

Bewertungsinstrument für die Analyse von Materialien der Niveaustufen A1 bis B1 der Kurse „Deutsch als Fremdsprache“

Evaluation Tool for the Analysis of Materials from Levels A1 to B1 of the Courses “German as a Foreign Language”

Begoña Velasco Arranz¹ ·  <https://orcid.org/0000-0002-9169-4970>

Universidad de las Islas Baleares (UIB)

Cra. Valdemosa, km. 7.5. Edifici Arxiduc, E-07122 Palma

ZUSAMMENFASSUNG

Die Inhalte der Unterrichtsmaterialien entsprechen nicht notwendigerweise denen, die das Goethe-Institut für die Referenzniveaus A1 bis B1 Deutsch gemäß eigenen Definitionen verlangt. Das für Berufszwecke erforderliche Mindestniveau ist B1. Materialgestalter, z. B. große Verlage, haben bei der Reihenfolge der Einführung von neuen Inhalten eine gewisse Freiheit. Die Anwendung eines quantitativen Bewertungsinstruments auf drei Materialreihen verschiedener Verlage ergibt im Vergleich zu getrennten Auswertungen einer jeden Stufe insgesamt eine höhere Deckungsrate der vom Goethe-Institut geforderten Inhalte, wenn die von Stufe A1 bis Stufe B1 behandelten Inhalte gemeinsam ausgewertet werden. Dennoch haben die Ergebnisse gezeigt, dass keines der untersuchten Materialien die vom Goethe-Institut gestellten Anforderungen als Ganzes zufriedenstellend erfüllt.

Schlüsselwörter: Sprachunterricht, Lehrmaterial, Materialanalyse, GER-Niveaus, Bewertungsinstrument.

ABSTRACT

The contents offered by didactic material do not necessarily correspond to those required for levels A1 to B1 of the German language, according to the Goethe Institut definitions. The lowest required level for professional purposes is B1, which provides some freedom for material designers, such as large publishers, in terms of the order in which content is presented. By applying a quantitative assessment tool to three educational series of material from different publishers, it has been verified that the degree of content coverage required by the Goethe Institut increases if the contents included between levels A1 and B1 are computed together instead of individually. However, the results have revealed that none of the material studied meets the requirements of the Goethe Institute as a whole.

Keywords: language teaching, teaching material, materials analysis, CEFR levels, evaluation tool.

1. Einleitung

In einer zunehmend globalisierten und digitalisierten Gesellschaft muss der Lehrplan für den Sprachunterricht an die Bedürfnisse der Lernenden und Lehrenden für den beruflichen Einsatz der Sprache angepasst werden.

Die Lehrplantheorie bietet eine Anleitung für die Umsetzung des Lehrplans im Sprachunterricht. Die Auswahl und Organisation der Inhalte sollte auf der Grundlage der jeweils für die Lernenden definierten Ziele entwickelt werden (Stern, 1983: 434).

¹ Corresponding author · Email: begona.velasco@uib.es

Die von den großen Verlagen angebotenen Materialien sollen auf die Sprachprüfungen der verschiedenen Niveaustufen des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GER) vorbereiten, denn der Arbeitsmarkt verlangt genau das. Die Materialien sind jedoch rigoros und unflexibel (Calero Ramirez, 2017: 67) und werden somit den Anforderungen nicht gerecht, was eine Anpassung des Lehrplans an den jeweiligen Unterricht erschwert (Núñez Paris, 2008: 27). Sprachlehrkräfte müssen deshalb dasjenige Material auswählen, das die Lernenden am besten darauf vorbereitet, das für den Berufseinstieg erforderliche Sprachniveau zu erreichen. Außerdem werden die Lehrkräfte bei der Auswahl nicht systematisch unterstützt und der Entscheidungsprozess muss rationalisiert werden, um die Kontrolle über die Lehrplaninhalte zu gewährleisten.

Die in den einzelnen Stufen erforderlichen Deutschkenntnisse definiert das Goethe-Institut auf der Grundlage des GER für die Niveaus A1 bis C2 (Goethe-Institut, 2017). Lehrkräfte wissen jedoch nicht mit Sicherheit, ob die ausgewählten Materialien den von der Einrichtung festgelegten Zielen entsprechen und ob sie vollständig sind, beispielsweise in Bezug auf den Wortschatz oder die Grammatik.

Um den Lehrenden die Auswahl der Materialien zu erleichtern, wurde in einer Doktorarbeit ein Instrument zur Bewertung des Lehrmaterials entwickelt. „Das Instrument zur Analyse und Bewertung von Software für das Erlernen der deutschen Sprache“ (Herramienta de análisis y evaluación de software para el aprendizaje de la lengua alemana) wurde bereits zur Untersuchung von Materialien für die Stufe A2 Deutsch eingesetzt. Im Ergebnis ist zu konstatieren, dass die für das Bestehen der offiziellen Prüfung erforderlichen Lerninhalte nicht so umfassend abgedeckt werden, wie es das Goethe-Institut vorschlägt (Velasco Arranz, 2019). Da die analysierten Lehrmaterialien aus Materialreihen stammen, ergibt sich die Frage, ob dieses Ergebnis auch für die gesamte Materialreihe Gültigkeit besitzt.

Hierzu wird in diesem Beitrag die Anwendung des Instruments erweitert: Die drei Sprachreferenzniveaus A1, A2 und B1 werden als Ganzes analysiert, da B1 das für berufliche Zwecke am häufigsten verlangte Niveau ist. Indem Materialreihen untersucht werden, deren Sprachniveaus aufeinander aufbauen, soll festgestellt werden, ob sich die inhaltlichen Mängel der Ergebnisse der Analyse von A2 bestätigen. Jene Defizite könnten nämlich darauf zurückzuführen sein, dass sie in einer anderen Stufe erscheinen, denn Verlage geben oft für die offiziellen Prüfungen miteinander verknüpfte Materialien heraus.

Diese mögliche Erklärung soll in diesem Artikel mithilfe des entwickelten Bewertungsinstruments überprüft werden, ob die Inhalte der Materialien im Sinne der Prüfungsanforderungen des Goethe-Instituts vollständig sind und ob sie somit den Referenzniveaus des GER gerecht werden. Die Materialanalyse konzentriert sich auf die Überprüfung der Inventare, d. h. der Sprachinhalte, die vom Goethe-Institut für das offizielle Sprachzertifikat verlangt werden (Hennemann & Van der Werff, 2013: 49). Die Inventare gliedern sich in sechs Bereiche: 1. Sprachhandlungen, 2. Strategien, 3. Notionen (Allgemeine Begriffe), 4. Themen, 5. Wortschatz und 6. Grammatik. Für jedes Inventar wurden alle Inhalte, die das Goethe-Institut auf jedem der Sprachniveaus definiert, in die Datenbank des Bewertungsinstruments aufgenommen. Im Falle von Wiederholungen auf zwei oder drei Niveaus wurden die Inhalte mehrfach genannt und das niedrigste Nennungslevel gekennzeichnet.

Es stellen sich die folgenden Fragen:

- (1) Ist die Materialreihe angemessen und vollständig?
- (2) Entspricht die Materialreihe als Ganzes den Anforderungen des Goethe-Instituts über alle eingeschlossenen Referenzstufen hinweg?

Verschiedene Autoren, die sich mit der Bewertung von Materialien befassen haben (Aguilar Juárez & Morales Gamboa, 2011; Cova et al., 2008; Galvis, 1992; Marquès Graells, 2002; Ruiz González et al., 2010; Soto &

Gómez, 2000), ordnen und gruppieren Bewertungskriterien auf ähnliche Weise und fassen Kriterien zu grösseren konzeptuellen Einheiten zusammen. Auch bei der Entwicklung des hier angewendeten Instruments wurden die Kriterien aufgrund ihrer Eigenschaften gruppiert, so dass sich vier Blöcke ergeben: 1. Obligatorische Inhalte, 2. Didaktische und psycho-pädagogische Aspekte, 3. Software-Funktionen und 4. Andere Aspekte, die nicht mit dem Lernen zusammenhängen, wie zum Beispiel Kosten. Die als Bewertungselemente identifizierten Kriterien wurden in ihrer Gesamtheit einer Expertenvalidierung unterzogen. Für den Zweck dieses Artikels spielt nur derjenige Teil des Instruments eine Rolle, der sich auf den obligatorischen Inhaltsblock bezieht, da das Goethe-Institut zu den anderen Kriterienblöcken keine Vorgaben macht.

Die Art der Analyse der Materialien ist quantitativ. Gemessen wird, inwieweit die Materialien die vom Goethe-Institut geforderten Inhalte in prozentualen Anteilen abdecken.

2. Erstellung des Bewertungsinstruments

Es handelt sich um ein Beurteilungsinstrument zur Analyse von Materialien, die Lernende auf eine offizielle Sprachstufenprüfung vorbereiten. Ausführlich hierzu in Velasco Arranz (2019).

Die Erstellung basiert auf der Scoring-Methode oder Nutzwertanalyse, da diese verschiedene Messkriterien, die beispielsweise auf Ordinal- oder Intervallskalen basieren, integrieren kann. Den verschiedenen Aspekten, Kriterien und jeweiligen Items werden Punkte zugeordnet, um die zur Auswahl stehenden Optionen vergleichbar zu machen (Rommelfanger & Eickemeier, 2002: 159). Von den bisher entwickelten Bewertungsformaten wird für die Analyse von Materialien die Produkttestmethode (Halbes, 2003) gewählt; sie funktioniert auf der Grundlage eines am Schulnotensystem orientierten Punktesystems. Diese Methode ermöglicht die korrekte Beurteilung und Messung eines Produkts oder auch einer Dienstleistung oder einer Kombination aus beiden. Nichts anderes ist ein Material zum Erlernen der deutschen Sprache. Es wird hier als eine Mischung aus Produkt und Dienstleistung verstanden, die von Lernenden und Lehrenden in Anspruch genommen wird.

Das Punktesystem besteht darin, die verschiedenen Kriterien einzeln zu messen und jedes Maß in eine einheitliche Bewertungsskala umzuwandeln (siehe Abschnitt 2.2.2).

Die Hauptziele dieser Methode sind die Schaffung von Transparenz und Orientierung für die Verbraucher (Silberer, 1979). Die Leitlinien für die Erstellung von Kriterien und deren Analyse sind wissenschaftliche Fundierung, Einfachheit und Relevanz, wobei die Ergebnisse gültig, zuverlässig und handhabbar sein und wesentliche Punkte abdecken müssen (Cary & Sproles, 1978).

Der Produkttest zeichnet sich durch vier methodische Aspekte aus (Stiftung Warentest, 2014: 10):

- (3) Messung von Kriterien nach absoluten und relativen Aspekten
- (4) Erfüllung mit einem Minimum an absoluten Kriterien, vorausgesetzt, das Produkt wird nicht als ungeeignet angesehen
- (5) Vergabe der höchsten Punktzahl, wenn das Produkt keinerlei Mängel aufweist
- (6) Jede Bewertung wird gemäß des Wertes gewichtet, der sich aus der Bedeutung des gemessenen Kriteriums ergibt

Der Aufbau des Instruments ermöglicht es, Materialien zu untersuchen, zu beurteilen und miteinander zu vergleichen. Hiermit ist definiert, welche generellen methodischen Aspekte bei der Erstellung eines Messinstruments von Produkten und Dienstleistungen zu beachten sind. Es gilt nun zu klären, welche formalen Kriterien für die Erarbeitung und Auswahl der zu messenden Eigenschaften beachtet werden müssen.

2.1. Kriterien für die Entwicklung des Instruments

Das Instrument muss eine Reihe von Kriterien erfüllen, um die einheitliche Qualität der Entscheidung zu gewährleisten. Das Ergebnis muss nachprüfbar und unabhängig davon sein, wer das Material untersucht. Diese Anforderung an das Instrument ist in Analogie zu dem formuliert, was Popper (1977) in seiner Definition der Regeln wissenschaftlichen Arbeitens und Verifizierens formulierte.

Vor zwei Jahrzehnten definierten Eisenführ und Weber (1999: 60-62) die folgenden Kriterien, die ein Instrument erfüllen muss und die auch heute noch gültig sind:

- (7) Wiederholbar: gleiches Ergebnis, wenn verschiedene Personen bewerten
- (8) Transparent: verständlich und in der Berechnung nachvollziehbar
- (9) Einfach: die Beteiligten müssen das Instrument problemlos verstehen können
- (10) Zuverlässig: die Ergebnisse müssen das Material genau widerspiegeln und die Entscheidungsfindung gemäß den Präferenzen und Bedürfnissen des Entscheiders ermöglichen
- (11) Offen: die Auswahl der Kriterien und die Berechnung müssen Änderungen zulassen können

Das Instrument misst die Eignung des Materials anhand einer Reihe von Kriterien, die jedes bewertete Material mehr oder weniger erfüllt. Die Auswahl der Kriterien im Einzelnen und in ihrer Gesamtheit muss folgende Anforderungen erfüllen: Vollständigkeit, Redundanzfreiheit, Messbarkeit, Unabhängigkeit von Präferenzen und Einfachheit (Eisenführ & Weber, 1999). Definiert werden einerseits jeder Schritt in der Entwicklung und andererseits die Messung und Bewertung sowohl aller Kriterien einzeln als auch als Gesamtheit, und zwar innerhalb eines formalen Entscheidungsrahmens, den Bhushan und Rai (2004) einem Bewertungsinstrument abverlangen. Mehrere Autoren schlagen Methoden zur Definition eines Zielsystems, einer Zielhierarchie (Eisenführ & Weber, 1999), Zielgewichtungen (Rommelfanger & Eickemeier, 2002) und Möglichkeiten zur Gruppierung und Gewichtung von Optionen vor (Ishizaka & Nemery, 2013).

Die einzelnen Materialien repräsentieren die verschiedenen Optionen, zwischen denen eine Entscheidung getroffen werden muss. Sie werden nach einer Reihe festgelegter Kriterien bewertet und verglichen, die die Eigenschaften der Materialien repräsentieren müssen.

2.2. Festlegung der Kriterien

Es werden Merkmale erfasst und festgelegt, anhand derer beurteilt werden kann, inwieweit ein Material die Lernenden auf das Bestehen der Prüfung zur Erlangung des Sprachzertifikats vorbereitet. Die Auswahl und Gruppierung der Kriterien wird unter Anwendung der in Punkt 2.1 aufgestellten Auswahlkriterien vorgenommen. Dabei werden die Merkmale folgender Quellen hinzugezogen: aus dem GER und dem Goethe-Institut, aus den wichtigsten Theorien des Sprachenlernens, aus der Analyse einschlägiger Lernsoftware wie Moodle, aus anderen Bewertungsinstrumenten und aus vorhandenen Materialien.

Für die Analyse von Materialien der Niveaustufen A1 bis B1, die Gegenstand dieses Artikels sind, interessieren uns die Merkmale des GER und des Goethe-Instituts, da beide Institutionen die Richtlinien und Inhalte vorgeben, die man zur Erlangung des Sprachzertifikats beherrschen muss.

2.2.1. Vom GER und vom Goethe-Institut abgeleitete Kriterien

Durch eine systematische Analyse der Referenzniveaus für das Sprachenlernen (A1 bis B2) mit den entsprechenden Deskriptoren wurde eine Referenz für die Entwicklung des Instruments erstellt. Der GER liefert eine detaillierte Beschreibung der Ziele und Inhalte für jede Stufe des Sprachenlernens und schlägt einen handlungsorientierten Ansatz vor; dieser basiert darauf, die Bedürfnisse der Lernenden zu erfüllen, um Aufgaben im wirklichen Leben zu bewältigen, und er sieht die Lernenden als soziale Akteure, die mit anderen Personen interagieren (Council of Europe, 2018: 27).

Das Goethe-Institut, das sich am GER orientiert, legt die Lernziele für die deutsche Sprache mit Hilfe seiner Deskriptoren, den Kannbeschreibungen, fest. Die Ziele sind in den so genannten **Inventaren**, d. h. Sprachinhalten, festgehalten, die jeder Deutschlernende beherrschen muss, um seine Sprachkenntnisse nachzuweisen. Die Kriterien wurden unter Berücksichtigung der Anforderungen der offiziellen Stelle für Sprachprüfungen festgelegt. Ausgehend von den Inhalten, die das Goethe-Institut auf der Grundlage eines kommunikativen Ansatzes entwickelt hat (Hennemann & Van der Werff, 2013: 10), und den Richtlinien des GER werden in dem Bewertungsinstrument eine Reihe von Kriterien festgelegt. Sie sind in sechs Bereiche unterteilt:

- Sprachhandlungen
- Strategien
- Notionen (Allgemeine Begriffe)
- Themen
- Wortschatz
- Grammatik

Diese sechs Bereiche bilden die 6 Items des Blocks der Anpassung von didaktischen Inhalten des Bewertungsinstruments zur Materialanalyse.

2.2.2. Messkriterien und Bewertungsskala

Um ein Material zu bewerten, muss es gemessen werden, d. h. es muss mit einem Standard verglichen werden (Sierra Bravo, 1991: 120). Die Messung geschieht jedoch getrennt von der Analyse und Bewertung der einzelnen Materialien.

Aufgrund der unterschiedlichen Merkmale der Kriterien ist es nicht möglich, eine einzige Maßeinheit zu verwenden. Im vorliegenden Fall wurde die Übereinstimmungs- oder Erfüllungsrate des Materials gemessen, indem quantitative Kriterien verwendet werden. Diese Kriterien werden gemessen und als Prozentsatz der Erfüllung ausgedrückt. Beispiel Wortschatz: Hier wird verglichen, inwieweit der vermittelte Wortschatz mit dem in der Prüfung geforderten Wortschatz übereinstimmt.

Nach Messung der quantitativen Kriterien werden die Messergebnisse in eine simple, verständliche und einheitliche Bewertungsskala umgewandelt. Die am einfachsten zu verstehende und zu interpretierende Bewertungsskala ist die des Schulsystems (Schneider, 2007). Im spanischen Schulsystem wird ein Notensystem von 0 bis 10 verwendet. Um zu bestehen, muss mindestens die Hälfte der Punkte erreicht werden. Danach gibt es vier Hauptstufen der Differenzierung verschiedener Notenbereiche: sehr gut, gut, befriedigend und

mangelhaft. Für die Zwecke dieses Instruments wird diese Skala mit 10 multipliziert und die Materialien werden mit 0 bis 100 Punkten bewertet.

Das deutsche Schulsystem zeigt ein sehr ähnliches Muster. Obwohl ein „Nicht ausreichend“ jede Punktzahl sein kann, die unter der Hälfte der Punkte liegt, ist das „Nicht ausreichend“-Ergebnis immer dasselbe. Dies führt zu der Schlussfolgerung, dass es eigentlich fünf Stufen der Schulbenotung gibt:

- 1 = Das Material bereitet in besonderer Weise und ohne Lücken auf die Prüfung vor.
- 2 = Das Material bereitet gut auf die Prüfungsanforderungen vor.
- 3 = Das Material bereitet zufriedenstellend vor.
- 4 = Das Material bereitet nicht zufriedenstellend vor.
- 5 = Das Material bereitet sehr unzureichend vor.

Ein Vergleich zwischen dem spanischen und dem deutschen Notensystem (Velasco Arranz, 2019: 120-124) zeigt, dass die Äquidistanz der Punkte zwischen den verschiedenen „Ausreichend“-Niveaus gleich groß sind. Die Höchstnote wird auch bei einer Punktzahl vergeben, die geringfügig unter der Gesamterfüllung der Anforderungen liegt, ungeachtet der Tatsache, dass für bestimmte Zwecke wie z. B. den Hochschulzugang auf die exakte Punktzahl zurückgegriffen werden kann. Die Bewertung, die sich aus dem Vergleich der spanischen und deutschen Benotungssysteme für jedes der fünf Niveaus ergibt, lautet wie folgt:

- | | | |
|----------------|---|-----|
| • sehr gut | = | 100 |
| • gut | = | 83 |
| • befriedigend | = | 67 |
| • ausreichend | = | 51 |
| • ungenügend | = | 0 |

Das Notensystem des Goethe-Instituts vergibt jedoch nur volle Noten. Um die Note „Sehr gut“ zu erhalten, müssen 85 % der Prüfungsfragen richtig beantwortet werden. Wendet man dieses Kriterium auf alle von einem Material angebotenen Inventare im Vergleich zu den Anforderungen des Goethe-Instituts an, so wird ein Material, das 85% eines bestimmten Inventars anbietet, als „ausreichend“ bewertet. Mit einer geringeren Punktzahl ergäbe sich ein „Ungenügend“ bzw. „Nicht ausreichend“. In Äquidistanz und unter Berücksichtigung der Tatsache, dass ein „Sehr gut“ nach dem deutschen Notensystem mit einer richtigen Antwort unter 100 % erreicht wird, werden die Bereiche für ein „Ausreichend“ mit einem Prozentsatz zwischen 85 % und 88,99 %, für ein „Befriedigend“ zwischen 89 % und 92,99 %, für ein „Gut“ zwischen 93 % und 96,99 % und für ein „Sehr gut“ zwischen 93 % - 100 % definiert (Velasco Arranz, 2019).

2.3. Validierungsprozess

Die im Instrument entwickelten Items, die die Eignung der Materialien für den Deutschunterricht messen, wurden mit der Technik der Validierung durch Expertenurteile überprüft (Cabero Almenara & Llorente Cejudo, 2013: 16). Die Expertenbefragung verfolgt zwei Ziele: die Überprüfung der Gültigkeit und die Gewährleistung der Zuverlässigkeit des Instruments (Robles Garrote & Del Carmen Rojas, 2015: 130).

Der für die Validierung entwickelte Fragebogen enthält 77 in vier Blöcke unterteilte Items, von denen wir uns in dieser Studie den 6 Items des ersten Blocks widmen, bei denen es sich um die vom Goethe-Institut definierten Inventare handelt.

Bei der Erstellung des Plans für die Auswahl der Experten wird darauf geachtet, dass sich zwei Wissensgebiete überschneiden: Sprachunterricht und Pädagogik. Es wird eine Stichprobe von 10 Experten durchgeführt, von denen fünf aus dem Bereich Sprachunterricht und fünf aus dem Bereich Forschung, Fachgebiet Bildungstechnologie, stammen. Die erhobenen Bewertungen werden nach Gruppen analysiert und die Ergebnisse der einen Gruppe werden denen der anderen gegenübergestellt. Dies ist eine Antwort auf die ambivalente Natur der Methode des Expertenurteils zwischen einer quantitativen und einer qualitativen Methode (Thangaratinam & Redman, 2005: 123) und zwischen möglichen Fehlern von nicht-anonymen Erhebungen und unter Personen, die die Forschenden persönlich kennen.

Es werden zwei Gruppen aufgestellt:

- Gruppe A: Auswahl von Sprachlehrenden nach bestimmten obligatorischen Kriterien, wie z. B. aktive Hochschullehrende, Fremdsprachenlehrende für Deutsch oder Englisch, Abschluss in Fremdsprachen-Philologie und Linguistik, sowie nach bestimmten zusätzlichen Kriterien, wie Teilnahme an IKT-Projekten oder Tätigkeit als Prüfungslehrende am Goethe-Institut.
- Gruppe B: Auswahl des Expertengremiums, das dem Internationalen Forschungsgremium im Bereich Bildungstechnologie (PI2TE) angehört.

Das Validierungsinstrument ist vom Internationalen Forschungsgremium im Bereich Bildungstechnologie (PI2TE) unter <http://edutec.es/panel> validiert worden. Die Kommentare und Änderungsvorschläge sind in der Endfassung des Analyinstruments berücksichtigt worden.

3. Anwendung des Instruments auf die Materialien A1-B1

Das Instrument wird auf die drei Materialreihen mit ihren angegebenen Referenzniveaus angewandt, und es wird ein Vergleich zwischen ihnen angestellt. Dabei handelt es sich um:

Material 1 (M1) - Hueber Verlag

- (1) Schritte International neu A1
- (2) Schritte International neu A2
- (3) Schritte plus B1

Material 2 (M2) – Klett Verlag

- (1) Netzwerk A1
- (2) Netzwerk A2
- (3) Netzwerk B1

Material 3 (M3) – Cornelsen Verlag

- (1) Studio 21 A1
- (2) Studio 21 A2
- (3) Studio 21 B1

Die konkreten Inhalte werden vom Goethe-Institut für jedes der sechs Inventare im Referenzhandbuch Profile Deutsch Stufen A1-C2 (Glabionat et al., 2017: 68-91) definiert. In einem ersten Schritt werden diese in eine Excel-Tabelle übertragen. Das vom Goethe-Institut für die Stufen A1, A2 und B1 verwendete Format, das zwischen den drei Stufen ähnlich, aber nicht identisch ist, wird vereinheitlicht. Die sich in zwei oder drei Stufen wiederholenden Inhalte werden erfasst und die Stufe, in der die Inhalte zum ersten Mal erwähnt werden, sowie die oberen Stufen, in denen sie erscheinen, werden registriert.

Die entsprechenden Listen werden mit den auf jeder Stufe angebotenen Materialinhalten verglichen. Es wird sowohl die Verwendung eines Inhalts in einem Material der erforderlichen Stufe als auch seine Verwendung in einer Stufe, in der der Inhalt nicht gefordert wird, erfasst.

Für die Analyse und Bewertung der Materialien werden die Inhalte wie folgt gemessen. Zunächst werden die vom Goethe-Institut geforderten Gesamtinhalte für jede Stufe ermittelt. An zweiter Stelle wird der Prozentsatz der Inhalte berechnet, die in mehr als einer Stufe verwendet werden, was wir als Wiederholung bezeichnen. Drittens wird der prozentuale Anteil einmaliger Inhalte, d. h. solcher, die nur in einem Niveau erforderlich sind, ermittelt. Viertens wird der Anteil der Wiederholungen des niedrigsten vom Goethe-Institut geforderten Niveaus berechnet. An fünfter Stelle wird der Anteil der neuen Inhalte jeder Stufe gemessen, also solcher, die im Goethe-Institut in den untersten Stufen nicht erscheinen. Am Ende wird die Gesamtpunktzahl berechnet. Tabelle 1 zeigt die Berechnung der vom Goethe-Institut geforderten Inhalte.

	Sprachniveau			
Wortschatz	A1	A2	B1	Gesamt
Auf dem Niveau erforderlich	803	1.323	929	3.055
% der mehrfach verwendeten Inhalte	81,6%	56,0 %	25,8%	53,6%
Bereinigter % pro Niveau	18,4%	44,0%	74,2%	46,4%
% der mehrfach verwendeten Inhalte auf 1. Niveau	81,6%	9,8%	0.0%	25,7%
% der neuen Inhalte auf Niveau	100,0%	53,7%	74,2%	72,1%
Neue Inhalte auf Niveau	803	711	689	2.203
Grammatik	A1	A2	B1	Gesamt
Auf dem Niveau erforderlich	68	50	31	149
% der mehrfach verwendeten Inhalte	30,9 %	42,0%	6,5 %	29,5%
Bereinigter % pro Niveau	69,1%	58,0 %	93,5%	70,5%
% der mehrfach verwendeten Inhalte auf 1. Niveau	30,9 %	0.0%	0.0%	14,1%
% der neuen Inhalte auf Niveau	100,0%	58,0 %	93,5%	84,6%
Neue Inhalte auf Niveau	68	29	29	126
Sprachhandlungen	A1	A2	B1	Gesamt
Auf dem Niveau erforderlich	48	49	37	134
% der mehrfach verwendeten Inhalte	83,3%	89,8%	48,6%	76,1%
Bereinigter % pro Niveau	16,7%	10,2%	51,4%	23,9%
% der mehrfach verwendeten Inhalte auf 1. Niveau	83,3%	4,1%	0.0%	31,3%
% der neuen Inhalte auf Niveau	100,0%	14,3%	51,4%	55,2%
Neue Inhalte auf Niveau	48	7	19	74

Strategien	Sprachniveau			
	A1	A2	B1	Gesamt
Auf dem Niveau erforderlich	25	25	29	79
% der mehrfach verwendeten Inhalte	64,0%	76,0 %	13,8%	49,4%
Bereinigter % pro Niveau	36,0%	24,0%	86,2%	50,6%
% der mehrfach verwendeten Inhalte auf 1. Niveau	64,0%	4,0 %	0.0%	21,5%
% der neuen Inhalte auf Niveau	100,0%	28,0%	86,2%	72,2%
Neue Inhalte auf Niveau	25	7	25	57
Begriffe	A1	A2	B1	Gesamt
Auf dem Niveau erforderlich	34	29	44	107
% der mehrfach verwendeten Inhalte	79,4%	93,1%	6,8%	53,3%
Bereinigter % pro Niveau	20,6%	6,9%	93,2%	46,7%
% der mehrfach verwendeten Inhalte auf 1. Niveau	79,4%	0.0%	0.0%	25,2%
% der neuen Inhalte auf Niveau	100,0%	6,9%	93,2%	72,0%
Neue Inhalte auf Niveau	34	2	41	77
Themen	A1	A2	B1	Gesamt
Auf dem Niveau erforderlich	47	56	47	150
% der mehrfach verwendeten Inhalte	34,0%	30,4%	10,6%	25,3%
Bereinigter % pro Niveau	66,0%	69,6%	89,4%	74,7%
% der mehrfach verwendeten Inhalte auf 1. Niveau	34,0%	5,4 %	0.0%	12,7%
% der neuen Inhalte auf Niveau	100,0%	75,0%	89,4%	87,3%
Neue Inhalte auf Niveau	47	42	42	131

Tabelle 1. Vom Goethe-Institut geforderte Inventare.

Beim Wortschatz werden insgesamt 3.055 Wörter verlangt, von denen 2.203 Wörter in einem der drei Niveaus neu sind. Die Differenz, also 852 Wörter, sind die vom Goethe-Institut in mindestens zwei Niveaus erwähnten Wörter. Der reale auf allen drei Niveaus erforderliche Wortschatz, 2.203 Wörter, entspricht 72,1% des in den drei Stufen angesammelten Wortschatzes, d. h., 27,9% des Wortschatzes wird wiederholt. Der Prozentsatz der Wiederholungen zeigt an, dass ein Teil des Inhalts vom Goethe-Institut in mehr als einer Stufe verlangt wird. Der Anteil an einmaligem Wortschatz pro Stufe zeigt den Teil des Inhalts, der nur in der angegebenen Stufe gefordert wird. In jedem Niveau beträgt die Summe der beiden Indikatoren 100%. Der prozentuale Anteil der Wiederholungen der 1. Stufe gibt die Stufe der Wortnennung durch das Goethe-Institut an. In Stufe A1 ist dieser Anteil identisch mit dem Prozentsatz der Wiederholungen und in B1 beträgt er immer 0. Der Prozentsatz neuer Inhalte gibt das Verhältnis der in dieser Stufe eingeführten Inhalte an, also derjenigen, die in einer unteren Stufe nicht verlangt werden. In Stufe A1 beträgt dieser Prozentsatz immer 100%. Dies sind nicht nur wichtige Indikatoren, sondern sie dienen auch der Überprüfung der Kohärenz der Datenbank.

In jedem der sechs Inventare wird ein deutlicher Anteil an wiederholten Inhalten festgestellt, wobei Sprachhandlungen mit 76,1% den höchsten und Themen mit 25,3% den niedrigsten Anteil aufweisen.

In einem ersten Schritt der Materialanalyse muss die Erfüllungsrate der Inhalte mit dem insgesamt Geforderten in jeder Stufe separat gemessen werden, wie in Tabelle 2 zu sehen ist.

Sprachniveau				
Wortschatz	A1	A2	B1	Gesamt
Auf der Stufe gefordert	803	1.323	929	3.055
M1	88,8%	50,5%	21,7%	51,8%
M2	84,9%	49,4%	30,9 %	53,0%
M3	76,1%	52,7%	18,6%	48,4%
Grammatik	A1	A2	B1	Gesamt
Auf der Stufe gefordert	68	50	31	149
M1	94,1%	86,0 %	71,0%	86,6%
M2	94,1%	84,0%	32,3%	77,9%
M3	86,8%	98,0 %	83,9%	89,9%
Sprachhandlungen	A1	A2	B1	Gesamt
Auf der Stufe gefordert	48	49	37	134
M1	75,0%	83,7%	54,1%	72,4%
M2	81,3%	73,5%	56,8%	71,6%
M3	58,3%	83,7%	45,9%	64,2%
Strategien	A1	A2	B1	Gesamt
Auf der Stufe gefordert	25	25	29	79
M1	84,0%	68,0%	65,5%	72,2%
M2	84,0%	64,0%	51,7%	65,8%
M3	80,0 %	88,0 %	58,6%	74,7%
Begriffe	A1	A2	B1	Gesamt
Auf der Stufe gefordert	34	29	44	107
M1	100,0%	93,1%	59,1%	81,3%
M2	82,4%	93,1%	63,6%	77,6%
M3	97,1%	72,4%	77,3%	82,2%
Themen	A1	A2	B1	Gesamt
Auf der Stufe gefordert	47	56	47	150
M1	74,5%	82,1%	14,9%	58,7%
M2	85,1%	87,5%	36,2%	70,7%
M3	78,7%	82,1%	31,9%	65,3%

Tabelle 2. Angebotenes Material nach Anforderungen je Stufe.

Die Analyse des Wortschatzes zeigt, dass nur das Material M1 der Stufe A1 den geforderten Mindestanteil von 85% erfüllt. Die Tendenzen in den einzelnen Stufen sind für alle drei Materialien ähnlich. Von den insgesamt 3.055 verlangten Wörtern sind die Anteile der Verwendung jedes Materials ähnlich und liegen zwischen 48,4 % und 53,0 %.

Bei der Grammatik ist die Erfüllungsrate insgesamt die höchste, während sie beim Wortschatz am niedrigsten ist. Auffällig niedrig ist bei der Grammatik mit 32,3% die Erfüllungsrate von M2 in Stufe B1. Dies erklärt bei M2 den mit 77,9 % geringsten Erfüllungsgrad der grammatikalischen Anforderungen des Goethe-Instituts im Vergleich zu M1 und M3. Alle Materialien liegen in Stufe A1 über 85%. In Stufe A2 liegen zwei der Materialien über 85 % und in Stufe B1 erreicht keines der Materialien das Minimum von 85 %. Insgesamt liegt nur bei

Grammatik der Durchschnitt der drei Niveaus bei M1 und M2 über 85 %. M2 erzielt in Grammatik ebenfalls seine höchste Punktzahl im Vergleich zu den anderen Inventaren.

Bei Sprachhandlungen wird in keiner Stufe das Minimum von 85 % erreicht. Bei Strategien wird es nur bei M3 in Stufe A2 erreicht. Das Ergebnis ist bei Notionen generell höher, aber über alle drei Stufen hinweg erreicht keines der Materialien im Durchschnitt 85 %. Und bei Themen ragen nur die Niveaus A1 und A2 von M2 heraus.

In der nachstehenden Tabelle 3 sind die Prozentwerte, die sich aus der obigen Tabelle 2 ergeben, als Ränge absteigend zwischen den Materialien, für jedes Material und für jedes Inventar dargestellt. Rang 1 entspricht dem höchsten Wert in jeder Zeile.

	Sprachniveau		
	A1	A2	B1
Wortschatz			
Auf dem Niveau erforderlich	803	1.323	929
M1	1	2	3
M2	1	2	3
M3	1	2	3
Grammatik			
Auf dem Niveau erforderlich	68	50	31
M1	1	2	3
M2	1	2	3
M3	2	1	3
Sprachhandlungen			
Auf dem Niveau erforderlich	48	49	37
M1	2	1	3
M2	1	2	3
M3	2	1	3
Strategien			
Auf dem Niveau erforderlich	25	25	29
M1	1	2	3
M2	1	2	3
M3	2	1	3
Begriffe			
Auf dem Niveau erforderlich	3.4	29	44
M1	1	2	3
M2	2	1	3
M3	1	3	2
Themen			
Auf dem Niveau erforderlich	47	56	47
M1	2	1	3
M2	2	1	3
M3	2	1	3

Tabelle 3. Ranking nach Niveaus des im verlangten Niveau präsentierten Materials.

Aus Tabelle 3 ist ersichtlich, dass die Erfüllungsquote mit einer einzigen Ausnahme am niedrigsten ist, und zwar immer in Stufe B1. Zwischen den Stufen A2 und A1 besteht mehr Gleichgewicht. Stufe A1 erhält zehnmal den Rang 1 und Stufe A2 achtmal. Bei Wortschatz und Themen sind die Ergebnisse für die drei Niveaus identisch. Bei Grammatik, Strategien und Notionen erhält Stufe A1 bei zwei Materialien Rang 1, wobei M1 immer Rang 1 erhält.

So übersteigt nur bei Grammatik die Erfüllungsrate auf allen gemessenen Niveaus das geforderte Minimum von 85 % bei zwei Materialien. Unter Anwendung der in Punkt 2.2.2. erörterten Skala decken die drei Materialien die geforderten Inhalte auf einem sehr niedrigen Niveau ab, sodass keine Garantie besteht, dass die Lernenden ein Mindestmaß an Inhalten lernen, um die Stufenprüfungen des Goethe-Instituts ohne zusätzliche Vorbereitung zu bestehen. Dies gilt für jede Stufe und insbesondere für die Gesamtheit der drei untersuchten Stufen.

In der folgenden Analyse soll geprüft werden, ob die Erfüllungsrate der Anforderungen des Goethe-Instituts ansteigt, sodass die Bewertung sich ändert. Dabei ist jedoch nur das tatsächliche Inventar für Stufe B1 zu berücksichtigen, das sich aus den Inhalten abzüglich ihrer Wiederholungen zwischen den Niveaus und sogenannten „einmaligen Inhalten“ zusammensetzt. Die anschließenden Analysen werden gemeinsam für die Niveaus A1, A2 und B1 durchgeführt.

In den Tabellen 4 bis 9 werden die Erfüllungsrate der einmaligen Inhalte in den Materialien M1 bis M3 der sechs Inventare verglichen, wobei die Inventare folgendermaßen abgekürzt werden:

- W = Wortschatz
- G = Grammatik
- H = Sprachhandlungen
- S = Strategien
- N = Notionen
- T = Themen

Die folgende Tabelle 4 zeigt die von Wiederholungen bereinigten einmaligen Inhalte. Jeder Inhalt wird in der Stufe gezählt, in der er zum ersten Mal erscheint. Es werden die gesamten Inhalte nach Niveaus aufgeschlüsselt gezeigt. Stufe A1 hat die höchste Anzahl von einmaligen Inhalten bei allen Inventaren. Die Niveaus A2 und B1 sind bei den Inventaren von Wortschatz, Grammatik und Themen ausgeglichen. In der Stufe A2 sind bei Sprachhandlungen, Strategien und Notionen nur wenige einmalige Inhalte vorzufinden.

	Inventar					
	W	G	H	S	N	T
Bereinigter Inhalt	2.203	126	74	57	77	131
A1	803	68	48	25	3.4	47
A2	711	29	7	7	2	42
B1	689	29	19	25	41	42

Tabelle 4. Einmalige Inhalte nach Inventar und Stufe.

Tabelle 5 zeigt die Anzahl der wiederholten Inventare, d. h. der Inventare, die vom Goethe-Institut in mehr als einer Stufe verlangt werden. In diesem Zusammenhang zeigt Tabelle 6 den Anteil der Inventare für jedes Material, unabhängig von der Stufe, in der es benötigt wird oder in der es präsentiert wird.

	Inventar					
	W	G	H	S	N	T
Bereinigter Inhalt	2.203	126	74	57	77	131
Mehrfach enthalten	3.055	149	134	79	107	150
% mehrfach verwendet	38,7%	18,3%	81,1%	38,6%	39,0%	14,5%

Tabelle 5. Verhältnis der wiederholten Inventare.

Tabelle 5 zeigt große Differenzen unter den Inventaren. Man würde erwarten, dass die Erfüllungsrate eines Materials mit den Anforderungen des Goethe-Instituts im Verhältnis zur Wiederholungsrate eines Inventars ansteigt. Handlungen weisen eine Wiederholungsrate von 81,1 % auf und man würde erwarten, dass der Erfüllungsgrad in den unterschiedlichen Niveaus höher wäre als zum Beispiel in Grammatik mit einer Wiederholungsrate von 18,3 %. Aber es besteht keine lineare Beziehung, wie später zu sehen ist. Grammatik hat den zweitniedrigsten Anteil wiederholter Inventare, aber die Erfüllungsrate ist von allen Inventaren die höchste.

In Tabelle 6 wird die Erfüllungsrate mit den Werten der letzten Spalte von Tabelle 2 verglichen. Der Ausschluss von wiederholten Inhalten hat den Effekt, dass M1 beim Wortschatz von 51,8 % auf 77,5 % ansteigt, genau wie es bei allen Materialien und Inventaren der Fall ist.

	Inventar					
	W	G	H	S	N	T
Bereinigter Inhalt	2.203	126	74	57	77	131
M1	77,5%	94,4%	90,5%	94,7%	88,3%	81,7%
M2	79,5%	96,0%	87,8%	93,0%	92,2%	78,6%
M3	68,8%	96,8 %	70,3%	91,2%	96,1%	79,4%

Tabelle 6. In jedweder Stufe präsentiertes einmaliges Inventar.

Tabelle 7 zeigt die Unterschiede in der Verwendungsrate der geforderten Inhalte; verglichen wird die Rate der ausschließlich in der ersten geforderten Stufe gemessenen Verwendung und die in jeder der drei Stufen gemessenen Verwendung. Die Erfüllungsrate steigt erheblich an, wenn auch mit deutlichen Unterschieden zwischen verschiedenen Inventaren und Materialien. Der geringste Anstieg wird in M3 in Grammatik verzeichnet (7,9 %) und der höchste in M1 in Themen (23,7 %).

	Inventar					
	W	G	H	S	N	T
Bereinigter Inhalt	2.203	126	74	57	77	131
M1	19,0%	6,3%	18,9 %	17,5 %	10,4%	23,7%
M2	22,3%	11,9%	10,8 %	22,8%	19,5%	8,4 %
M3	18,3%	7,9%	21,6%	17,5 %	9,1%	15,3%

Tabelle 7. Unterschiede bei der Messung einmaliger Inhalt in jeder Stufe.

Tabelle 8 zeigt die Punktzahlen, die die Materialien erhalten, wenn man nur den Inhalt berücksichtigt, der in jeder Stufe zum ersten Mal verlangt wird. Nur M1 und M3 bei Grammatik und M3 bei Notionen erfüllen den Mindestinhalt, allerdings mit niedrigen Punktzahlen.

	Inventar					
	W	G	H	S	N	T
Bereinigter Inhalt	2.203	126	74	57	77	131
M1	0	63	0	0	0	0
M2	0	0	0	0	0	0
M3	0	67	0	0	59	0

Tabelle 8. Bewertung der Materialien in der niedrigsten geforderten Stufe.

Tabelle 9 zeigt die Punktzahlen der auf der Grundlage der Ergebnisse von Tabelle 6 berechneten Bewertung.

	Inventar					
	W	G	H	S	N	T
Bereinigter Inhalt	2.203	126	74	57	77	131
M1	0	89	73	90	64	0
M2	0	95	62	83	80	0
M3	0	98	0	76	95	0

Tabelle 9. Bewertung der Materialien in jeder Stufe.

Tabelle 9 zeigt die Bewertung unter Einschluss der in anderen Niveaus verwendeten Inhalte. Die Erfüllungsrate steigt bei den drei genannten Inhalten deutlich an und der Anstieg in den anderen Fällen führt zu einem „Ausreichend“ aller drei Materialien bei Grammatik, Strategien und Notionen und zu einem „Ausreichend“ von M1 und M2 bei Handlungen. Die Anstiege bei Wortschatz und Themen und bei Sprachhandlungen von M3 sind zwar beträchtlich, reichen aber für ein „Ausreichend“ nicht aus. Tabelle 10 schließlich zeigt die Ergebnisse der einzelnen Inventare, ausgedrückt in Schulnoten.

	W	G	H	S	N	T
M1	sehr schlecht	sehr empfehlenswert	empfehlenswert	sehr empfehlenswert	empfehlenswert	sehr schlecht
M2	sehr schlecht	sehr empfehlenswert	nicht empfehlenswert	sehr empfehlenswert	empfehlenswert	sehr schlecht
M3	sehr schlecht	sehr empfehlenswert	sehr schlecht	empfehlenswert	sehr empfehlenswert	sehr schlecht

Tabelle 10. Noten nach Inventartyp.

Die Ergebnisse zeigen, dass alle drei bewerteten Materialien nur in Grammatik „Gut“ sind. M1 und M2 schneiden bei Strategien ebenfalls mit „Gut“ ab, aber M3 nur mit „Befriedigend“. Bei Notionen hingegen erhält M3 ein „Gut“, während M1 und M2 nur mit „Befriedigend“ abschneiden.

4. Schlussfolgerungen

In allen sechs Inventaren gibt es einen hohen Anteil der vom Goethe-Institut geforderten Inhalte in mehr als einer Sprachstufe. In der entsprechenden Dokumentation wird der Grund für die Wiederholungen nicht angesprochen (Glabionat et al., 2017).

Eine frühere Untersuchung der Stufe A2 derselben Materialreihe ergab für alle drei Materialien und Inventare ein niedriges Level bei der Inhaltserfüllung (Velasco Arranz, 2019: 178). Die vorliegende Arbeit bestätigt das gleiche Ergebnis für die Stufen A1 bis B1.

Durch die Bereinigung wiederholter Inhalte in mehr als einer Stufe steigt die Erfüllungsrate in allen drei Stufen deutlich an. Das einzige Inventar, das von allen Materialien gut behandelt wird, ist Grammatik, obwohl es auch nicht die besten Noten erhält. Wortschatz und Themen werden nicht annähernd ausreichend abgedeckt.

Im Idealfall spiegeln die Prüfungen zur Erlangung des Stufenzertifikats die Fähigkeit des Kandidaten wider, effektiv in der deutschen Sprache gemäß den Anforderungen des GER zu kommunizieren. Dies sollte in den Leitlinien für die Entwicklung der amtlichen Prüfungen berücksichtigt werden. Darüber hinaus sollten Leitlinien für die Angemessenheit der entsprechenden Kenntnisse und Fähigkeiten sicherstellen, dass der Deutschunterricht dieses Ziel verfolgt. Abgesehen von den wiederholten Zweifeln an der Angemessenheit der Richtlinien des Goethe-Instituts diesbezüglich, würde man von Materialien, die den Anspruch erheben, auf eine konkrete Sprachstufe vorzubereiten, eine umfangreiche Orientierung über die erforderlichen Inhalte erwarten. In den untersuchten Materialien sind aber erhebliche Mängel erkennbar.

Es bleiben Zweifel, ob die drei Materialien als repräsentativ angesehen werden können. Die Verlage selbst bieten eine breite Palette unterschiedlicher Materialien an, die auf eine bestimmte Sprachstufe vorbereiten. Es sollte auch darauf hingewiesen werden, dass das Ziel des vom GER eingerichteten Systems in erster Linie eine allgemeine Qualifizierung der Sprachstufe für den Zugang zum Arbeitsmarkt ist. Folglich gibt es nur eine Art von Prüfung und Anforderungen, unabhängig von der Vielfalt der angebotenen Studiengänge und abgesehen von dem Sonderfall der Hochschulzugangsberechtigung. Mit anderen Worten: Egal, welches Material die Studierenden verwenden und für welche Zwecke – der Test ist für alle derselbe.

In diesem Zusammenhang muss man sich fragen, warum die Materialinhalte so erheblich von den Anforderungen abweichen. Künftige Forschungen sollten sich auf die Entwicklung von Hypothesen und deren anschließende empirische Überprüfung konzentrieren. Ungeachtet der Zweifel an der Angemessenheit des Ansatzes des Goethe-Instituts sind dessen Anforderungen der Maßstab für die Prüfungen. Es muss davon ausgegangen werden, dass alle geforderten Inhalte die gleiche Wahrscheinlichkeit haben, geprüft zu werden, und bei der Berechnung der Note die gleiche Bedeutung haben. Allerdings liegen hierzu keine einschlägigen Studien vor.

Folgende Forschungsfragen werden vorgeschlagen. Das Goethe-Institut nimmt für sich in Anspruch, Inventare nach der Häufigkeit ihrer Nutzung, d. h. ihrer relativen Bedeutung für die alltägliche Kommunikation, auszuwählen. Möglicherweise identifizieren die Verleger zeitgemäße Inhalte treffender und orientieren sich stärker an der Realität als an den Vorgaben des Goethe-Instituts. Eine alternative Erklärung für das Abweichen der Inhalte kann an den Anforderungen der modernen Didaktik liegen, mit authentischen Materialien zu arbeiten. Es könnte schlichtweg sehr schwierig oder kostenintensiv sein, geeignete Materialien zu finden, die den geforderten Inhalten des Goethe-Instituts entsprechen. Es gilt also zu untersuchen, wie die Unterrichtsmaterialien erstellt werden, wie sehr die Vorgaben des Goethe-Instituts hierbei berücksichtigt werden und welche Kontrollmechanismen ggf. existieren, um eine weitgehende Übereinstimmung mit den Anforderungen zu gewährleisten.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

About the author

PhD in Educational Technology from the UIB and Bachelor of German Philology by the UVA. Since 2017, assistant professor at the UIB. Cooperation in teaching innovation projects such as 'Internet application for learning German for the UIB Faculty of Tourism in 2008, preparation and development of a German MOOC for tourism at the UIB (2019). Publications related to the teaching of Languages and New Technologies applied to the teaching of languages.

Literaturverzeichnis

- Aguilar Juárez, I., & Morales Gamboa, R. (2011). “Análisis de los Modelos de producción de CODIE (Contenidos Digitales Educativos)”. In T. Morales Reynoso, C. Serrano Barquín, J. M. Serrano García, D. A. Miranda Gracia, E. Estrada Laredo, P. Troche Hernández, & A. Rojas García (Eds.), *1er Coloquio Internacional Aplicaciones educativas de las Tecnologías de la Información y la Comunicación, Universidad de Guadalajara*, 569–583. https://www.academia.edu/11874006/Análisis_de_los_Modelos_de_producción_de_CODIE_Contentos_Digitales_Educativos_ [2.5.2020].
- Bhushan, N., & Rai, K. (2004). *Strategic decision making. Applying the Analytic Hierarchy Process*. London: Springer.
- Cabero Almenara, J., & Llorente Cejudo, M. C. (2013). La Aplicación del Juicio de Experto como Técnica de Evaluación de las Tecnologías de la Información y Comunicación (TIC). *Eduweb. Revista de Tecnología de Información y Comunicación En Educación*, 7(2), 11–22. <https://dialnet.unirioja.es/servlet/articulo?codigo=4857163> [12.3.2021].
- Calero Ramirez, C. (2017). Neue Medien im DaF-Unterricht: Theorie und Praxis zum Hörverstehenstraining mit Podcasts. *Informationen Deutsch Als Fremdsprache*, 38(1), 36–69. <https://doi.org/10.1515/infodaf-2011-0105>
- Cary, R. T., & Sproles, G. B. (1978). Evaluating Product Testing Methods: A Theoretical Framework. *Home Economics Research Journal*, 7(1), 66–75. <https://doi.org/10.1177/1077727X7800700201>
- Council of Europe. (2018). Common European framework of reference for languages: Learning, teaching, assesment. Companion volume with new descriptors. <https://www.coe.int/en/web/common-european-framework-reference-languages> [14.6.2020].
- Cova, Á., Arrieta, X., & Aular de Duran, J. (2008). Revisión de modelos para la evaluación de software educativos. *Revista Electrónica de Estudios Telemáticos*, 7(1), 93–114. <http://ojs.urbe.edu/index.php/telematique/article/view/1969/1866> [23.4.2021].
- Eisenführ, F., & Weber, M. (1999). *Rationales Entscheiden*. Berlin: Springer.
- Galvis, A. (1992). *Ingeniería de Software educativo*. Santa Fe de Bogotá: Ediciones Uniandes.
- Glabionat, M., Müller, M., Rusch, P., Schmidt, H., & Wertenschlag, L. (2017). *Profile Deutsch. Gemeinsamer europäischer Referenzrahmen. Lernzielbestimmungen, Kannbeschreibungen und kommunikative Mittel. Niveaustufen A1-A2, B1-B2, C1-C2*. Stuttgart: Klett Sprachen.
- Halbes, S. (2003). Der vergleichende Warentest zur Unterstützung des nachhaltigen Konsums Inhaltsverzeichnis. Lehr- Und Forschungsbericht, 52. <https://docplayer.org/184121440-Der-vergleichende-warentest-zur-unterstuetzung-des-nachhaltigen-konsums.html> [25.3.2020].
- Hennemann, D., & Van der Werff, F. (2013). Goethe-Zertifikat A2. Fit in Deutsch 2. Prüfungsziele. Testbeschreibung. Goethe-Institut. https://www.goethe.de/pro/relaunch/prf/ko/Pruefungsziele_Testbeschreibung_A2_Fit2.pdf [30.3.2020].
- Ishizaka, A., & Nemery, P. (2013). *Multi-Criteria Decision Analysis: Methods and Software*. John Wiley & Sons Ltd. <https://doi.org/10.1002/9781118644898>
- Marquès Graells, P. (2002). Evaluación y selección de software educativo. Comunicación y Pedagogía: *Nuevas Tecnologías y Recursos Didácticos*, 185, 31–37. <https://dialnet.unirioja.es/servlet/articulo?codigo=306631> [10.2.2018].

- Núñez Paris, F. (2008). Teoría del currículum y didáctica de las lenguas extranjeras. *Revista Nebrija de Lingüística Aplicada a La Enseñanza de Lenguas*, Número 4, 1–30. <https://doi.org/10.26378/mlael2496>
- Popper, K. (1977). *La lógica de la investigación científica*. Madrid: Editorial Tecnos.
- Robles Garrote, P., & Del Carmen Rojas, M. (2015). La validación por juicio de expertos: dos investigaciones cualitativas en Lingüística aplicada. *Revista Nebrija*, 9(18), 124–139. <https://doi.org/https://doi.org/10.26378/mlael918259>
- Rommelfanger, H. J., & Eickemeier, S. H. (2002). *Entscheidungstheorie*. Berlin: Springer Verlag.
- Ruiz González, R., Muñoz Arteaga, J., & Álvarez Rodríguez, F. (2010). Evaluación de Objetos de Aprendizaje a través del Aseguramiento de Competencias Educativas. In *Virtual Educa* (pp. 1–18). Universidad Nacional Autónoma de México. <http://repositorial.cuaieed.unam.mx:8080/xmlui/handle/20.500.12579/1187> [10.2.2020].
- Schneider, W. (2007). *Marketing*. Heidelberg: Physica-Verlag.
- Sierra Bravo, R. (1991). *Diccionario práctico de estadística*. Madrid: Editorial Paraninfo.
- Silberer, G. (1979). Ansatzpunkte und Probleme einer Erfolgsbeurteilung beim neutralen Gütertest. *Zeitschrift Für Verbraucherpolitik/Journal of Consumer Policy*, 2, 110–126. <https://doi.org/10.1007/BF02278744>
- Soto, F., & Gómez, M. (2000). EVALÚA: Un instrumento de evaluación de recursos multimedia para la atención a la diversidad. <http://diversidad.murciaeduca.es/tecnoneet/docs/2002/5-22002.pdf> [12.5.2018].
- Stern, H. (1983). *Fundamental Concepts of Language Teaching*. Oxford: Oxford University Press.
- Stiftung Warentest. Wie testet und bewertet die Stiftung Warentest? (2014). <https://www.test.de/unternehmen/testablauf-5017344-0/> [24.10.2019].
- Thangaratinam, S., & Redman, C. W. (2005). The Delphi technique. *The Obstetritian & Gynecologist*, 7, 120–125. <https://doi.org/10.1576/toag.7.2.120.27071>
- Velasco Arranz, B. (2019). *Elaboración de una herramienta de análisis y evaluación de software para el aprendizaje de la lengua alemana*. [Tesis doctoral, Universitat de les Illes Balears]. <https://www.educacion.gob.es/teseo/mostrarRef.do?ref=1913679>