrete or slightly abstract topics • topics of personal as well as general interest • expected content points listed	Specific Purpose • pass on information • provide reasons for actions and comments Text Types: • write notes and messages • write personal letters, simple formal letters and emails • write reports/ articles	• Narrative • Descriptive • Instructive • Expository • Simple argumentative	• short, straightforward, detailed texts (110-140 words) • logical development of ideas • coherent text • common, frequent grammatical features • task-specific vocabulary

Harsch & Rupp 2011

Task Development Cycle

Test Development Cycle
(Downing & Haladyna, 2006)

- Needs analysis
- Item writer training
- Test specifications
- Task development and moderation
- Rating scale development
- Trial (1, 2, 3..)
- Analysis and revision: rating /analysing performances
- Benchmarking, rater training
- Publish sample materials
- Live administration
- Post test analysis
- Washback, feedback

References

Brown, G., Anderson, A., Shillcock, R., & Yule, G. (1984). *Teaching talk: Strategies for production and assessment.* Cambridge, UK: Cambridge University Press.

Downing, T.M. & Haladyna, S.M. (2006). *Handbook of Test Development*. London: Routledge.

Fulcher, G. Webinar at https://larc.sdsu.edu/glennfulcher/assessingspeaking4-12.pdf

Fulcher, G. (2003). *Testing Second Language Speaking*. London: Routledge.

Hamp-Lyons, L. (1996). The Challenges of Second-Language Writing Assessment. In E. White, W. Lutz & S. Kamusikiri (Eds.), *Assessment of writing: Politics, Policies, Practices* (pp. 226-240). New York: MLAA.

Hamp-Lyons, L., & Kroll, B. (1996). Issues in ESL Writing Assessment: An Overview. *College ESL, 6*(1), 52-72.

Harsch, C., & Rupp, A. (2011). Designing and scaling level-specific writing tasks in alignment with the CEFR: a test-centered approach. *Language Assessment Quarterly*, 8(1), 1-34

Norris, J. M., Brown, J. D., Hudson, T. D., & Bonk, W. (2002). Examinee abilities and task difficulty in task-based second language performance assessment. *Language Testing*, 19(4), 395-418

Prabhu, N. S. (1987). *Second language pedagogy. Oxford, UK: Oxford University Press.*

Robinson, P. & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *IRAL: International Review of Applied Linguistics in Language Teaching*, Vol. 45 Issue 3, p161-176.

Skehan, P. & Foster, P. (2001). Cognition and tasks. In Robinson, P. (Ed.) *Cognition and second language instruction*. (pp. 183-205) Cambridge: Cambridge University Press.

Weigle Cushing, S. (2002). *Assessing Writing*. CUP.