

Allgemeine Überlegungen zur Leistungsmessung im kompetenzorientierten Spanischunterricht

Nicht alles, was gezählt werden kann, zählt und nicht alles was zählt, kann gezählt werden.

Albert Einstein

Competence, like truth, beauty and contact lenses, is in the eye of the beholder (*Betrachter*).

Laurence J. Peter

1. Was wird beurteilt: Performanz und/oder Kompetenzen?

- Kompetenzen können nicht direkt getestet werden, sie zeigen sich im Gebrauch (Performanz).
- Geht es um die **Beurteilung der Performanz**, wird von den SuS verlangt, dass sie in einem direkten Test Beispiele ihrer (schriftlichen/mündlichen) Sprachproduktion zeigen. (vgl.: GER Deskriptoren Kapitel 4)
- Geht es um die **zugrunde liegenden sprachlichen Kenntnisse/Kompetenzen** eines S insgesamt, so lassen durch die Bearbeitung eines breiten Spektrums von Aufgabentypen Rückschlüsse darauf gewinnen. (vgl: GER Deskriptoren Kapitel 5)

2. Probleme bei der Leistungsmessung im Bereich mündlicher Sprachmittlung

- grundsätzliches Problem der Fremdsprachen mit den Gütekriterien für Tests:

Offene, integrative Aufgaben besitzen ein hohes Maß an Validität (Gültigkeit), sind aber wenig reliabel und objektiv.

Geschlossene und halboffene Aufgaben dagegen haben ein hohes Maß an Reliabilität (Zuverlässigkeit) und Objektivität (und werden deshalb oft als Testformate eingesetzt), sind aber nicht wirklich valide.
- Authentisch ist die mündliche Leistungsmessung in der Fremdsprache dann, wenn sich der Lernende als Sprecher/Hörer in einer sprachlichen Situation bewähren muss. Dazu braucht es möglichst realistische, also offene Aufgabenstellungen.
- Mündliche Handlungssituationen wie Simulationen und Dolmetschaufgaben sind interaktiv und integrativ (bilden „echtes“ Kommunikationsverhalten ab, enthalten spontane, nicht planbare Impulse). Sie sind in hohem Maße valide, aber wenig reliabel und objektiv.
- Das macht die Bewertung simulierter Kommunikationsformen schwierig.
- Zusätzliche Schwierigkeit bei der Sprachmittlung: Je nach Art der Aufgabe, des situativen Kontextes, der Intention der Beteiligten, der Rollenbeschreibung und Rollenverteilung, der Textsorte des Ausgangs- bzw. Zieltextes müssen die S ganz Unterschiedliches leisten. (Konstruktvalidität?)
- Ein Beispiel für die Vielfalt: Blick auf mögliche Kombinationen von Textsorten mit Arten der Sprachmittlung:

Ausgangstext D/SN	Art der Sprachmittlung D/ SN / beides	Zieltext D/SN
<p><u>Schriftlicher Text</u> z.B: Sachtext Fiktionaler Text Brief e-Mail Diagramme</p> <p><u>Mündlicher Text</u></p> <p>interaktiv z.B: Dialog Gespräch Interview Diskussion</p> <p>nicht interaktiv z.B: Rede Stadtführung Lautsprecher- durchsage TV/Radiosendung</p>	<p><u>Mündlich – schriftlich</u> Textnahe Übersetzung Zusammenfassung (global / selektiv) Freie Übertragung</p> <p><u>Schriftlich – mündlich</u> Textnahe Übersetzung Zusammenfassung (global / selektiv) Freie Übertragung</p> <p><u>Mündlich-mündlich</u> Dolmetschen Zusammenfassung (global / selektiv) Freie Übertragung</p> <p><u>Schriftlich-schriftlich</u> Textnahe Übersetzung Zusammenfassung (global / selektiv) Freie Übertragung</p>	<p><u>Schriftlicher Text</u> z.B: Sachtext Fiktionaler Text Brief e-Mail</p> <p><u>Mündlicher Text</u></p> <p>interaktiv z.B: Dialog Gespräch Interview Diskussion</p> <p>nicht interaktiv z.B: Rede Stadtführung Lautsprecher- durchsage TV/Radio- sendung</p>

-
-

- Dazu kommen interkulturelle, personale, interpersonale Faktoren u.a.m
- Fazit: Mündliche Sprachmittlung ist schwer testbar. Sie gehört aber zum Kernbestand des FU, hat für den außerschulischen Kontext (gerade in Spanisch) eine hohe Relevanz und birgt ein großes Motivationspotential. Außerdem bietet sie hervorragende Ansatzpunkte um das spontane dialogische Sprechen zu üben.
- Frage: Reicht es, sie in Lernaufgaben zu berücksichtigen und auf die Leistungsmessung zu verzichten?

3.Arbeit mit Kriterienrastern/Skalen/Checklisten* und Deskriptoren

*Zur Unterscheidung s. Europarat: Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen, Straßburg, 2001, S. 184

- Die **Qualität** der erbrachten mündlichen Leistung muss eingeschätzt werden (*rating= Einschätzung der Qualität von Lernertexten*) – diese Einschätzung kann naturgemäß nur subjektiv ausfallen.
- **Kriterienraster (Skalen/Checklisten)** definieren mit Hilfe verbaler Beschreibungen die Anforderungen auf verschiedenen Leistungsniveaus.

- Sieermöglichen die Überführung nicht-systematischer, subjektiver Einschätzungen in reflektierte Urteile („gelenktes Urteil“).
- Sie können dadurch dazu beitragen, unter Kolleg/innen einen gemeinsamen Bezugsrahmen für die Leistungsbewertung zu schaffen.
- „Die Erfahrung (...) war, dass weder ein Addieren von Fehlern noch die präzisesten Kriterien Garanten für Objektivität sein können, sondern dass ein **ständiger kollegialer Austausch über Anwendung und Definition von Kriterien stattfinden muss.**“

Schinschke, Andrea: Korrigieren nach Schablone – geht das?, in: Praxis Fremdsprachenunterricht, Heft 4/2008, S. 21 – 25, Zitat S. 22

- **Fazit: Deskriptoren** sind die Grundlage für die Bewertung, sie verringern die Subjektivität der Einschätzung, **aber:** sie müssen **vergleichbar** eingestuft/interpretiert werden.

4. Vorteile der Arbeit mit Bewertungsrastern

- erhöht die Transparenz von Qualitätsmaßstäben
- SuS entwickeln Qualitätskriterien für ihre eigenen Produkte und die von Mitschüler/innen
- Übersicht über eigene Stärken und Schwächen und die Leistungsentwicklung (z.B: Praxis am Institut Beatenberg: Klebepunkte auf den Kompetenzrastern)
- Folge: SuS bauen im Zuge der Selbstevaluation ein aktives Verhältnis zum eigenen Lernprozess auf. Sie lernen ihre Lernfortschritte selbst zu erkennen, Folge: positives Gefühl der Selbstwirksamkeit = motivationssteigernder Effekt
- Förderung der Lernerautonomie: Die Definitionen für die einzelnen Niveaustufen ermöglichen es, sich erreichbare Ziele zu setzen. Die SuS bauen durch die Selbstevaluation ein aktives Verhältnis zur Frage von Kontrolle und Bewertung auf.

vgl. Schinschke, Andrea: Korrigieren nach Schablone – geht das?, in: Praxis Fremdsprachenunterricht, Heft 4/2008, S. 21 – 25, S. 24-25

5. Die Arbeit mit Deskriptoren

- Deskriptoren dienen als Entscheidungshilfe dafür, die Performanz der SuS auf der Bewertungsskala auf einem bestimmten Niveau (A1, A2...) einstuft (oft *summativ = Einschätzung des Lernerfolgs am Ende eines Kurses*) oder um zu entscheiden, ob ein bestimmtes Lernziel erreicht wurde (*formativ = kursbegleitende Lernfortschrittskontrolle*).
vgl: Gemeinsamer europäischer Referenzrahmen für Sprachen, S. 181

• Merkmale guter Deskriptoren

- „Die Erfahrung hat gezeigt, dass die Konsistenz, mit der Lehrende und Lernende die Deskriptoren interpretieren können, verbessert wird, wenn die Deskriptoren nicht nur beschreiben WAS die Lernenden tun können, sondern auch, WIE GUT sie es tun können.“
Europarat: Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen, Straßburg, 2001, S. 175

- Die Niveaubeschreibungen können auch für sich alleine stehen, d.h. die beschriebenen Unterscheidungen sollten real sein, nicht nur verbal
- Sie ermöglichen Ja/Nein Entscheidungen
- Sie sind konkret, klar und kurz
- Sie enthalten möglichst wenig Fachterminologie
- Sie beschreiben Merkmale, Verhalten oder Können auf positive Weise
- So können sie in Checklisten für die Selbstbeurteilung und für die Beurteilung durch Lehrende aufgenommen werden. (*Sie müssen allerdings ausreichender ausdifferenziert sein, nur dann können auch kleinere Fortschritte sichtbar gemacht werden – ein nicht zu unterschätzender motivierender Faktor*)

- **Problem negativ formulierter Deskriptoren**

„Es ist jedoch eine Schwäche der meisten existierenden Skalen, dass die Deskriptoren auf den unteren Niveaus oft negativ und im mittleren Bereich der Skalen normorientiert formuliert sind. Sie machen oft rein verbale Unterscheidungen zwischen den Niveaus, indem sie in aneinander grenzenden Beschreibungen lediglich ein oder zwei Wörter ersetzen, die außerhalb des Kontexts der jeweiligen Skala wenig aussagen.“

Gemeinsamer europäischer Referenzrahmen für Sprachen, Zitat S. 175

Gegenargument: Die unteren Niveaus zeichnen sich oft durch Fehlendes oder Fehlerhaftes aus, das für die Interimssprache dieser Lernenden kennzeichnend ist und deswegen auch zur Bewertung herangezogen werden kann.

- **Problem vager Deskriptoren:**

- Sie sehen zwar gut aus, aber jede/r interpretiert sie anders.

- **Probleme bei der Umsetzung im Unterrichtsalltag**

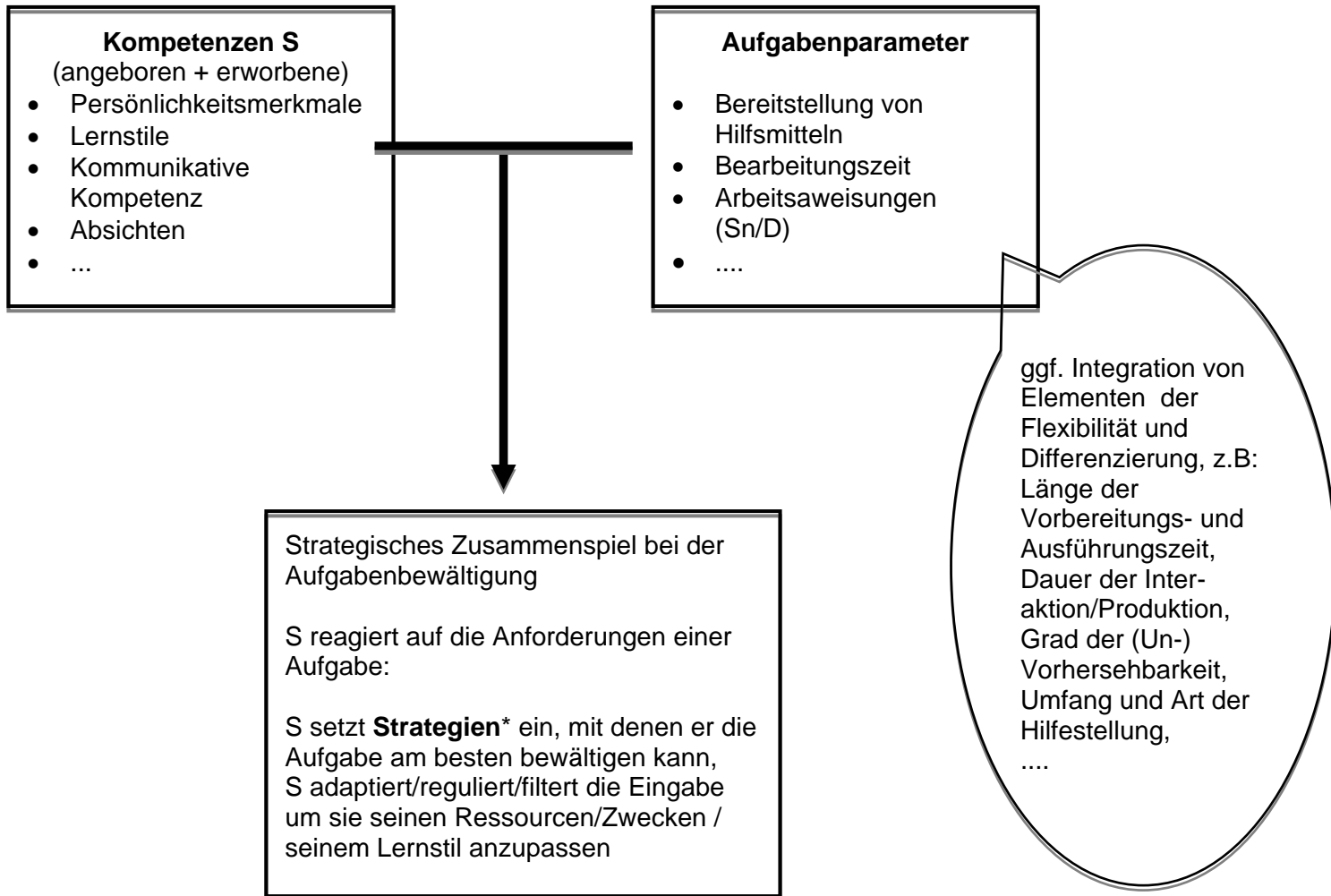
- Nicht alle relevanten Merkmale können durch kurze Deskriptoren beschrieben werden.
- Für jede Prüfungssituation müssen die jeweils relevanten Kriterien neu bestimmt werden. : Die Konstruktion valider *Rating-Skalen* ist langwierig und komplex.
- Die Praktikabilität einer Skala ist für den Unterrichtsalltag entscheidend.
- Folge: Normalerweise werden im Alltag schon vorhandene Skalen adaptiert.

6. Überlegungen zum Schwierigkeitsgrad von Aufgaben

Der Schwierigkeitsgrad einer Aufgabe lässt sich „nicht mit Sicherheit vorhersagen, am allerwenigsten für einen bestimmten Lerner“.

Gemeinsamer europäischer Referenzrahmen für Sprachen: Zitat S. 155

#



*„Strategien (allgemeiner sowie kommunikativer Art) stellen ein ganz wesentliches Bindeglied zwischen den verschiedenen (angeborenen oder erworbenen Kompetenzen eines Lernenden und der erfolgreichen Bewältigung der Aufgabe dar.“

Gemeinsamer europäischer Referenzrahmen für Sprachen: Zitat S. 155

Überlegungen zum Schwierigkeitsgrad von Sprachmittlungsaufgaben

- Sprachmittlung steht der Textproduktion nahe, es sind immer unterschiedliche Lösungen möglich, aber: im Bereich Inhalt besteht u.U. kaum Spielraum (*je nach Sprachstand schwierig für Lerner*).
- Die Gewichtung der Teilkompetenzen sollte je nach Aufgabenform und Anspruchsniveau verschieden sein.
- Es sind globalsprachliche Fertigkeiten gefordert, aber eine Bepunktung für Teilkompetenzen wird der Schülerleistung eher gerecht.

- Eine **analytische Beurteilung** verschiedener Aspekte der Sprachproduktion – in diesem Fall der Sprachmittlung - erlaubt eine entsprechende Rückmeldung an SuS und liefert die Metasprache für ein Gespräch über die Notenfindung.

7. Weitere Fragen zur Leistungsmessung im kompetenzorientierten Spanischunterricht

Passung Aufgabe – Bildungsstandards (*curriculare Validität*)?

Welchem Niveau entspricht die Leistung?

Vgl. hierzu die entsprechenden Beurteilungsraster zur mündlichen Kommunikation im Europäischen Referenzrahmen

Nach welchen Kriterien soll die Leistung bewertet werden?

Mögliche Kriterien (Zusammenschau der im Referenzrahmen und der vom ISB-Staatsinstitut für Schulqualität und Bildungsforschung München verwendeten Kriterien*:

Sprache	Strategie	Inhalt**
Aussprache und Intonation Grammatische Korrektheit Wortschatzbeherrschung Spektrum der sprachlichen Mittel <i>ISB: Idiomatik</i>	Flüssigkeit Sprecherwechsel Kooperation Flexibilität Um Klärung bitten <i>ISB: Kontrolle und Reparaturen, Kompensation</i>	Kohärenz Themenentwicklung Genauigkeit <i>ISB: Relevanz, Aufgabenerfüllung, Sachwissen</i>

***taucht in international anerkannten Sprachzertifikaten als Kriterium nicht auf – dort wird von Aufgabenerfüllung (taskachievement) gesprochen. Sachwissen ist integraler Bestandteil des Bildungsauftrages – anders als bei internationalen schul- und altersunabhängigen sowie lehrplanunabhängigen Sprachprüfungen, Inhalt lässt sich nicht auf den Nachweis von Sachwissen reduzieren, ist nicht gleichbedeutend mit dem Reproduzieren von Faktenwissen – dieses Wissen muss kohärent, argumentativ stringent und in klarem Bezug zur Aufgabenstellung erfüllt werden.*

**Vgl: Time totalk!... – Eine Handreichung zur Mündlichkeit im Unterricht der modernen Fremdsprachen, Institut für Schulqualität und Bildungsforschung München, S. 10*

Europarat: Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen, S. 18

8. Beispiel für ein Bewertungsraster (A2 Dolmetschen)

	Aussprache/ Intonation	Sprachliche Mittel
1	Müheles wirkende Aussprache und Intonation	Erweitertes Spektrum vertrauter und einfacher Strukturen durchgehend sicher verwendet
2	Natürlich wirkende Aussprache und Intonation; geringe Aussprachefehler	Erweitertes Spektrum vertrauter und einfacher Strukturen sicher verwendet
3	Gut verständliche Aussprache, erkennbarer muttersprachlicher Einfluss; gelegentliche Verstöße gegen die Phonologie beeinträchtigen die Verständlichkeit nur unwesentlich	Erweitertes Spektrum vertrauter und einfacher Strukturen meist korrekt verwendet
4	Starker muttersprachlicher Akzent sowie muttersprachliche Intonation; phonologische Fehler, die das Verständnis bisweilen erschweren; noch verständlich	Erweitertes Spektrum vertrauter und einfacher Strukturen nicht immer korrekt verwendet
5	Sehr starker muttersprachlicher Einfluß; häufige phonologische Fehler, die das Verständnis beeinträchtigen	Erweitertes Spektrum vertrauter und einfacher Strukturen mit sehr vielen Verstößen verwendet
6	Die Kommunikation ist aufgrund starker Mängel in einem oder mehreren Bereichen sehr stark beeinträchtigt oder nicht mehr gewährleistet	

	Aufgaben-/Rollenerfüllung; Inhalt	Strategie
4	Erfüllt die gestellte Aufgabe in vollem Umfang ; Verständigung jederzeit gewährleistet; Flüssige Wiedergabe auch komplexer Sachverhalte in der Fremdsprache	sicheres, flexibles Gesprächsverhalten zügiges Sprechen; erfolgreiche Selbstkorrektur; bei sprachlichen Unsicherheiten sehr oft Kompensation durch adäquate Ersatzwendungen
3	Erfüllt die gestellte Aufgabe weitgehend Verständigung in der Regel gewährleistet	gelegentliches Stocken; gelungene Überbrückung von Pausen/Lücken; sprachliche Defizite häufig durch adäquate Ersatzwendungen kompensiert
2	Erfüllt die gestellte Aufgabe trotz einiger Schwächen Verständigung noch gewährleistet	sprachliche Defizite zum Teil durch Ersatzwendungen kompensiert; Korrektur eigener Fehler nur bei Verständnisproblemen
1	Erfüllt die gestellte Aufgabe nur unzulänglich	häufiges Stocken, kaum Überbrückung von Pausen/Lücken Sprachliche Defizite kaum durch adäquate Ersatzwendungen kompensiert;
0	Die Kommunikation ist aufgrund starker Mängel in einem oder mehreren Bereichen sehr stark beeinträchtigt oder nicht mehr gewährleistet	