



Introduction to basic issues in foreign language assessment

Dr Claudia Harsch
University of Warwick

Overview



- Basic concepts
- Assessment purposes, use and interpretation
- Educational frameworks and reference criteria
- Principles of 'good' assessment practice
- Assessment tools and approaches
- Designing assessment in line with 'good principles' for the language classroom
- Integrating teaching, assessment and evaluation

Overview



- **Basic concepts**
- Assessment purposes, use and interpretation
- Educational frameworks and reference criteria
- Principles of 'good' assessment practice
- Assessment tools and approaches
- Designing assessment in line with 'good principles' for the language classroom
- Integrating teaching, assessment and evaluation



Evaluation

Assessment

Testing

Assessment – what is it?



Assessment in education is best described as an action “to determine the importance, size, or value” (of a program, proposal or a course)

Wikipedia

Assessment

the act of **making a judgment** about something...

[count]

-The school uses a variety of tests for its annual student assessments. [=evaluations] [noncount].

Merriam-Webster's learnerdictionary.com


Brainstorm



- What assessment concepts do you know?
- Which tools have you experienced so far?
Think of your learner and teacher experience
- Which assessment types enhance learning?
Why?
- Are there any ways of assessment which enhance teaching?

Evaluation – Definition?



eval·u·ate  *transitive verb* \i-'val-yə-,wāt, -yü-,āt\

eval·u·at·ed | eval·u·at·ing

Definition of EVALUATE



1 : to determine or fix the value of

2 : to determine the significance, worth, or condition of usual by careful appraisal and study

- eval·u·a·tion  *noun*
- eval·u·a·tive  *adjective*
- eval·u·a·tor  *noun*

Examples of EVALUATE

- We need to *evaluate* our options.
- *evaluate* a training program as effective

Origin of EVALUATE

back-formation from *evaluation*, from French *évaluation*, from Middle French *evaluacion*, from *esvaluer* to evaluate, from *e-* + *value* value

First Known Use: 1842

<http://www.merriam-webster.com/dictionary/evaluate>

Evaluation is the **structured interpretation** ... of results. It looks at original objectives, and at what is either predicted or what was accomplished and how it was accomplished. ... Evaluation can be **formative**, that is taking place during the development of a ... project or organization, with the intention of improving the value or effectiveness of ... (it). It can also be **summative**, drawing lessons from a completed action or project or an organisation at a later point in time ...

Wikipedia

A 'Culture of Evaluation'



Evaluation

Informal Assessment

Informal tests
Feedback (peer/teacher)

Achievement
Curriculum
Norm- /criterion-orientation

Planning
Diagnosis
School grades

Self-Assessment

Individualised
Portfolio

Proficiency /
achievement
Planning
Processes

Reflection
Self-regulation
Autonomy

External Assessment

Standardised Tests
Empirical basis

Proficiency
Criterion-orientation

Comparison
Educational
monitoring
Planning

Two worlds



Blaska & Krekeler, 2009

Differences between World of Language Classrooms – World of Testing Institutions



Think about...

- What are the stakes, consequences?
- Who is involved in the assessment?
 - Development, administration, marking
 - Research
 - Reporting and interpretation
- How systematic are assessments, results and consequences monitored?
- Who will use the results for which purposes?

Overview



- Basic concepts
- **Assessment purposes, use and interpretation**
- Educational frameworks and reference criteria
- Principles of ‘good’ assessment practice
- Assessment tools and approaches
- Designing assessment in line with ‘good principles’ for the language classroom
- Integrating teaching, assessment and evaluation

Assessment Use



- Use assessment results to inform decisions about....

individual
students

planning
teaching

developing
curricula

improving
schools

Bailey & Brown, 1996

Assessment Purposes



- To measure development
- To diagnose where learners are
- To measure proficiency
- To gain feedback on our teaching
- To give feedback to our learners
- To report to external bodies for evaluation, monitoring
- To certify language proficiency
- To regulate entrance to programmes (high stakes)
- To compare learners
- To maintain standards
- ...

Assessment Types



formative

summative

proficiency

achievement/progress

diagnostic

placement

entrance/selection

exit/certification

criterion-referenced

norm-referenced

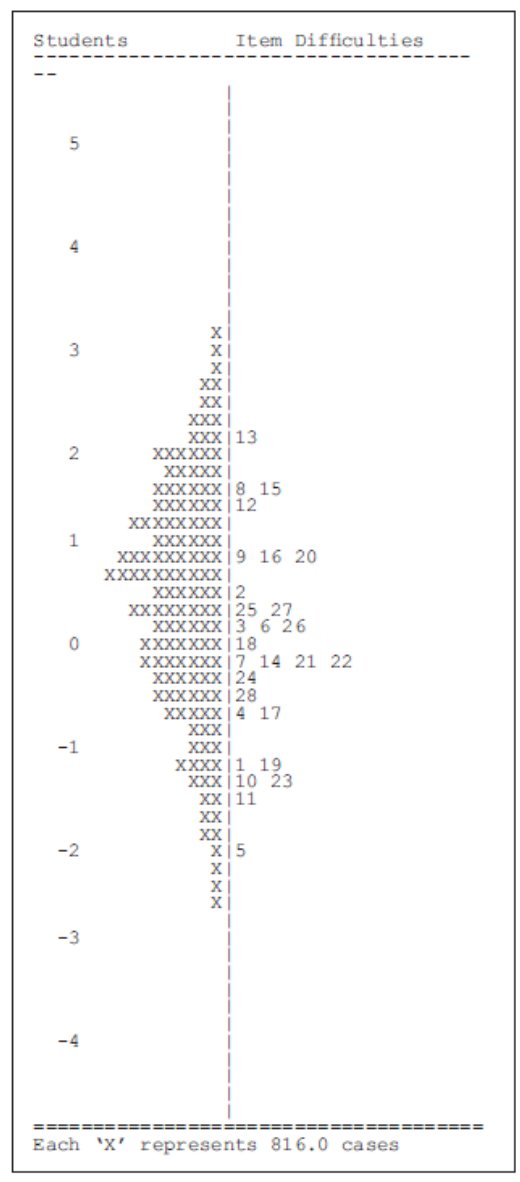


If we want to make informed and justifiable decisions based on assessment results, we need to ensure the quality of the assessment.

Quality of Assessment Re

- Quality of Instruments
 - Quality control?
 - Empirical analysis?
- Quality of Marking
 - Standardised?
 - Marker training and answer keys?
 - Quality monitoring?
- Calculating results
 - Add up raw scores?
 - Statistics? CTT or IRT?

Figure 13.2 ■ Item plot for reading items



Assessment Interpretation



- Interpretation
 - Understandi
 - Understandi
 - Understandi results
- Interpretation
 - Who interpr
 - Norm- or cri
 - Who reports
 - Which instru results and i

| Ihr Kind | Kompetenzstufen – Leseverständnis Englisch | Klasse | Land |
|---|---|--------|------|
| | B1.2 Kann weitgehend selbstständig verschiedene Texte aus Themenfeldern des eigenen Interessen- und Erfahrungsbereiches lesen und verstehen. | 9 % | 18 % |
| | B1.1 Kann unkomplizierte Sachtexte über Themen, die mit den eigenen Interessen und Fachgebieten in Zusammenhang stehen, mit befriedigendem Verständnis lesen. | 17 % | 18 % |
| | | 34 % | |
| | A2.2 Kann kurze, einfache Texte zu vertrauten konkreten Themen verstehen, in denen gängige alltags- oder berufsbezogene Sprache verwendet wird. | 39 % | 33 % |
| | A2.1 Kann kurze, einfache Texte lesen und verstehen, die sehr häufig genutzte und einen gewissen Anteil international bekannter Wörter enthalten. | | 8 % |
| A1 Kann sehr kurze, einfache Texte Satz für Satz lesen und verstehen, indem er/sie bekannte Namen, Wörter und einfachste Wendungen herausucht und, wenn nötig, den Text mehrmals liest. | 23 % | | |

Overview



- Basic concepts
- Assessment purposes, use and interpretation
- **Educational frameworks and reference criteria**
- Principles of 'good' assessment practice
- Assessment tools and approaches
- Designing assessment in line with 'good principles' for the language classroom
- Integrating teaching, assessment and evaluation

Educational frameworks



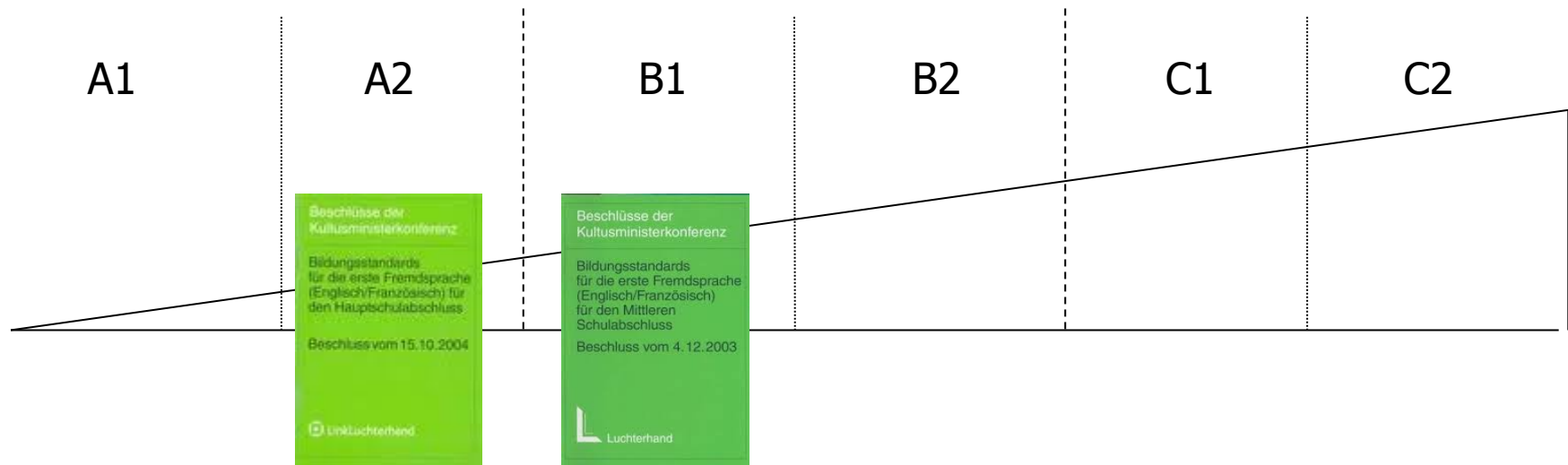
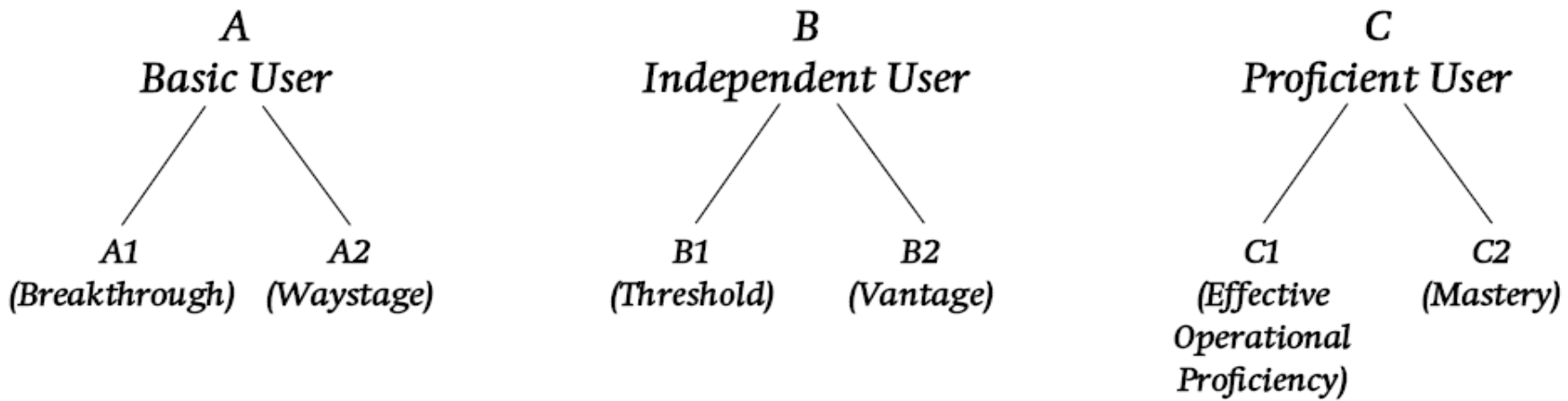
- Frameworks such as curricula, educational standards or the CEF help interpreting assessment results
- Frameworks as reference tools for teachers, self-assessment, external assessment
=> enhancing communication with regard to curricula, teaching goals, assessment criteria

Educational frameworks



- Competency model with levels of proficiency
- Description of what learners' abilities as positive 'Can do' statements
- Framework allows qualitative description of assessment results (rather than only numerical scores) – reference criteria

CEF and Educational standards: Levels of Proficiency



ES: Competency Model



| Functional communication competencies | |
|---|---|
| Language Activities | Linguistic Competencies |
| <ul style="list-style-type: none"> • Listening comprehension and audio-visual comprehension • Reading comprehension • Speaking <ul style="list-style-type: none"> -Participation in conversations -Coherent speech • Writing • Language mediation | <ul style="list-style-type: none"> Vocabulary Grammar Pronunciation and intonation Spelling |
| Intercultural competencies | |
| <ul style="list-style-type: none"> • Socio-cultural orientation knowledge • Sensitive approach to cultural diversity • Practical approach to intercultural encounters | |
| Methodological competencies | |
| <ul style="list-style-type: none"> • Text reception (listening-, audio-visual- and reading comprehension) • Interaction • Text production (speaking and writing) • Learning strategies • Presentation and use of different media • Conscious learning and organisation of learning | |

CEF: Illustrative Scale Reading



| OVERALL READING COMPREHENSION | | Ihr Kind | Kompetenzstufen – Leseverständnis Englisch | Klasse | Land |
|-------------------------------|--|-------------|---|--------|------|
| C2 | Can understand and interpret abstract, structurally complex texts and understand a wide range of explicit and implicit as well as explicit | [Empty box] | B1.2 Kann weitgehend selbstständig verschiedene Texte aus Themenfeldern des eigenen Interessen- und Erfahrungsbereiches lesen und verstehen. | 9 % | 18 % |
| C1 | Can understand in detail texts in an area of speciality, provided | | B1.1 Kann unkomplizierte Sachtexte über Themen, die mit den eigenen Interessen und Fachgebieten in Zusammenhang stehen, mit befriedigendem Verständnis lesen. | 17 % | 18 % |
| B2 | Can read with a large degree of fluency different texts and purposes and use a broad active reading vocabulary and idioms. | | A2.2 Kann kurze, einfache Texte zu vertrauten konkreten Themen verstehen, in denen gängige alltags- oder berufsbezogene Sprache verwendet wird. | 34 % | 33 % |
| B1 | Can read straightforward texts of personal interest with a satisfactory degree of fluency. | | A2.1 Kann kurze, einfache Texte lesen und verstehen, die sehr häufig genutzte und einen gewissen Anteil international bekannter Wörter enthalten. | 39 % | 8 % |
| A2 | Can understand short, simple texts on a high frequency everyday or shared interest and a proportion of shared interest | | A1 Kann sehr kurze, einfache Texte Satz für Satz lesen und verstehen, indem er/sie bekannte Namen, Wörter und einfachste Wendungen herausucht und, wenn nötig, den Text mehrmals liest. | | 23 % |
| A1 | Can understand very short texts on names, words and basic phrases | | | | |



CEF as reference framework in Germany

- Curriculum development
- Educational standards
- Reference for external assessment tools
- Planning teaching
- Textbooks use CEF as orientation
- Self-assessment
- European Language Portfolio

=> linking element, reference for evaluation

Overview



- Basic concepts
- Assessment purposes, use and interpretation
- Educational frameworks and reference criteria
- **Principles of 'good' assessment practice**
- Assessment tools and approaches
- Designing assessment in line with 'good principles' for the language classroom
- Integrating teaching, assessment and evaluation

Good Practice in Assessment I



- Be **fair** and **transparent**
- Be **objective** – clear guidelines for marking
- Use approaches/instruments which are **fit for purpose**
 - Types and formats aligned to aims and purposes
 - Familiarity: Preparation
 - Right level
- Use **reliable** and **valid** approaches/instruments
 - Construct
 - Instructions
 - Marking

Good Practice in Assessment II



- Do assessment regularly and **systematically**
- Attend **training** to produce appropriate tasks
- **Evaluate** the quality of an external instrument (fitness for purpose, look at research)
- Think of how to **interpret** and **use** the assessment

Source: Ainslee (2004)

- See also e.g. [EALTA Guidelines for good practice – pdf](#)

Two worlds – many aims



World of Language Classroom

– World of Testing Institutions: Different principles apply



| | external | classroom |
|---|---|---|
| Test alignment to classroom, curriculum | Not necessarily, can be theory-based | Alignment to classroom, curriculum and theory |
| Pilot and calibration | yes | No (but there are ways to do it!) |
| Assesseees - assessor | Large scale, sample, unknown | Small scale, known |
| Development | Central, standardised, trained item writers | Decentral, not standardised, teachers (but training available!) |

World of Language Classroom

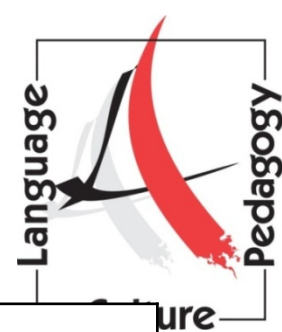
– World of Testing Institutions: Different principles apply



| | external | classroom |
|-----------------------------|-------------------------------------|--|
| Purpose | Proficiency | Performance, achievement, development, learning |
| Consequences | High-stakes (sometimes low-stakes) | Low-stakes (sometimes high-stakes) |
| Basis of measurement | One test | Several tests, assessment points |
| Aims | Certificate, educational monitoring | Enhancing learning, teaching, gain school grades |

Tasks for Teaching and Tasks for Testing

Different aims and principles



| Tasks for Learning | Tasks for Testing |
|--|---|
| Focus on processes | Focus on products |
| Cooperative learning | Individual performance |
| Interaction between learners and teacher-learner(s), individualised or cooperative | Standardised conditions |
| Errors as possibilities for learning | Avoidance of errors |
| Complex tasks, allowing for diverse ways of solving them, subjectively scored | Tasks focusing on rather specific (sub-)skills with unambiguous solutions, objectively scored |

Teaching and Testing



- How do tests affect teaching and learning?
=> teaching to the test – what effects can it have?
- In which context will it be **positive**, i.e. enhance learning and teaching?
- Should this be the purpose of a “good” test?

Overview



- Basic concepts
- Assessment purposes, use and interpretation
- Educational frameworks and reference criteria
- Principles of ‘good’ assessment practice
- **Assessment tools and approaches**
- Designing assessment in line with ‘good principles’ for the language classroom
- Integrating teaching, assessment and evaluation

Assessment tools



- Tests
- Portfolios
- Feedback
(written or oral)
- Role plays
- Projects
- (Drama) performances
- Presentations
- Interviews
- Questionnaires
- Exams
- Student observations
-



Assessment approaches

- Paper-pencil
- Computer-based
- Performance
- Process / product
- Direct / indirect ...
 - Individual
 - Group
 - Interactive
 - Dynamic
 - Self-assessment
 - Peer assessment
 - Teacher assessment
 - External assessment

Assessing language skills



- Four skills – what about mediation or intercultural communication? (e.g. Paran & Sercu, 2010)
 - Enabling skills like grammar, vocabulary, orthography?
 - Separate or integrated? Reflection of ‘real world’ tasks?
 - Feasible approaches for the different skills?
Direct or indirect?
- ⇒ See the *Cambridge Assessment Series* for a helpful first overview
- ⇒ See the *Into Europe Materials* for excellent test

Learning outcomes – beyond assessing language skills



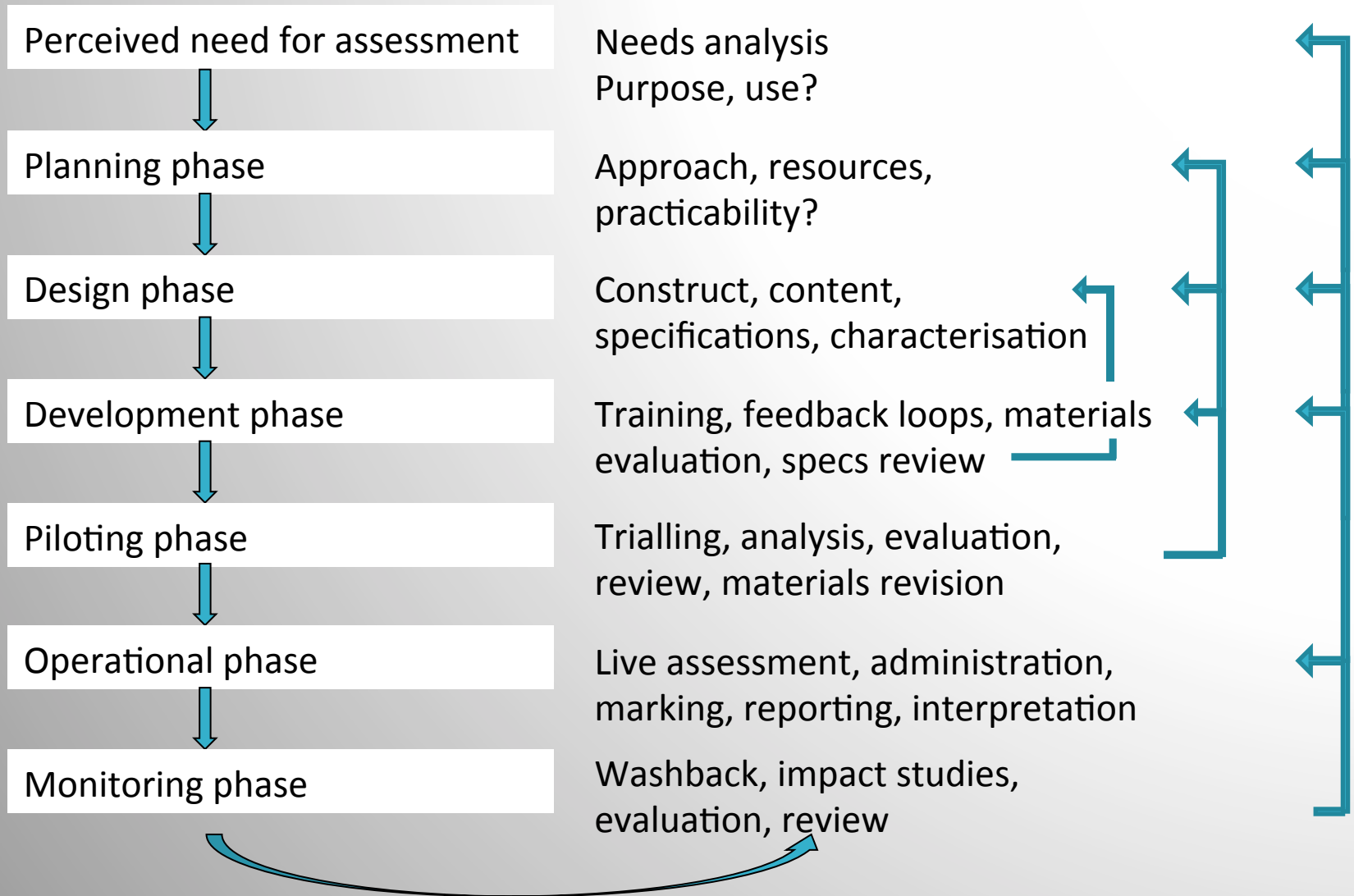
- (Intercultural) Communicating
 - Designing, creating, performing
 - Demonstrating knowledge and understanding
 - Accessing and managing information
 - Managing and developing oneself
 - Performing procedures and demonstrating techniques
 - Solving problems and developing plans
 - Thinking critically and making judgements
- => See e.g. [OCSLD](#) for an overview of suitable methods

Overview



- Basic concepts
- Assessment purposes, use and interpretation
- Educational frameworks and reference criteria
- Principles of 'good' assessment practice
- Assessment tools and approaches
- **Designing assessment in line with 'good principles' for the language classroom**
- Integrating teaching, assessment and evaluation

Cyclical Model of development process (Milanovic, 2002)



Classroom assessment



- Who would be responsible for which aspects?
- Which phases would be considered?
- How could you as teacher ensure validity, reliability, fairness, objectivity, appropriate use and interpretation, and positive washback?

=> **Helpful resources:**

e.g. Black et al. (2003), Brooks (2002), Dlaska & Krekeler (2009), Fulcher & Davidson (2007), Hattie & Timperley (2007), Hattie (2012), Hughes (2002, 2nd ed.), Stobbart & Gipps (1987, 3rd ed.), etc.

Bear in mind



- Importance of **teamwork**: collaborate with colleagues (across schools) when designing, developing, piloting or marking assessment
- Importance of a construct: **define** what it is you want to assess; **specify** (= describe, justify) your tools
- Importance of planning for intended impact and **positive washback**; anticipating potential unintended effects
- Importance of **piloting**, also in the classroom: e.g. ask colleagues, other schools; collect data over the years; do simple statistical analyses with excel
- Importance of **'research'**: discuss assessment, criteria with your students: how students perceive it, and how students solved your tasks (insight into validity); systematically **monitor and revise** your instruments

Overview



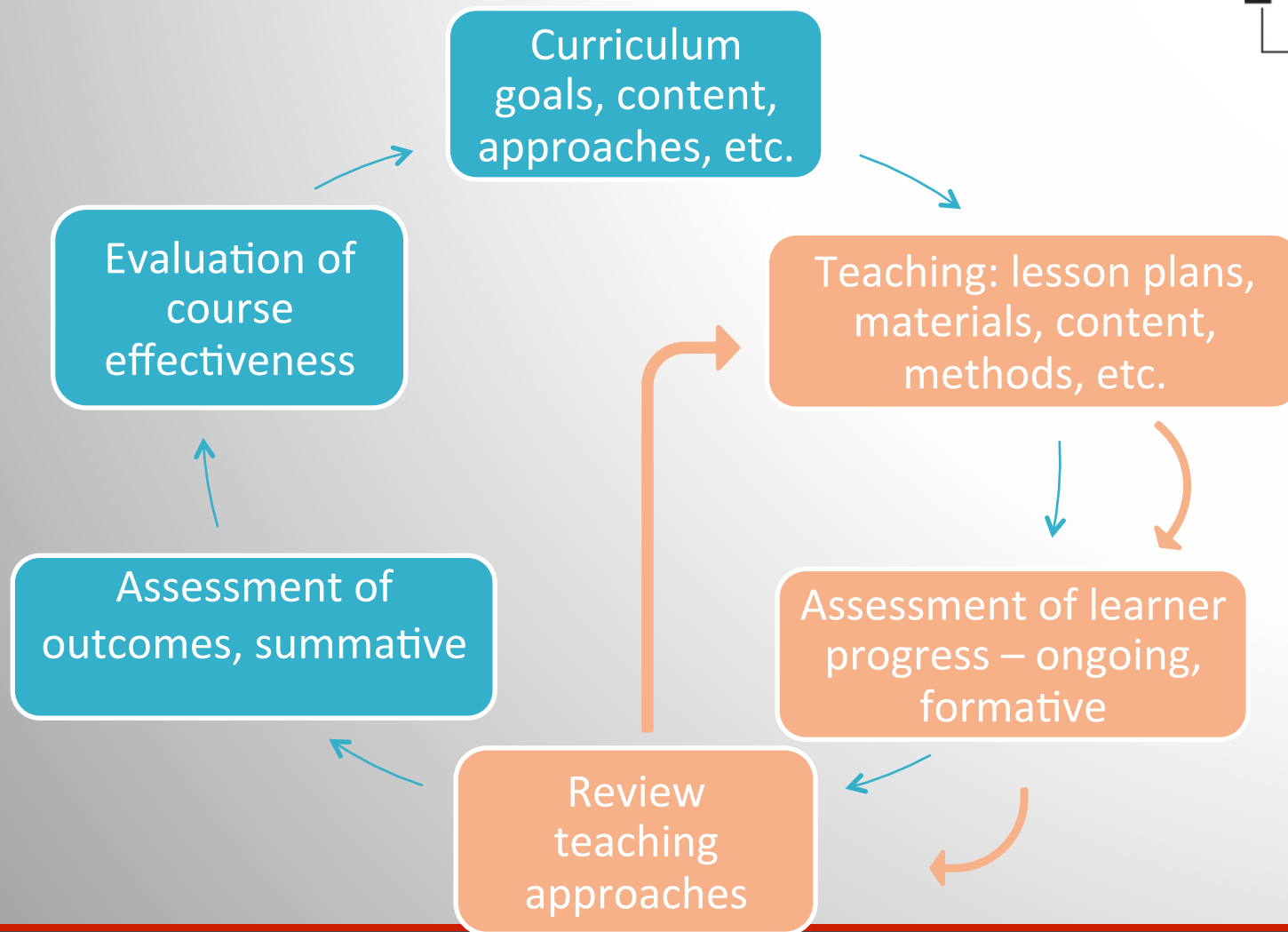
- Basic concepts
- Assessment purposes, use and interpretation
- Educational frameworks and reference criteria
- Principles of 'good' assessment practice
- Assessment tools and approaches
- Designing assessment in line with 'good principles' for the language classroom
- **Integrating teaching, assessment and evaluation**

Teaching – Assessment – Evaluation



- Objectives – curriculum, educational standards
- Teaching – methods, materials, goals
- Learning – processes and outcomes
- Assessment – purpose, formats, approaches
 - formative; summative
 - achievement; progress; proficiency
 - feedback on teaching effectiveness, learning progress
- Evaluation of teaching effectiveness:
Different purposes, approaches, formats

Evaluative Cycle





Thank you for your attention!

Dr. Claudia Harsch
Associate Professor
The Centre for Applied Linguistics
University of Warwick
Coventry CV4 7AL
Tel. 024765 75912
C.Harsch@warwick.ac.uk

Selected References

- Ainslee, S. (2004). Measuring Learning. In Harnisch H. & Swanton P. (Eds.) *Adults learning languages: A CILT guide to good practice*. 93-112
- Bachmann, L. & Palmer, A. (2010). *Language Assessment in Practice*. Oxford: OUP.
- Bachmann, L. (2004). *Statistical Analyses For Language Assessment*. Cambridge: CUP.
- Bailey, K., & Brown, J. D. (1996). Language testing courses: What are they? In A. Cumming & R. Berwick (Eds.), *Validation in language testing* (pp. 236-256). Clevedon, UK and Philadelphia: Multilingual Matters.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it Into Practice*. Buckingham, U.K.: Open University Press.
- Brooks, V. (2002). *Assessment in Secondary Schools: The New Teacher's Guide to Monitoring, Assessment, Reporting and Accountability*. Buckingham: Open University Press.
- Council of Europe (2011). *Common European Framework of Reference for Language Learning and Teaching. Language Examination and Test Development*. Strasbourg. Online: http://www.coe.int/t/dg4/linguistic/ManualLangageTest-Alte2011_EN.pdf
- Council of Europe (2001). *A Common European Framework of Reference for Language Learning and Teaching*. Cambridge: CUP.
- Dlaska, A., & Krekeler, C. (2009). *Sprachtests. Leistungsbeurteilungen im Fremdsprachenunterricht evaluieren und verbessern*. Hohengehren: Schneider.

Selected References

- Douglas, D. (2010). *Understanding Language Testing*. London: Hodder.
- Emmrich, R & Dietrich, S. (2011). Vergleichsarbeiten schreiben und dann? Zum Umgang mit Rückmeldungen aus Vergleichsarbeiten im Fremdsprachenunterricht. In: *Praxis English* 3/2011, 49-53.
- Fulcher, G., & Davidson, F. (2007). *Language Testing and Assessment: An Advanced Resource Book*. London and New York: Routledge.
- Harsch, C. (2012). Der Einsatz von Sprachtests in der Fremdsprachenforschung: Tests als Untersuchungsgegenstand und Forschungsinstrument. In S. Doff (Ed.), *Empirisch basierte Methoden in der Fremdsprachenforschung: Eine Einführung* (pp. 150-183). Tübingen: Narr.
- Harsch, C., & Schröder, K. (2010). Hoffnungsvoller Aufbruch. Neue Zeiten für Leistungserhebung, Leistungsmessung und Benotung. *Praxis Englisch*(3), 44-46.
- Hattie, J. and H. Timperley. (2007). The Power of Feedback. *Review of Educational Research* 77/ 1, 81–112.
- Hattie, J. (2012). *Visible learning: Maximizing Impact on Learning*. London: Routledge.
- Henning, G. (1987). *A guide to language testing: Development, evaluation, research*. Rowley, MA: Newbury House.

Selected References

- Hughes, A. (2002). *Testing for Language Teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402.
- Lynch B. (2003). *Language assessment and programme evaluation*. Edinburgh: Edinburgh University Press.
- McNamara, Tim. (2000). *Language Testing*. Oxford: Oxford University Press.
- Paran, A., & Sercu, L. (Eds.). (2010). *Testing the untestable in language education*. Bristol: Multilingual Matters.
- Stobart, G. & Gipps, C. (1997). *Assessment: A Teacher's Guide to the Issues* (3rd ed.). London: Hodder & Stoughton.
- Taylor, L. (2009). Developing Assessment Literacy. *Annual Review of Applied Linguistics*, 29, 21-36.
- Watanabe, Y. (2011). Teaching a course in assessment literacy to test takers: Its rationale, procedure, content and effectiveness. *Cambridge ESOL Research Notes*(46), 29-34.

Recommendations

The following books are recommended for assessing different language skills:

Alderson, J.C. (2000): *Assessing Reading*. Cambridge: CUP.

Buck, G. (2001): *Assessing Listening*. Cambridge: CUP.

Luoma, S. (2005): *Assessing Speaking*. Cambridge: CUP.

Weigle, S.C. (2001): *Assessing Writing*. Cambridge: CUP.

The following practical resources are highly recommended:

http://www.brookes.ac.uk/services/ocsd/2_learntch/briefing_papers/methods_assessment.pdf

<http://www.lancs.ac.uk/fass/projects/examreform/Pages/Projects.html>

Here, you find calibrated examples of language tests for reading, use of English, listening, speaking and writing, together with helpful discussions of assessment issues. The four book volumes are available as pdf online:

<http://www.lancs.ac.uk/fass/projects/examreform/Pages/Exams.html>

URLs of assessment associations:

ALTE: <http://www.alte.org/>

EALTA: <http://www.ealta.eu.org/>; <http://www.ealta.eu.org/guidelines.htm>