



**Gemeinsamer  
europäischer  
Referenzrahmen  
für Sprachen:**  
lernen, lehren, beurteilen

**Begleitband**

Niveau A1·A2·B1·B2·C1·C2



## Anhang 6

# Entwicklung und Validierung der erweiterten Beispieldeskriptoren

Die Anhänge dieses Bandes sind digital zur Ansicht bzw. zum Download unter dem folgenden Link abrufbar:  
[www.klett-sprachen.de/GER-Begleitband-Downloads](http://www.klett-sprachen.de/GER-Begleitband-Downloads)

Die in diesem Werk angegebenen Links wurden von der Redaktion sorgfältig geprüft, wohl wissend, dass sie sich ändern können. Die Redaktion erklärt hiermit ausdrücklich, dass zum Zeitpunkt der Linksetzung keine illegalen Inhalte auf den zu verlinkenden Seiten erkennbar waren. Auf die aktuelle und zukünftige Gestaltung, die Inhalte oder die Urheberschaft der verlinkten Seiten hat die Redaktion keinerlei Einfluss. Deshalb distanziert sie sich hiermit ausdrücklich von allen Inhalten aller verlinkten Seiten, die nach der Linksetzung verändert wurden. Diese Erklärung gilt für alle in diesem Werk aufgeführten Links.

Alle Anfragen in Bezug auf Nachdruck oder Übersetzung aller Teile dieses Dokuments müssen gerichtet werden an: Directorate of Communication (F-67075 Strasbourg Cedex oder an [publishing@coe.int](mailto:publishing@coe.int)). Alle weitere Korrespondenz zu dieser Veröffentlichung sollte an die Education Policy Division, Education Departement, Council of Europe gerichtet werden.

© 2020 Council of Europe, Straßburg

© 2020 Goethe-Institut München für die deutsche Übersetzung

Diese deutsche Buchausgabe des Titels ist erschienen bei:

Ernst Klett Sprachen GmbH, Rotebühlstraße 77, 70178 Stuttgart 2020

[www.klett-sprachen.de](http://www.klett-sprachen.de)

Alle Rechte vorbehalten.

Übersetzung: Jürgen Quetz, Rudi Camerer

Projektkoordination und Administration Council of Europe: Ahmet Murat Kilic

Projektkoordination deutsche Buchausgabe: Dr. Michela Perlmann-Balme (Goethe-Institut),

Sebastian Weber (Ernst Klett Sprachen)

Redaktion & Korrektorat: Sibylle Krämer, Bayreuth

Layoutkonzeption: Sabine Kaufmann

Gestaltung und Satz: Joachim Schrimm, Frielzheim

Druck und Bindung: Elanders GmbH, Waiblingen

Printed in Germany

ISBN 978-3-12-676999-0

## Entwicklung und Validierung der erweiterten Beispieldeskriptoren

### Aktualisierung der Skalen von 2001

Die Skalen mit Beispieldeskriptoren aus dem Jahr 2001 sind einer der am häufigsten benutzten Aspekte des GeR, und die Relevanz der ursprünglichen Deskriptoren ist im Lauf der Jahre bemerkenswert stabil geblieben. Deshalb wurde der Ansatz gewählt, dieses System von 2001 zu ergänzen, statt Deskriptoren darin zu verändern. Es gibt jedoch bei einigen wenigen dieser Deskriptoren aus Kapitel 4 und 5 des GeR auch Vorschläge für erhebliche Veränderungen. Die Änderung einer kleinen Zahl von „absoluten“ Formulierungen auf C2 soll stärker reflektieren, dass die Beispieldeskriptoren des GeR nicht einen idealisierten Muttersprachler/eine idealisierte Muttersprachlerin als Bezugspunkt für die Kompetenz von Sprachverwendenden/Lernenden nehmen. Diese kleinen Änderungen sind im erweiterten System von Beispieldeskriptoren eingeschlossen, das hier veröffentlicht wird, und sind in Anhang 7 aufgelistet. Die gewählte Arbeitsweise begann mit einer kleinen Autorengruppe aus der *Eurocentres Foundation*, welche relevante kalibrierte Materialien aus den im Vorwort des Begleitbandes zitierten Quellen auswählte, einfügte und, wenn nötig, adaptierte. In einer Reihe von Treffen mit einer kleinen Gruppe von Expertinnen und Experten, die als Sondierungsgremium agierten, wurde das so entstandene System von Deskriptoren präzisiert, bevor es einer größeren Gruppe von Beratern und Beraterinnen zur Begutachtung vorgelegt wurde.

### Neue Skalen

In dieser Phase des Projekts wurden neue Skalen für *Lesen als Freizeitbeschäftigung* (unter „Leseverstehen“), für *Telekommunikationsmittel benutzen* (unter „mündliche Interaktion“) und für *Zusammenhängendes monologisches Sprechen: Informationsvermittlung* (unter „mündliche Produktion“) hinzugefügt. Während dieses Prozesses wurden auch einige bereits bestehende Deskriptoren, die eher monologisches Sprechen definieren, aus der Skala *Informationsaustausch in Zusammenhängendes monologisches Sprechen: Informationsvermittlung* verschoben.

### vor A1

„Vor A1“ stellt eine „Wegmarke“ auf halbem Weg zum Niveau A1 dar, ein Kompetenzbereich, auf dem Lernende noch kaum<sup>1</sup> produktive Fähigkeiten erworben haben, sondern sich auf ein Repertoire an Wörtern und Wendungen verlassen müssen. Dass ein solcher Kompetenzbereich vor A1 existiert, wird am Anfang des Abschnitts 3.5 des GeR erwähnt. Eine kurze Liste von Deskriptoren, die unterhalb von A1 kalibriert wurden, wird im Schweizer Nationalen Forschungsprojekt erwähnt. Eine umfangreichere Beschreibung der Kompetenzen von Lernenden auf A1 und der Einbezug eines Niveaus unter A1 war für Benutzer/innen wichtig, wie eine Reihe von Projekten zu Deskriptoren bewies, die auf diese niedrigeren Niveaus fokussierten. Deshalb ist eine Bandbreite von Kompetenzen, die „vor A1“ genannt werden, in der Mehrheit der Skalen mit eingeschlossen.

### Phonologie

Für *Beherrschung der Aussprache und Intonation*, eine bereits im GeR existierende Skala, wurde ein völlig neues System von Deskriptoren entwickelt (siehe den Bericht von Enrica Piccardo: *Phonological Scale Revision Process Report*, Piccardo

1 Anm. der Übersetzer: Im engl. Original heißt es: „... has not yet acquired a generative capacity ...“ Diese Formulierung scheint uns nicht ganz angemessen, denn ohne eine zumindest rudimentäre „generative“ Fähigkeit wäre auch die Verwendung einzelner Wörter und Wendungen u.E. nicht möglich.



2016). In der Forschung, die den ursprünglichen Deskriptoren zugrunde lag, war die Skala für Aussprache und Intonation am wenigsten erfolgreich. Diese Skala war die einzige Beispielskala im GeR, für die muttersprachliche Normen als Bezugspunkt gewählt wurden, obgleich dies eher implizit geschah. Bei einer Aktualisierung schien es angemessener, auf Verständlichkeit als das primäre Konstrukt bei der Beherrschung von Phonologie (inkl. Aussprache und Intonation) zu fokussieren. Dies geschah in Übereinstimmung mit aktueller Forschung, besonders im Kontext von Deskriptoren für *Auf einem plurilingualen/plurikulturellen Repertoire aufbauen*. Das Phonologieprojekt, das daraus resultierte, durchlief alle drei Validierungsphasen, die weiter unten mit Bezug auf andere neue Skalen beschrieben werden, wobei in jeder Phase über 250 Probanden involviert waren.

## Junge Lernende<sup>2</sup>

Es gibt einen weithin anerkannten Bedarf an Instrumenten, die die Anpassung des GeR an das Lehren und Lernen bei jungen Lernenden unterstützen. Es wurde jedoch bewusst entschieden, während dieses Projekts eine parallele Entwicklung und Kalibrierung von neuen Deskriptoren für junge Lernende zu vermeiden, weil Deskriptoren für junge Lernende ohnehin meist aus den Beispieldeskriptoren des GeR abgeleitet und je nach Alter und Kontext adaptiert werden. Darüber hinaus wurde ein großer Teil der Arbeit schon von Fachleuten auf diesem Gebiet aus den Mitgliedstaaten bei der Entwicklung und Validierung des Europäischen Sprachenportfolios für junge Lernende geleistet. Deshalb wurde der Ansatz gewählt, Deskriptoren für junge Lernende zu sammeln und zu sortieren und diese den wichtigen beiden Altersgruppen (7–10 und 11–15) zuzuordnen, die in der Mehrzahl der validierten Beispiele für Sprachenportfolios zu finden sind.

Obgleich nicht voll umfassend, trägt das Projekt eine repräsentative Auswahl von Portfoliodeskriptoren für junge Lernende aus einem Spektrum der Mitgliedsstaaten des Europarats zusammen. Dabei werden im wesentlichen Materialien benutzt, die aus Beispielen der Datenbank für Portfolios gewonnen sind, die durch den Europarat akkreditiert worden sind, und/oder Beispiele, die auf der Website des Europarats registriert sind, zusammen mit Deskriptoren für das Assessment von jungen Lernenden, die von *Cambridge English Language Assessment* zur Verfügung gestellt wurden. Diese Materialien wurden einzeln auf die Niveaus der Beispieldeskriptoren aus dem GeR von 2001 abgestimmt, wobei Entsprechungen zwischen den Deskriptoren für junge Lernende und den Beispieldeskriptoren des GeR identifiziert wurden. Diese wurden dem Sondierungsgremium zur Begutachtung vorgelegt. Diese Sammlung und Angleichung soll die weitere Entwicklung von Curricula, Portfolios und Instrumenten für das Assessment junger Lernender unterstützen, im Bewusstsein, dass lebenslanges Lernen zu den Kompetenzen führt, die im GeR beschrieben werden.

Zusätzlich wurden die erweiterten Beispieldeskriptoren in das Dokument aufgenommen, um Lehrern und Lehrerinnen auf deren Relevanz für Programme für junge Lernende hinzuweisen. Eine Beurteilung der vorgeschlagenen Relevanz für jede der beiden Altersgruppen wurde den erweiterten Beispieldeskriptoren des GeR beigefügt. Diese Beurteilungen wurden auch durch das Sondierungsgremium sowie in einem gesonderten Beratungsworkshop bestätigt.

Die Deskriptoren wurden in zwei Dokumenten präsentiert, einem für jede Altersgruppe. Die Dokumente sind identisch strukturiert und präsentieren die Deskriptoren in Niveaus, beginnend mit „vor A1“, wobei die nicht relevanten GeR-Beispieldeskriptoren herausgefiltert wurden, wenn sie als jenseits der für diese Altersgruppe typischen kognitiven und sozialen Erfahrungen liegend eingeschätzt wurden (vor allem auf den höheren Niveaus). Die Dokumente zeigen somit, welcher Deskriptor im GeR mit dem für junge Lernende verbunden ist, begleitet von einer Angabe, welche Relevanz ein GeR-Deskriptor für die Altersgruppe hat, sofern es noch keine eigenen Beispiele für solche Deskriptoren gibt. Weiterhin speichert ein Archiv, das in Skalen organisiert ist, alle geplanten Deskriptoren für beide Altersgruppen.

2 Eine Zusammenstellung der Deskriptoren für junge Lernende findet sich auf <https://www.coe.int/en/web/common-european-framework-reference-languages/bank-of-supplementary-descriptors>.

Abbildung 1: Entwicklungsdesign des Junge-Lernende-Projekts

<b>Vorbereitende Arbeiten</b>	<b>Anfängliches Zusammentragen von validierten Deskriptoren für Portfolio und Assessment</b> 35 Quellen in der engeren Wahl evaluiert/zusammengestellt aus 19 Mitgliedsstaaten	<b>Kategorisierung der Quellen</b> Abstimmung der Quellen nach Altersgruppen	<b>Expertenbefragung</b> Relevante Altersgruppen finalisiert
<b>Entwicklung</b>	<b>Beurteilung der Übereinstimmung mit Beispieldeskriptoren von 2001</b> Durch federführende Forscher/innen, gemäß den validierten Niveaus für Deskriptoren für junge Lernende	<b>Prüfung der Reihenfolge durch das Steuerungskomitee</b> Durch Mitglieder der erweiterten GeR-Experten- und Autorengruppe	<b>Hinzufügung eines erweiterten Satzes von Deskriptoren mit Relevanzbeurteilungen</b> Die erweiterten Deskriptoren beurteilt in Hinblick auf ihre Relevanz für die jeweilige Altersgruppe
<b>Qualitative Validierung</b>	<b>Beratungen mit Experten</b> Peer-Beurteilung der Abstimmung von Deskriptoren für junge Lernende und Relevanzbeurteilungen	<b>Nachbesserung/Workshop mit Experten</b> Aktualisierung der Abstimmung Relevanzbeurteilungen und Kommentare	
<b>Fertigstellung</b>	<b>Getrennte Referenzdokumente für jede Altersgruppe</b> Einführungskapitel Neuorganisation der Deskriptoren nach Niveaus statt nach Skalen Für beide Altersgruppen Archivversion beibehalten, die nach Skalen organisiert ist	<b>Abschließende Aktualisierung</b> Anpassung an die abschließenden Versionen des Dokuments, um die validierten Veränderungen am erweiterten System der Beispieldeskriptoren im GeR widerzuspiegeln	Eine neue Sortierung von Deskriptoren für junge Lernende

### Mediation<sup>3</sup>

Das im GeR von 2001 sowie im Projekt von 2013–2017 gewählte Konzept von Mediation wird im nächsten Abschnitt erläutert. In der Pilotfassung des GeR von 1996, die während der letzten Phasen des Schweizer Forschungsprojekts veröffentlicht wurde, wurden die Kategorien für Skalen mit Beispieldeskriptoren für Mediation umrissen, um diejenigen für Rezeption, Interaktion und Produktion zu ergänzen. Jedoch wurde kein Projekt begonnen, sie weiterzuentwickeln. Ein wichtiges Ziel der vorliegenden Aktualisierung war deshalb, jetzt solche Beispielskalen für Mediation zur Verfügung zu stellen, da deren Bedeutung im Bildungsbereich ständig zunimmt. Unter Berücksichtigung von Mediation wurden Deskriptoren für *Auf einem plurilingualen/plurikulturellen Repertoire aufbauen* hinzugefügt. Die im Vorwort aufgeführten Institutionen trugen zur Validierung dieser neuen Deskriptoren für Mediation, Online-Interaktion, Reaktion auf Literatur (bzw. kreative Texte) und *Auf einem plurilingualen/plurikulturellen Repertoire aufbauen* bei.

Der Schwerpunkt bei der Entwicklung neuer Skalen lag auf Mediation, für deren verschiedene Aspekte nunmehr 23 Beispielskalen vorliegen (Mediationsaktivitäten: 18; Mediationsstrategien: 5). Der gewählte Ansatz zu Mediation war umfassender als jener im GeR 2001, wo in Abschnitt 2.1.3 Mediation als vierte Kategorie kommunikativer Sprachaktivitäten eingeführt wurde, zusätzlich zu Rezeption, Interaktion und Produktion:

3 Anm. der Übersetzer: Wie schon erwähnt, benutzen wir in der Übersetzung des Begleitbandes den Terminus „Mediation“, weil dieser eher als „Sprachmittlung“ die vielfältigen Aspekte widerspiegelt, von denen hier die Rede ist. Einen guten Überblick findet man am Beginn des Kapitels über Mediation im Begleitband.



Sowohl bei der rezeptiven als auch bei produktiven Sprachverwendung ermöglichen die mündlichen und / oder schriftlichen Aktivitäten der **Sprachmittlung** Kommunikation zwischen Menschen, die aus irgendwelchen Gründen nicht direkt miteinander kommunizieren können. Übersetzung oder Dolmetschen, die Zusammenfassung oder der Bericht ergeben eine (Neu)Fassung eines Ausgangstextes für Dritte, die keinen unmittelbaren Zugriff darauf haben. Sprachmittelnde Aktivitäten, also die Umformung eines schon vorhandenen Textes, nehmen eine wichtige Stellung im alltäglichen sprachlichen Funktionieren unserer Gesellschaften ein.

Diese Beschreibung wird im GeR 2001 in Abschnitt 4.4.4 weiterentwickelt:

Bei **sprachmittelnden Aktivitäten** geht es den Sprachverwendenden nicht darum, ihre eigenen Absichten zum Ausdruck zu bringen, sondern darum, Mittler/innen zwischen Gesprächspartnern/-partnerinnen zu sein, die einander nicht direkt verstehen können, weil sie Sprechende verschiedener Sprachen sind (was der häufigste, aber nicht der einzige Fall ist). Zu den sprachmittelnden Aktivitäten gehören Dolmetschen und Übersetzen sowie das Zusammenfassen und Paraphrasieren von Texten in derselben Sprache, wenn derjenige, für den der Text gedacht ist, den Originaltext nicht versteht.

Das Augenmerk im Text des GeR 2001 liegt somit auf Informationsweitergabe und auf Vermittlungstätigkeiten entweder innerhalb einer Sprache oder über mehrere Sprachen hinweg.

Ähnlich wie in weiten Teilen der Erziehungswissenschaft, steht der konzeptionelle Ansatz in diesem Projekt dem Ansatz von Coste und Cavalli näher, den sie 2015 in einer Veröffentlichung für den Europarat gewählt hatten: *Education, mobility, otherness – The mediation functions of schools* (Coste und Cavalli 2015). Die vollständige Konzeptionalisierung von Mediation wird in *Developing illustrative descriptors of aspects of mediation for the CEFR* (North und Piccardo 2016) beschrieben. Bei der Entwicklung von Kategorien für Mediation übernahm die Autorengruppe Coste und Cavalli Unterscheidung zwischen:

- „Beziehungsrelevante Mediation“: der Prozess der Entwicklung und Gestaltung zwischenmenschlicher Beziehungen, um eine positive, kooperative Situation zu schaffen (für den sechs Skalen entwickelt wurden);
- „Kognitive Mediation“: der Prozess, der Zugang zu Wissen und Konzepten ermöglicht, insbesondere wenn jemand dies nicht direkt und selbstständig bewältigen kann, sei es, weil die Konzepte neu und unvertraut sind und / oder wegen sprachlicher oder kultureller Barrieren.

Allerdings ist es praktisch unmöglich, kognitive Mediation zu leisten, ohne entsprechende beziehungsrelevante Aspekte aufzugreifen. Echte Kommunikation erfordert eine ganzheitliche Integration beider Aspekte. Aus diesem Grund werden die Mediations-Skalen in einer praktischen Unterteilung in vier Gruppen präsentiert:

- Mediation von Texten
- Mediation von Konzepten
- Mediation von Kommunikation
- Mediationsstrategien

Schließlich führten Überlegungen zu sprachübergreifender und kultureller Mediation zu einem speziellen Interesse an der Fähigkeit, plurilinguales oder plurikulturelles Repertoire einzusetzen, für die drei zusätzliche Skalen entwickelt wurden:

- *Auf einem plurikulturellen Repertoire aufbauen*
- *Plurilinguales Verstehen*
- *Auf einem plurilingualen Repertoire aufbauen*

Mit der Entwicklung von Deskriptoren für plurilinguale und plurikulturelle Kompetenz, die mit den GeR-Niveaus verknüpft sind, sollen Lehrende ermutigt werden, in ihren Planungen den Erwerb plurilingualler und plurikultureller Kompetenzen entsprechend dem Leistungsniveau ihrer Lernenden zu berücksichtigen.

## Zur Forschungsmethodik

Das Projekt setzte die Methoden ein, die bei der Erforschung der ursprünglichen GeR-Deskriptoren von Brian North und Günther Schneider in der Schweiz benutzt wurden, und entwickelte sie weiter. Das Projekt nutzte einen ähnlichen multimethodischen Ansatz (*mixed method design*) unter Einbezug qualitativer und quantitativer Verfahren, wie in Abbildung 2 zusammengefasst. Auf eine ausführliche Analyse der relevanten Fachliteratur folgte eine Phase, in der die Deskriptoren zunächst intuitiv formuliert und dann von einem Sondierungsgremium begutachtet wurden. Von Februar 2015 bis Februar 2016 gab es drei Phasen der Validierung, an denen 1.000 Probanden teilnahmen. Auf die Validierung folgten von Juli 2016 bis Februar 2017 drei Runden der Beratung und eine Pilotierung von Januar bis Juli 2017.

Die Forschungsmethode für die Entwicklung und Validierung der neuen Skalen spiegelt, wenn auch in größerem Umfang, jene des ursprünglichen Schweizerischen Forschungsprojekts wider (GeR 2001 Anhang B). Wie schon das ursprüngliche Forschungsprojekt, hatte auch dieses Projekt drei breit angelegte Phasen:

- Erste Forschung und Entwicklungen (intuitive Phase)
- Überprüfen und Verbesserung der Kategorien und der Qualität der Deskriptoren (qualitative Phase)
- Kalibrierung der besten Deskriptoren auf einer mathematische Skala und Bestätigung der Grenzwerte zwischen den Stufen (quantitative Phase)

Diese Aufgaben fanden zwischen Januar 2014 und März 2016 statt und wurden anschließend durch Konsultationen und Pilotierungen ergänzt.

## Vorbereitende Arbeiten

Als erster Schritt wurden vorhandene Instrumente und Artikel zum Thema Mediation gesammelt; ebenfalls wurden die Mediations-Deskriptoren von *Profile Deutsch* und einigen anderen Quellen ins Englische übersetzt. In mehreren Treffen mit Daniel Coste und Marisa Cavalli, den Autoren von *Education, mobility, otherness – The mediation functions of schools* (Coste und Cavalli 2015) wurde ein erster Satz von Kategorien entwickelt sowie eine erste Sammlung von Deskriptoren für Mediation von Texten und Mediation von Konzepten zusammengestellt und dann in einer Entwurfsfassung entworfen. Die Haupt-Kategorien für die Gruppierung der Skalen in diesem frühen Stadium waren:

- Kognitive Mediation (Zugang zu Wissen, Bewusstheit und Fertigkeiten erleichtern)
- Interpersonale Mediation<sup>4</sup> (Entwickeln und Pflege positiver Beziehungen; Rollen und Konventionen definieren, um Offenheit<sup>5</sup> zu fördern, Konflikte zu vermeiden bzw. aufzulösen und Kompromisse auszuhandeln)
- Mediation von Texten (Informationen und Argumente weitergeben: klären, zusammenfassen, übersetzen usw.)

Die vollständige erste Sammlung enthielt auch eine Anzahl von Skalen-Entwürfen mit Bezug zu institutioneller Mediation (zum Beispiel: Neuankömmlinge integrieren, in Institutionen mit Betroffenen umgehen, Entwicklung und Pflege institutioneller Beziehungen), zusammen mit einer Anzahl von Skalen zu verschiedenen Aspekten von Mediation durch Lehrende – wobei beide Aspekte den Fokus von Coste und Cavalli auf die Rolle von Mediation in Schulen widerspiegeln. Allerdings bestand beim ersten Beratungstreffen im Juli 2014 Übereinstimmung darüber, dass diese Skalen Aspekte von Interaktion und Produktion wiederholten, die bereits im GeR enthalten waren und daher keine neuen Perspektiven eröffneten. Aus diesem Grund wurde die Entwicklung mit Fokus auf die oben erwähnten Kategorien konzeptioneller, interpersoneller und textlicher Mediation vorangetrieben. Die Sammlung wurde für ein Expertentreffen überarbeitet, das im September 2014 eine Autorengruppe installierte.

4 Anm. der Übersetzer: Wir benutzen hier den für Kommunikation oft benutzten Terminus „interpersonal“. Die Gesprächspartner/innen müssen die diversen kulturspezifischen Zeichen verstehen, wie zum Beispiel Grußformeln, Stimmlage, Mimik und Gestik, Distanz und Nähe, also nicht nur verbale, sondern auch paraverbale, non-verbale sowie extraverbale, die jeweils kulturspezifisch sein können.

5 Anm. der Übersetzer: Im engl. Text wird hier das Wort *receptivity* benutzt, das auch „Aufnahmefähigkeit“ bedeuten kann.



## Entwicklung

Die Autorengruppe führte anschließend eine gründliche Literaturrecherche durch und überarbeitete die ursprüngliche Sammlung in mehreren Arbeitstreffen zwischen September 2014 und Februar 2015. Untergruppen befassten sich mit Online-Interaktion, plurilingualer/plurikultureller Kompetenz und Phonologie. Die Arbeit zu plurilingualen und plurikulturellen Kompetenzen ergab sich quasi natürlich aus den Betrachtungen zu sprachübergreifender Mediation, insbesondere in der Rolle von Vermittlern.

Arbeiten zur Phonologie wurden begonnen, weil die existierenden GeR-2001-Skalen zur Aussprache und Intonation als einzige unter den GeR-Beispielskalen implizite Muttersprachler/innen als Bezugspunkt wählten und dadurch unrealistische Erwartungen weckten (B2: „Hat eine klare, natürliche Aussprache und Intonation erworben“). Dies wurde als unvereinbar mit der plurilingualen Perspektive angesehen. Ein Sondierungsgremium unterstützte die Arbeit der Autorengruppe mit zahlreichen Input-Materialien und Feedback.

Im Februar 2015 stand ein Satz von 427 Deskriptoren-Entwürfen zu Online-Interaktion, Mediationsaktivitäten und Mediationsstrategien und für plurilinguale/plurikulturelle Kompetenz für die erste Validierungsrunde bereit. Da die Arbeit zu plurilingualen/plurikulturellen Kompetenzen und Phonologie später gestartet worden war, wurden zu diesem Zeitpunkt nur einige Deskriptoren für erstere und keine Deskriptoren für letztere einbezogen. Die Phonologie-Deskriptoren wurden erstmals in einem Workshop und in Konsultationen mit Phonologie-Experten im Juni 2015 erprobt.

## Qualitative Validierung

Zu diesem Zeitpunkt waren 137 Institutionen für die Teilnahme an der Validierung gewonnen worden. Diese erste Phase fand in diesen Institutionen in Präsenz-Workshops statt, an denen fast 1.000 Personen teilnahmen. Die Aufgabe war eine etwas systematischere Variante als diejenige, die in den 32 Workshops des ursprünglichen GeR-Deskriptoren-Forschungsprojekts verwendet worden war. Die Teilnehmenden diskutierten paarweise rund 60 Deskriptoren für drei bis fünf verwandte Bereiche, entschieden, welchen Bereich die Beschreibungen betrafen, bewerteten sie nach (a) Klarheit, (b) pädagogischer Relevanz und (c) ihrem Bezug zum praktischen Sprachgebrauch und schlugen verbesserte Formulierungen vor. In der Folge wurden 60 Deskriptoren fallen gelassen, darunter eine komplette Skala. Sehr viele der übrigen Deskriptoren wurden neu formuliert, meistens gekürzt, und auf Anraten der Workshop-Teilnehmenden wurden zwei neue Skalen entworfen (*Einen geschriebenen Text mündlich übersetzen*; *Komplizierte Informationen in kleinere Einheiten aufgliedern*). An dieser Stelle wurden einige der Details, die aus den Deskriptoren gelöscht worden waren, zu den Beispielen für verschiedene Domains genommen (s. Anhang 5). Qualitative Validierung für Phonologie, an der online 250 Projekt-Teilnehmende mit den gleichen (vertrauten) Aktivitäten beteiligt waren, kam viel später im Jahr, im November und Dezember 2015.

## Das Rasch-Modell

Das Rasch-Modell wurde nach dem dänischen Mathematiker Georg Rasch benannt. Es ist das meistbenutzte einer Gruppe von Probabilitätsmodellen, die Latent-Trait-Theorien (auch IRT = Item-Response-Theorien genannt) operationalisieren. Das Modell analysiert das Ausmaß, in dem ein Item dem ihm zugrunde liegenden Konstrukt (= *latent trait*) entspricht, das gemessen werden soll. Es schätzt erstens die Schwierigkeitswerte (= wie schwierig jedes Item ist) und zweitens die Fähigkeitswerte (= wie kompetent eine Person in Bezug auf das jeweilige Merkmal ist) ein. Das Modell wird für viele Zwecke benutzt, aber die beiden wichtigsten sind<sup>6</sup>:

- der Aufbau von Itembanken für Tests
- die Analyse von (ausgefüllten) Fragebogen<sup>7</sup>

6 Anm. der Übersetzer: Der wichtigste Zweck des Rasch-Modells ist die stichprobenunabhängige Analyse von Testitems, und die ist nützlich z. B. für den Aufbau von Itembanken. Itembanken können wiederum polytome „Fragebogen“-Items enthalten, je nachdem, was es für eine Itembank ist, ob man z. B. 0, 1, 2, 3 oder mehr Punkte erhalten kann. Es kann auch zur Analyse von Antworten auf Fragen wie diese benutzt werden: „Wie gut finden Sie X auf einer Skala von 0 bis 5?“

8 7 Anm. der Übersetzer: Der engl. Ausdruck *questionnaire* bedeutet hier eher „ausgefüllter Fragebogen“ oder „Antwortbogen“.

Für die Analyse von (ausgefüllten) Fragebogen wird eine Variante benutzt, das *rating scale model* (RSM, „Rating-skalenmodell“). Eine solche „Multifacetten“-Variante<sup>8</sup> des Rasch-Modells kann die Subjektivität aus der Beurteilung von Prüfenden eliminieren. Detaillierte Erklärungen findet man im *Reference Supplement* des *Manual for Relating Examinations to the CEFR* (Handbuch zum Bezug von Sprachprüfungen auf den GeR).

Der Hauptvorteil des Rasch-Modells liegt – im Gegensatz zur klassischen Testtheorie – darin, dass die gewonnenen Werte für andere Gruppen generalisiert werden können, die man als Teil der allgemeinen Gesamtpopulation betrachten kann (d.h. die eine hinreichende Menge der gleichen Merkmale teilen).

Die objektive Skalierung und die potentielle Übertragbarkeit der Skalenwerte macht das Modell in besonderem Maße geeignet für die Bestimmung des Niveaus, auf dem man Kann-Beschreibungen auf gemeinsamen Referenzskalen wie denen der GeR-Niveaus verorten kann.

## Quantitative Validierung

An der nächsten Phase nahmen 189 Institutionen mit insgesamt 1.294 Teilnehmern/Teilnehmerinnen aus 45 Ländern teil. Wiederum organisierten alle der teilnehmenden Institutionen einen Präsenz-Workshop. Nach Einarbeitungsaktivitäten, ähnlich denen, die im Handbuch des Europarts *Relating Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) – A Manual* (Council of Europe 2009) empfohlen werden, nahmen die Teilnehmer/innen an einem Workshop zur Standardsetzung teil, in dem sie individuell und nach Diskussion die Deskriptoren im Entwurf den GeR-Niveaus zuwiesen. Zu diesem Zweck wurde das gesamte Spektrum der Bandbreiten aus den ursprünglichen Forschungen zu Deskriptoren des GeR benutzt (= 10 Niveaustufen von „vor A1“ bis C2). Die Teilnehmer/innen markierten ihre Entscheidungen auf PDF-Ausdrucken, und erst am Ende trugen sie die individuellen Entscheidungen, die sie in Betracht gezogen hatten, endgültig in einen Online-Fragebogen ein. Bei der Analyse wurde zunächst der Prozentsatz der Antwortenden für jeden Deskriptor, den sie einem Niveau oder Unterniveau zugewiesenen hatten, berechnet, und danach wurde eine Skalierungsanalyse nach dem Rasch-Modell durchgeführt, so wie bei den ursprünglichen Forschungen zu den Deskriptoren im GeR. Für eine Rasch-Analyse braucht man eine Matrix von untereinander verbundenen Daten, und jedes einzelne Item (hier Deskriptoren) sollte idealerweise 100 Antworten aufweisen. Die folgende Matrix wurde benutzt, um überlappende Datensätze zu schaffen, die an verschiedene teilnehmende Institutionen verteilt wurden. Auf diese Art und Weise erreichte man für alle Skalen mit Deskriptoren das Ziel: Die niedrigste Zahl von Antworten für jede beliebige Skala war 151, und die höchste 273.

8 Anm. der Übersetzer: Der Begriff Multifacetten-Rasch-Modell wird z. B. von Thomas Eckes verwendet in *Zeitschrift für Sozialpsychologie* 2006/37, pp. 185-195 (<https://doi.org/10.1024/0044-3514.37.3.185>)

Modell der Datensammlung für jede der drei Validierungsphasen (Diese genaue Verteilung ist Phase 2: Zuordnung zu Niveaus)

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Gemeinsame Konstruktion von Bedeutung	3	X	X		X																			
Konzeptuelle Gespräche entwickeln <sup>9</sup>	3	X		X		X																		
Spezifische Informationen mündlich weitergeben	3							X	X														X	
Daten mündlich erklären (Grafiken, Schaubilder usw.)	3			X				X	X															
Mündliche Verarbeitung von Texten	3							X						X										
Dolmetschen/Als Sprachmittler handeln	3													X									X	X
Einen geschriebenen Text mündlich übersetzen	4							X						X									X	X
Kooperative Interaktion unterstützen	3		X																			X		
Interaktion organisieren	4			X		X	X																	
Kommunikation in heiklen Situationen und bei Meinungsverschiedenheiten auflösen	3					X	X													X				
Spezifische Informationen schriftlich weitergeben	3									X														
Daten schriftlich erklären (Grafiken, Schaubilder usw.)	3			X									X											
Schriftliche Verarbeitung von Texten	3									X			X											
Einen geschriebenen Text schriftlich übersetzen	3										X			X										
Eine positive Atmosphäre schaffen	3																							
Etwas mit Vorwissen verbinden	3		X																					
Einen (dichten) Text weiter ausführen	4										X	X					X							X

9 Einige der hier aufgeführten Skalen heißen im endgültigen Text anders; die Abweichungen sind kursiv geschrieben.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
Einen Text glätten	3	9									X			X	X									
Komplizierte Informationen in kleinere Einheiten aufgliedern	5	6	X					X	X	X					X									
Informationen visuell darstellen	4	6	X					X	X	X					X									
Sprache anpassen	3	11									X	X		X										
Online-Konversation und -Diskussionen	3	39					X								X				X					
Zielorientierte Online-Transaktionen und Zusammenarbeit	3	23						X	X						X			X						
Persönliche Reaktion auf kreative Texte (einschließlich Literatur)	3	19															X		X					
Analyse und Kritik kreativer Texte (einschließlich Literatur)	3	17									X						X	X						
Auf einem plurikulturellen Repertoire aufbauen und interkulturelles Wissen aufbauen	3	18				X											X	X						
Auf einem plurilingualen Repertoire aufbauen	3	20																X				X		X
GeR Zusammenarbeit in Diskussionen	5	X	X	X	X														X					
GeR Soziolinguistische Kompetenz	4					X											X	X				X		X
GeR Verarbeitung	6		X							X											X			X
GeR Großer Anker	10	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<b>TOTAL</b>		<b>62</b>	<b>65</b>	<b>66</b>	<b>69</b>	<b>67</b>	<b>66</b>	<b>67</b>	<b>66</b>	<b>67</b>	<b>66</b>	<b>70</b>	<b>70</b>	<b>66</b>	<b>67</b>	<b>72</b>	<b>68</b>	<b>69</b>	<b>62</b>	<b>68</b>	<b>68</b>	<b>71</b>	<b>69</b>	
Andere in den Kategorien eingeschlossene GeR-Anker						7	1					6	6	6								1		
Alle GeR-Anker	15	15	16	15	15	14	17	11	10	16	16	10	16	16	16	10	14	14	15	10	16	15	20	



Eine solche Matrix wurde in jeder Validierungsphase benutzt, mit dem bewussten Ziel, die Verteilung der Kategorien von Deskriptoren auf solche Nutzergruppen zu erreichen, von denen man wusste, dass sie sich für die betreffenden Kategorien interessieren. Die Vorteile der Rasch-Analyse waren zunächst, jene Deskriptoren, die sich als unbrauchbar erwiesen hatten, und jene Teilnehmer/innen, die die Aufgaben nicht lösen konnten, zu identifizieren und auszuschließen, und weiterhin, dass es für jeden Deskriptor einen arithmetischen Wert gab. Dieser Wert konnte dann umgewandelt werden für die Skala, die den ursprünglichen GeR-Deskriptoren zugrunde lag, indem einige der ursprünglichen GeR-Deskriptoren als „Anker-Items“ benutzt wurden.

Ergebnisse dieser vorläufigen quantitativen Analyse wurden bei einem Beratungstreffen im Juli 2015 diskutiert. 36 Deskriptoren wurden fallen gelassen, und ungefähr die Hälfte wurde einer erneuten Kalibrierung unterzogen, häufig nach Änderungen. Ein größeres Problem war das Fehlen von Deskriptoren für Mediation und plurilinguale/plurikulturelle Kompetenz auf A1 und A2. Man bemühte sich, diese vor der nächsten Phase zu verfassen.

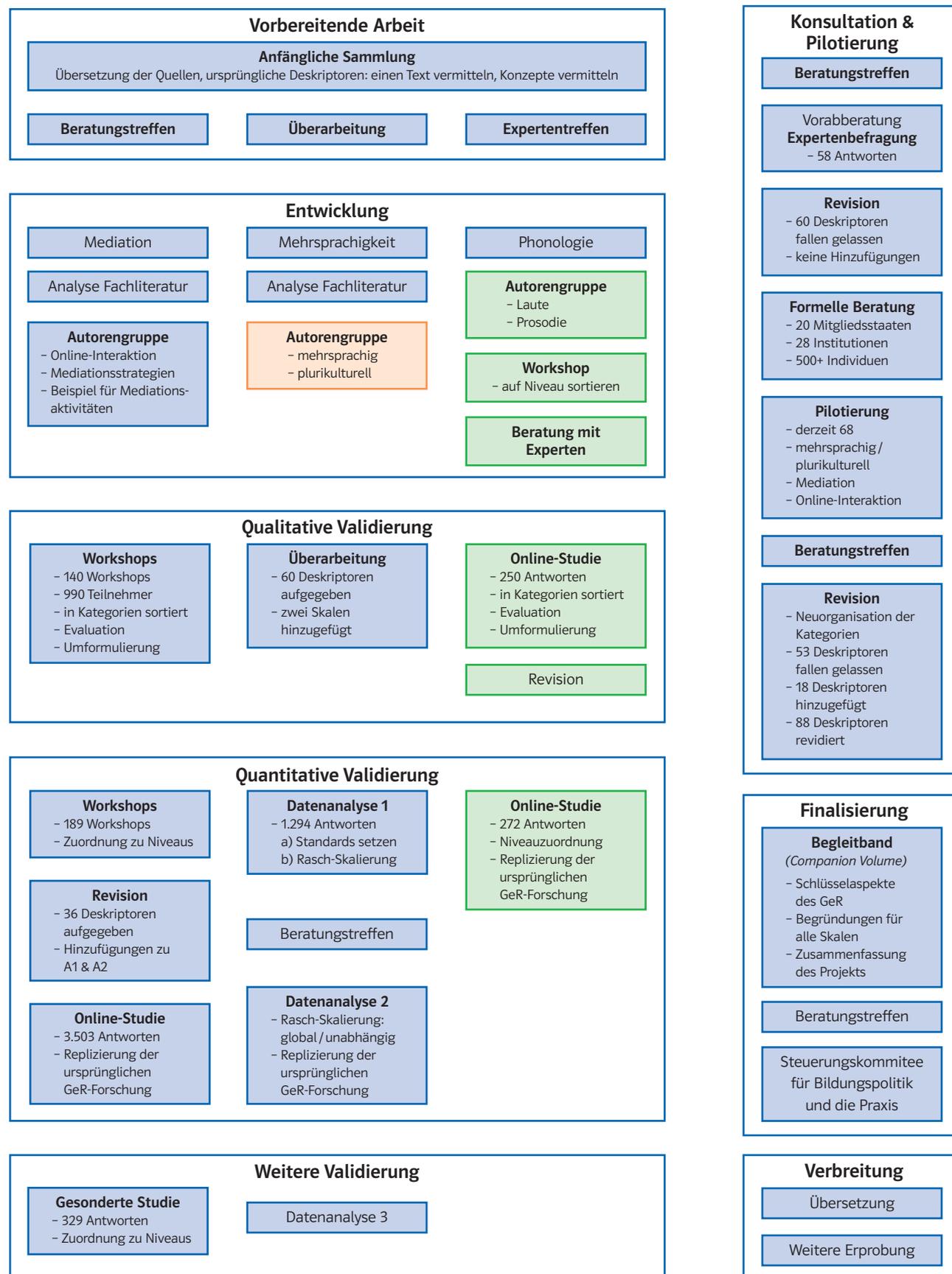
Die zentrale qualitative Datensammlung erfolgte dann in einer Online-Befragung auf Englisch und Französisch von Oktober bis Dezember 2015. Dieses Mal antworteten die Teilnehmer/innen an der Erhebung individuell auf die Frage: „Könnten Sie oder eine Person, an die Sie denken, tun, was in dem Deskriptor beschrieben wird?“ Sie wurden aufgefordert, dieses dreimal zu tun, für ihre verschiedenen plurilingualen angenommenen oder realen Personen und/oder für Personen, die sie sehr gut kennen (Partner, Kinder usw.), und dieses resultierte in 3.503 brauchbaren Antworten von ungefähr 1.500 Personen. Die Aufgabe war eine etwas adaptierte Replikation derjenigen, die bei der Kalibrierung der ursprünglichen GeR-Deskriptoren benutzt wurde, welche auf der Beurteilung von Lehrkräften mit Deskriptoren für ein repräsentatives Sample von Studierenden in ihren Klassen basierte. Zwei Analysen wurden durchgeführt: eine globale Analyse mit allen Deskriptoren und eine zweite Analyse, bei der jede Hauptkategorie getrennt analysiert wurde. Entscheidungen über das Niveau eines jeden Deskriptors wurden dann auf der Basis aller verfügbaren Informationen getroffen.

Eine quantitative Validierung für Phonologie folgte im Januar 2016, wobei 272 Personen teilnahmen. Es gab zwei Aufgaben: (a) Zuordnung zu Niveaus und (b) Lernerleistungen in Videoclips zu beurteilen („Können die Lernenden in dem Video tun, was im Deskriptor beschrieben wird?“). Verschiedene Techniken der Standardsetzung wurden benutzt; für Details verweisen wir Leser/innen erneut auf *Phonological Scale Revision Process Report* (Piccardo 2016).

## Weitere Validierung von plurilingualer und plurikultureller Kompetenz

Schließlich wurde aus drei Gründen im Februar 2016 eine zusätzliche Erhebung durchgeführt. Erstens war es eine Gelegenheit, Deskriptoren für rezeptive Strategien und plurilinguales Verstehen mit einzuschließen, die vor allem aus dem MIRIADI-Projekt<sup>10</sup> adaptiert wurden; zweitens hatte die Aufgabe für Plurilingualismus bei der zentralen Online-Umfrage nicht besonders gut funktioniert, sodass diese zusätzliche Erhebung eine Wiederholung mit unterschiedlichen Aufgaben darstellte; und schließlich war es eine Gelegenheit, weitere Deskriptoren für plurikulturelle Kompetenz hinzuzufügen, vor allem auf den unteren Niveaus. Die Erhebung wurde in zwei völlig unterschiedlichen Parallelversionen durchgeführt. 267 Freiwillige aus dem Teilnehmerkreis des Projekts füllten den einen Fragebogen aus, während 62 Experten für plurilinguale Bildung den anderen ausfüllten. Die Ergebnisse wurden dann einander gegenübergestellt. Sie erwiesen sich für beide Gruppen als identisch, und die Kalibrierung auf Niveaus war ebenfalls in hohem Maße kompatibel mit der Skala für Soziolinguistische Angemessenheit im GeR.

Abbildung 2: Multimethodisches Forschungsdesign





## Probleme und Lösungen

Eine große Zahl von Rückmeldungen kam in Beratungstreffen und während der erweiterten Beratung und Pilotierung in den Jahren 2016–17 von Personen, die schon an den Validierungsaktivitäten im Jahr 2015 teilgenommen hatten. Dieser Abschnitt fokussiert auf einige der Schlüsselaspekte, die im Verlauf des Projekts angesprochen wurden, und darauf, wie jeder von ihnen angegangen wurde.

## Beziehungen zwischen den Skalen für Mediation und denen im GeR 2001

Ogleich der Schwerpunkt des Projekts darin lag, solche Deskriptoren für Aktivitäten und Strategien zur Verfügung zu stellen, die nicht schon im GeR 2001 abgedeckt wurden, erinnerten einige Aspekte in den Skalen für Mediation, besonders auf den unteren Niveaus, an die Arten von Aktivitäten, die in den bereits vorhandenen Skalen des GeR beschrieben werden. Das liegt daran, dass einige Aspekte von Mediation (in der erweiterten Interpretation, die jetzt verwendet wird) bereits in den Beispielskalen von 2001 vorhanden sind. Die neuen Skalen unter Mediation von Texten, z. B. *Spezifische Informationen weitergeben*, *Daten erklären* und *Verarbeitung von Texten*, sind eine Ausarbeitung von Konzepten, die in der bereits existierenden Skala *Texte verarbeiten* in Abschnitt 4.6.3 des GeR unter „Text“ stehen. In ähnlicher Weise sind die Skalen, die besonders Gruppeninteraktion betreffen, nämlich *Interaktion und Zusammenarbeit in einer Gruppe erleichtern*, *Gemeinsame Konstruktion von Bedeutung* und *Gespräche über Konzepte und Ideen fördern*, in vielerlei Hinsicht eine Weiterentwicklung von Konzepten in der bereits existierenden Skala *Kooperieren* unter Interaktionsstrategien. Dies unterstreicht die Schwierigkeiten aller Kategorisierungssysteme. Man sollte niemals die Tatsache unterschätzen, dass Kategorien nur zweckmäßige, erfundene Artefakte sind, die es uns erleichtern, die Welt zu interpretieren. Grenzen sind unscharf und Überschneidungen sind unvermeidlich.

## Sprachenübergreifende Mediation

Früherer Versionen der Deskriptoren hatten mit verschiedenen Formulierungen experimentiert, die diesen Aspekt in Betracht zu ziehen versuchten. Es erwies sich jedoch als außerordentlich schwierig, eine klare Unterscheidung zu treffen. „Muttersprache“ und „Erstsprache“ und „Schulsprache“ sind oft nicht synonym, und sogar Begriffe wie „Ausgangssprache“ und „Zielsprache“ erwiesen sich als verwirrend (wenn man z. B. aus einer anderen Sprache mittelt, könnte das in die Muttersprache erfolgen, die andere Sprache ist in solchen Fällen die Ausgangssprache und die Muttersprache wäre die Zielsprache). Versuche, auf diese Varianten einzugehen, bedeuteten auch, dass sich irgendwann die Sammlung der Deskriptoren unnötigerweise verdreifachte, bei nur geringfügigen Veränderungen in der Formulierung.

Deshalb entschied sich die Projektgruppe dafür, so wie bei den ursprünglichen Beispieldeskriptoren von 2001, die Kalibrierung darauf zu beziehen, welche Schwierigkeiten in der funktionalen Sprachkompetenz zu beobachten sind, unabhängig davon, welche Sprachen involviert sind. Es wird empfohlen, dass diese Sprachen von den Benutzern und Benutzerinnen spezifiziert werden, wenn Deskriptoren für den praktischen Gebrauch adaptiert werden.

Die Skalen für Mediation von Texten enthalten einen Verweis auf „Sprache A“ und „Sprache B“, was erweiterte Bezeichnungen für die Mediation von kommunikativen Quellen bzw. kommunikativem Output sind. In Anmerkungen wird dazu festgestellt, dass die Mittlung innerhalb einer Sprache oder über Sprachen, Varietäten oder Register hinweg erfolgen kann (oder jede denkbare Kombination davon), und Benutzer/innen die betreffenden spezifischen Sprachen benennen könnten. Benutzer/innen könnten auch gleichermaßen Beispiele für ihren jeweiligen Kontext zur Verfügung stellen, vielleicht angeregt durch die Beispiele in Anhang 5 für die vier Domänen von Sprachgebrauch: öffentlich, persönlich, beruflich und den Bildungsbereich.

So könnte z.B. der erste Deskriptor in der Skala *Spezifische Informationen mündlich oder in Gebärdensprache weitergeben*:

*Kann (in Sprache B) die Relevanz einer bestimmten Information erklären, die in einem bestimmten Abschnitt eines langen und komplexen Textes (aus Sprache A) vorkommt.*

zu diesem werden:

*Kann auf Französisch die Relevanz einer bestimmten Information erklären, die in einem bestimmten Abschnitt eines langen und komplexen Textes auf Englisch vorkommt (z.B. in einem Artikel, auf einer Website, in einem Buch oder in einem direkten Gespräch/online, das Zeitgeschehen betreffend oder einen Bereich des persönlichen Interesses oder ein persönliches Anliegen).*

oder, sofern es die Kommunikation in nur einer Zielsprache betrifft:

*Kann die Relevanz einer bestimmten Information erklären, die in einem bestimmten Abschnitt eines langen und komplexen Textes vorkommt (z.B. in einem Artikel, auf einer Website, in einem Buch oder in einem direkten Gespräch/online, das Zeitgeschehen betreffend oder einen Bereich des persönlichen Interesses oder ein persönliches Anliegen).*

Alle Deskriptoren für Mediation eines Textes umfassen integrierte Fertigkeiten, eine Mischung aus Rezeption und Produktion. Der Fokus liegt *nicht* auf Rezeption, für die es bereits Skalen im GeR gibt. Das Niveau, auf das Deskriptoren kalibriert sind, spiegelt das Niveau der Verarbeitung und die erforderliche Produktion. Wenn Rezeption und Produktion in verschiedenen Sprachen erfolgen, ist das Niveau des Deskriptors dasjenige, das für die Verarbeitung und die Artikulation der Ausgangsbotschaft in der oder den Zielsprache(n) erforderlich ist.

## Allgemeine und kommunikative Sprachkompetenzen

In allen Beispielskalen im GeR definieren die Deskriptoren auf einem bestimmten Niveau das, was Sprachverwendende/Lernende üblicherweise erreichen können, wenn sie über eine kommunikative Sprachkompetenz (GeR 2001, Abschnitt 5.2) in der oder den betreffenden Sprache(n) gemäß den im GeR angegebenen Niveaus verfügen, sofern die betreffende Person auch über die persönlichen Eigenschaften, die Kenntnisse, die kognitive Reife und Erfahrung – d.h. über die Allgemeinen Kompetenzen (GeR 2001, Abschnitt 5.1) – verfügt, die dafür notwendig sind. Die Skalen des GeR sollen die Fähigkeiten als Profil darstellen können. Es ist unwahrscheinlich, dass alle Sprachverwendenden, die allgemein auf „B1“ sind, exakt alles tun können, was auf B1 in allen Beispielskalen des GeR beschrieben wird, nicht mehr und nicht weniger. Es ist wesentlich wahrscheinlicher, dass Menschen, deren allgemeines Niveau „B1“ ist, sich in Bezug auf einige Aktivitäten tatsächlich auf A2 oder A2+ befinden, und in Bezug auf andere auf B1+ oder sogar auf B2, je nach ihrem persönlichen Profil und ihrer allgemeinen Kompetenz, was wiederum von ihrem Alter, ihrer Erfahrung usw. abhängt. Dies ist der Fall bei vielen Beispielskalen des GeR, die kognitive Fähigkeiten betreffen, wie zum Beispiel *Notizen anfertigen, Information und Argumentation verstehen, Formelle Diskussion und Besprechungen, Die Ansprache von Publikum sowie Berichte und Aufsätze schreiben*. Dies ist auch der Fall bei vielen Mediationsaktivitäten. Einige der Skalen zu Mediation von Texten (z.B. *Verarbeitung von Texten*) oder Mediationsstrategien (z.B. *Einen Text straffen*) umfassen Aktivitäten, die einen hohen Grad an kognitiver Gewandtheit erfordern, die man nicht in gleichem Maße bei allen Menschen erwarten kann.

Ganz ähnlich werden die Profile von Sprachverwendenden/Lernenden auf beispielsweise B1 sich erheblich unterscheiden bezüglich *Auf einem plurilingualen/plurikulturellen Repertoire aufbauen*, je nach ihrem persönlichen Entwicklungsstand sowie den Erfahrungen und Kompetenzen, die sie bereits erworben haben. Statt zu versuchen, den Einfluss von individuellen Unterschieden zu eliminieren, erkennt dieser Ansatz zur Formulierung der Deskriptoren an, dass sie ein Schlüsselfaktor sind, der zur Entwicklung eines einzigartigen Profils von kommunikativen Fähigkeiten von Lernenden beiträgt.



## Allgemeine und kommunikative Sprachkompetenzen bei *Auf einem plurikulturellen Repertoire aufbauen*

Wie bei Mediation umfasst der Gebrauch des plurikulturellen Repertoires ein Spektrum an allgemeinen Kompetenzen (GeR 2001, Abschnitt 5.1), üblicherweise in enger Verbindung mit pragmatischen und soziolinguistischen Kompetenzen (GeR 2001, Abschnitte 5.2.2 und 5.2.3). Deshalb kommen in dieser Skala, ebenso wie in den Skalen für Mediation und vielen anderen Skalen im GeR, auch andere als nur Sprachkompetenzen ins Spiel. Die Grenzen zwischen Weltwissen, soziokulturellem Wissen und interkulturellem Bewusstsein sind, wie im GeR 2001 erklärt, nicht wirklich eindeutig. Das sind auch weder die Grenzen zwischen praktischen Fertigkeiten und prozeduralem Wissen – welche auch soziale Fertigkeiten mit einschließen – noch die zwischen soziokulturellem Wissen und interkulturellen Fertigkeiten und Wissen. Auf dem Gebiet der Soziopragmatik untersucht man diese Aspekte auch aus einer eher „linguistischen“ Perspektive. Was wichtiger ist als mögliche Überschneidungen zwischen Kategorien, ist die Tatsache, dass Sprachverwendende/Lernende alle diese verschiedenen Aspekte bei der Schaffung von Bedeutung in einer kommunikativen Situation mit ins Spiel bringen, verbunden mit einer angemessenen kommunikativen Sprachkompetenz. Einige Menschen sind eher dazu in der Lage als andere, sofern ein bestimmtes Niveau der Sprachkompetenz ihnen das erlaubt, vielleicht auch aufgrund ihrer verschiedenen Begabungen und Erfahrungen.

### Konsultation und Pilotierung

Der oben beschriebenen Entwicklung und Validierung folgte ein Prozess der Konsultation und der Pilotierung in drei Phasen:

- Workshop mit Experten
- der Konsultation vorausgehende Online-Erhebung mit Experten
- formelle Konsultation

Nach einem Treffen mit Experten des Europarats im Juni 2016 und einer detaillierten, der Konsultation vorausgehenden Online-Erhebung mit GeR-Experten im Sommer 2016, wurden die Deskriptoren revidiert, bevor zwischen Oktober 2016 und Februar 2017 eine formelle Konsultation auf Englisch und Französisch stattfand. Es gab zwei parallele Erhebungen für Einzelpersonen und Institutionen. Ungefähr 500 einzelne Auskunftspersonen füllten den Fragebogen aus, zusammen mit einer Anzahl eingeladener Institutionen und Curriculum- oder Assessment-Einrichtungen. Zusammen mit anderen Fragen wurden die Teilnehmer/innen an der Erhebung gebeten zu sagen, in welchem Umfang sie die beiden neuen Skalen als hilfreich empfanden, und zu den Deskriptoren Stellung zu nehmen. Alle vorgeschlagenen neuen Skalen wurden von 80 % der Teilnehmer/innen als hilfreich oder sehr hilfreich eingeschätzt, wobei die Institutionen/Einrichtungen zu positiveren Antworten tendierten. Die beliebtesten der neuen Skalen betrafen Mediation von Texten, Zusammenarbeit in kleinen Gruppen und Online-Interaktion. Bei zwei Beispielskalen gab es beträchtliche Unterschiede bei der Meinung von Einzelpersonen und Institutionen: *Zielorientierte Online-Transaktionen und -Kooperation* sowie *Auf einem plurilingualen Repertoire aufbauen*. Während 96 % der Institutionen diese beiden Skalen als hilfreich oder sehr hilfreich befand, waren es nur 81 bis 82 % der Einzelpersonen.

Bei der formellen Konsultation begrüßten zwei Drittel der Teilnehmer/innen an der Erhebung definitiv die Tatsache, dass die Beispielskalen für Mediation über den Bereich des klassischen modernen Sprachunterrichts hinausgingen (hin zu CLIL und Sprache der Bildungseinrichtung), mit über 90 % weitgehender Zustimmung bei sowohl Einzelpersonen als auch Institutionen. Eine Vielzahl von Kommentaren und Vorschlägen gingen ein, die alle zu den endgültigen Formulierungen der Deskriptoren und der Skalentitel sowie der Art beitrugen, in der die Skalen präsentiert werden.

Die Pilotierung erfolgte zwischen Februar und Juni 2017, und die Ergebnisse wurden ständig in die Formulierung und die Präsentation der Beispielskalen eingefügt. Die große Mehrheit der Pilotierungen wählte Deskriptoren aus den relevanten Skalen aus, um darauf das Design der kommunikativen Aufgaben im Unterricht zu basieren, und benutzten danach die Deskriptoren, um den Sprachgebrauch der Lernenden zu beobachten. Das Feedback zu den Deskriptoren war sehr positiv, mit einigen nützlichen Vorschlägen für kleine Überarbeitungen. Die beliebtesten Bereiche bei der

Pilotierung waren Zusammenarbeit in Gruppen, Mediation von Texten und plurilinguale/plurikulturelle Kompetenz. Bei einer Pilotierung wurden auch die beiden Beispieskalen für Online-Interaktion in einer getrennten Erhebung mit 1.175 italienischen Lehrkräften für Englisch benutzt, die einen Online-Kurs über die Benutzung digitaler Ressourcen abgeschlossen hatten<sup>11</sup>. 94,8 % der Antwortenden fanden die Deskriptoren sehr klar oder ziemlich klar, und 80,8 % berichteten, dass sie sehr leicht oder ziemlich leicht für die Selbsteinschätzung zu benutzen waren.

Zur gleichen Zeit wie die formelle Konsultation wurde ein Fragebogen an Mitgliedsstaaten versandt, in dem nach der Benutzung des GeR in den jeweiligen Ländern, der Vertrautheit mit unterstützenden Materialien, die in letzter Zeit von der bildungspolitischen Abteilung des Europarats (Sprachenpolitisches Programm) zur Verfügung gestellt wurden, sowie nach ihrer Reaktion zu den vorgeschlagenen neuen Beispieskalen gefragt wurde. Die Mitgliedsstaaten wurden auch gebeten, Institutionen für die Pilotierung vorzuschlagen. Die Ergebnisse waren sehr positiv, außer einigen Vorbehalten betreffs der Benutzung des GeR in der Eingangsphase der Lehrerbildung – nur die Hälfte der Antwortenden sagten, sie sei in hohem Maße hilfreich gewesen. Wie zu vermuten war, wurden von den Dimensionen des GeR, auf die in offiziellen Dokumenten am häufigsten Bezug genommen wurde und die in der Praxis am häufigsten umgesetzt wurden, die Deskriptoren (bis zu 83 %), die Referenzniveaus (bis zu 75 %) und der handlungsorientierte Ansatz (bis zu 63 %) genannt. Auf die Frage, ob sie die neuen Skalen begrüßten, gab es den höchsten Wert für plurilinguale/plurikulturelle Kompetenz (79 %), gefolgt von Online-Interaktion (75 %) und danach Mediation (63 %) sowie Literatur (58 %).

### Einbeziehung von Deskriptoren für Gebärdensprache<sup>12</sup>

Menschen, die von Geburt an gehörlos sind, mögen eine Gebärdensprache als ihre erste Sprache lernen, sofern ihre Eltern und Altersgenossen entsprechenden Input geben. Gebärdensprachen sind nicht einfach eine Form gestenbasierter Kommunikation und nicht einfach ein anderes Medium, durch das gesprochene Sprache ausgedrückt wird. Die linguistische Forschung hat umfangreiche Belege für die Annahme geliefert, dass Gebärdensprachen menschliche Sprachen eigener Art sind, wie gesprochene Sprachen, und linguistische Merkmale, Mittel, Regeln und Einschränkungen aufweisen ähnlich denen in gesprochener Sprache. Solche Merkmale beinhalten Spracherwerb, Entwicklung, Verlust und alle anderen psychologischen Prozesse und sprachspezifischen Repräsentationen, die auch auf gesprochene Sprachen zutreffen.

Zeitgleich zum oben erwähnten Hauptprojekt wurden Deskriptoren für gebärdensprachliche Kompetenz entwickelt, wobei ähnliche Methoden angewandt wurden wie bei einem Projekt der Zürcher Hochschule für Angewandte Wissenschaften (ZHAW), das vom Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung finanziert worden war.<sup>13</sup> Das Projekt identifizierte und kalibrierte Deskriptoren für produktive und rezeptive gebärdensprachliche Kompetenz. Diese Deskriptoren beziehen sich speziell auf Gebärdensprachen und ergänzen die vorhandenen GeR-Deskriptoren.

Es sollte allerdings daran erinnert werden, dass die GeR-Deskriptoren, die Handlungsfähigkeit in der Sprache ausdrücken, auf alle menschlichen Sprachen zutreffen. Gebärdensprachen werden ebenso zur Ausübung von Handlungen verwendet wie gesprochene Sprachen. Daher können die gleichen Deskriptoren für beide Sprach-Modalitäten verwendet werden, und deshalb wurden alle GeR-Deskriptoren umformuliert, um modalitäts-neutral zu sein.

Schon seit der Einführung des GeR bestand ein Bedarf an definierten Lernzielen, Curricula und Niveaus für die Ausbildung in Gebärdensprachen. Tatsächlich wird der GeR zunehmend genutzt, um Kurse in Gebärdensprache zu

11 „Techno-CLIL 2017“, Moderation: Letitia Cinganotto & Daniela Cuccunulio <https://moodle4teachers.org/enrol/index.php?id=90>

12 Dieser Abschnitt wurde von der Autorengruppe an der ZHAW, Jörg Keller, Petrea Bürgin, Aline Meili und Dawei Ni auf Deutsch, Englisch und Französisch verfasst; die deutsche Fassung ist im Begleitband abgedruckt. Anm. der Übersetzer: In der Terminologie weicht diese Gruppe ein wenig vom Sprachgebrauch im GeR ab, z. B. „Validation“ statt „Validierung“.

13 Der Europarat dank dem Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung für die Ermöglichung des Projekts durch eine Zuwendung von € 385 000.



strukturieren. Die meisten gehörlosen Kinder (95 %) werden von hörenden Eltern zur Welt gebracht. Obgleich die Gemeinschaft der Gehörlosen insgesamt klein ist, besteht großer Bedarf an solchen Kursen nicht nur für Familien gehörloser Kinder, sondern auch für Bildungszwecke (Dolmetscher, gehörlose Migranten, Schwerhörige, Lehrende, Linguisten usw.). Darüber hinaus spielt der GeR inzwischen eine Rolle bei der Ausbildung und Qualifizierung von Gebärdensprachen-Lehrenden und Dolmetschern und insbesondere auch im Bemühen um offizielle Anerkennung von Gebärdensprachen und professionell Tätigen in diesem Bereich. Die Initiative, Deskriptoren für Gebärdensprache in den GeR zu integrieren, erhielt daher starke Unterstützung von zahlreichen Vereinigungen der Gehörlosen-Gemeinschaft.

Das ZHAW-Projekt für Gebärdensprache<sup>14</sup> (= GS) „Gemeinsamer Europäischer Referenzrahmen für Gebärdensprachen: Empirie-basierte Grundlagen für grammatische, pragmatische und soziolinguistische Deskriptoren in Deutschschweizer Gebärdensprache“ folgte einem anderen Zeitplan und schloss die Forschung im Juni 2019 ab, also 3 Jahre nach Abschluss des Hauptprojekts. Auch das GS-Projekt nutzte einen Mixed-Methods-Ansatz, bei dem intuitive, qualitative und quantitative Analysen im Forschungsdesign kombiniert wurden. Die drei Hauptphasen des Projekts sind in Abbildung 3 dargestellt.

Das GS-Projekt war gänzlich empirisch ausgelegt. Anstatt existierende Deskriptoren für Lautsprachen für GS zu adaptieren, zielte das GS-Projekt der ZHAW auf gebärdensprachliche Kompetenzbeschreibungen mit Analysen von Videoaufzeichnungen von Experten der Gebärdensprache. Die Gebärden der Experten bei verschiedenen Textsorten wurden gefilmt, und diese Aufzeichnungen wurden anschließend in einer Reihe von Workshops mit GS-Lehrkräften diskutiert. Das ZHAW-Projektteam formulierte auf der Basis der gesammelten Aussagen über die Videos und anhand der Analysen aus den Workshops Deskriptoren zur GS. Auf diese Weise wurden über 300 Deskriptoren für produktive Kompetenzen und 260 Deskriptoren für rezeptive Kompetenzen entwickelt. Ebenso wie im „Mediationsprojekt“ wurden zu diesem Zeitpunkt noch keinerlei Levels berücksichtigt: Ziel war es vielmehr, wichtige Aspekte von GS-Kompetenzen als Aussagen zu sammeln. Und wie im „Mediationsprojekt“ wurden die Deskriptoren später durch einen iterativen Prozess von Besprechung und Beratung im Rahmen von Workshops verbessert.

In einem einfachen Validierungsexperiment konnte anschließend gezeigt werden, dass hörende Personen, die nicht gebärden können, sowie gehörlose Personen, die keine GS-Lehrpersonen waren, ein signifikant anderes Verständnis der Deskriptoren hatten als gehörlose Lehrpersonen. Daher wurden die Deskriptoren ausschließlich von GS-Lehrkräften kalibriert, die gehörlos geboren waren oder denen L1-Kompetenz von der GS-Gemeinschaft auf der Basis der aufgezeichneten Performanz (Videos) zugesprochen worden war.

Die Deskriptoren wurden sodann in Kategorien eingeteilt. Ursprünglich war die Idee, verschiedene Skalen für Texttypen zu erstellen (erzählend, beschreibend, erklärend usw.).<sup>15</sup> Es zeigte sich jedoch, dass viele der Deskriptoren für mehrere Texttypen relevant sind, denn sie beschrieben transversale Kompetenzen. Deshalb wurden die Deskriptoren in einem Workshop durch das Projektteam in Sets auf der Grundlage von Ähnlichkeiten eingeteilt. Drei voneinander unabhängige Gruppen entwickelten dabei Kategorien, die ähnliche oder verwandte Kompetenzen umfassen sollten. Im Anschluss einigten sich drei Gruppen auf eine endgültige Kategorisierung. Die charakteristischen Merkmale jedes Sets wurden festgestellt und verfeinert, was zur Definition von Kategorien für 9 Skalen führte:

14 Zürcher Hochschule für Angewandte Wissenschaften. Autorengruppe: Jörg Keller, Petrea Bürgin, Aline Meili und Dawei Ni.

15 Keller, J., Meili, A., Bürgin, P. und Ni, D. (2017). „Auf dem Weg zum Gemeinsamen Europäischen Referenzrahmen (GER) für Gebärdensprachen. Empirie-basierte Bestimmung von Deskriptoren für Textkompetenz am Beispiel der Deutschschweizer Gebärdensprache (DSGS)“.

*DAS ZEICHEN* 105, 86–97.

Keller, J., Meili, A., Bürgin, P. und Ni, D. (2018). „Deskriptoren zur gebärdensprachlichen Textstrukturierung im GER für Gebärdensprachen“.

*DAS ZEICHEN* 109, 242–251.

Keller, J. (2019). „Deskriptoren für Textkompetenz in Gebärdensprachen“. In: Barras, Małgorzata; Karges, Katharina; Studer, Thomas und Wiedenkeller, Eva (Hrsg.): *IDT 2017*, Band 2. Berlin: ESV, 111–117.

Keller, J., Bürgin, P. Meili, A., Ni, D. und Hermann-Shores, P. (2018). *Kompetenzprofile für Gebärdensprachen*. Video-Dok. (68 Min.), Hochschule für Heilpädagogik, Zürich. Online: [https://www.youtube.com/watch?v=1YLovf46V2A&feature=youtu.be&list=PLRLsJ4CFI-AZDuyRlfu9o\\_lzjTi\\_Lyb01](https://www.youtube.com/watch?v=1YLovf46V2A&feature=youtu.be&list=PLRLsJ4CFI-AZDuyRlfu9o_lzjTi_Lyb01).

*Linguistische Kompetenz:*

1. Gebärdensprach-Repertoire (rezeptiv / produktiv)
2. Diagrammatische Korrektheit (rezeptiv / produktiv)

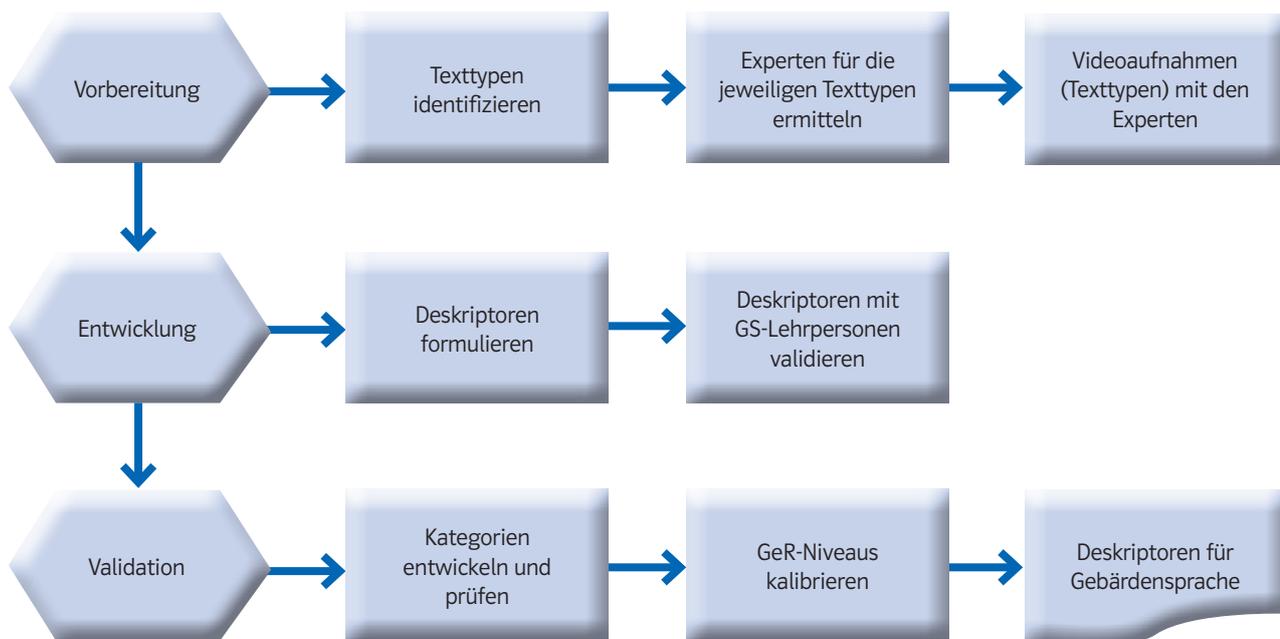
*Soziolinguistische Kompetenz:*

3. Soziolinguistische Angemessenheit und kulturelles Vorwissen (rezeptiv / produktiv)

*Pragmatische Kompetenz:*

4. Gebärden-Textstruktur; (rezeptiv / produktiv)
5. Setting und Perspektive (produktiv und rezeptiv)
6. Sprachbewusstheit und Interpretation (rezeptiv)
7. Auftritt und Wirkung (produktiv)
8. Verarbeitungstempo (rezeptiv)
9. Gebärdenfluss (produktiv)

**Abbildung 3: Die Phasen des ZHAW-GS-Projektes**



Der letzte Schritt bestand darin, die GeR-Niveaus zu kalibrieren. Um eine Deskriptoren-Skala zu erhalten, wurde ebenso wie im Mediationsprojekt, im Phonologie-Projekt und im ursprünglichen GeR-Projekt das Rasch-Modell genutzt. Im Unterschied zu diesen wurden jedoch Deskriptoren auf Videos beurteilt. Für die Kalibrierung wurden diese Deskriptoren-Videos auf DSGS (Deutschschweizer GS) und International Sign (IS) produziert. Letzteres ist eine Art *lingua franca* für GS-Lehrpersonen aus verschiedenen europäischen Ländern, die ebenfalls an der Online-Umfrage teilnahmen. Nach einer erfolgreichen Erprobung der Beurteilungsskala durch die Projektgruppe wurden die Teilnehmenden an der Umfrage aufgefordert, den Schwierigkeitsgrad der dargestellten Deskriptoren auf einer 4-Punkt-Skala von 1 (nicht schwierig) bis 4 (sehr schwierig) zu beurteilen.

Das gesamte Daten-Set (N = 223) wurde nach Fällen mit sehr wenigen oder keinen Bewertungen gefiltert. Diese Fälle wurden anschließend aus dem Set gelöscht. Die verbliebenen Bewertungen wurden in zwei Gruppen (Schweiz und Europa) aufgeteilt. In der Schweizer Gruppe (N = 53) waren alle Deskriptoren (über 300) von fast allen Teilnehmenden



bewertet worden. In der europäischen Gruppe (N = 37) wurde nur ein Teil der Deskriptoren evaluiert, was durchschnittlich 15 Bewertungen pro Deskriptor ergab.<sup>16</sup>

Wie bereits oben in der kurzen Beschreibung des Rasch-Modells erwähnt, werden Deskriptoren den entsprechenden Niveaus genauer zugeordnet, wenn inkonsistent beurteilende Personen und inkonsistent beurteilte Items aus der Analyse ausscheiden. Dieses Verfahren wurde in diesem Projekt wie auch im Hauptprojekt angewendet.

In einem abschließenden Schritt wurden die Niveau-Grenzen auf der Gebärdensprachen-Skala bestimmt. Zur Vereinfachung wurden bereits kalibrierte „Anker-Deskriptoren“ aus dem lautsprachlichen GeR von 2001 in die GS-Kalibrierung aufgenommen. Dies sollte eine mathematische Transformation der Kalibrierungswerte der neuen GS-Skala auf die bereits existierende GeR-Skala ermöglichen. Eine Erklärung dieses Vorgehens findet man in den Abschnitten über quantitative Validierung in: *Developing illustrative descriptors of aspects of mediation for the CEFR* (North und Piccardo 2016) und *Phonological Scale Revision Process Report* (Piccardo 2016). Anders als bei diesen beiden Projekten waren die mathematischen Ergebnisse für die lautsprachlichen Anker-Deskriptoren im GS-Projekt jedoch nicht plausibel und glaubwürdig, selbst nachdem instabile Anker entfernt worden waren. Deshalb wurde eine andere Methode zur Bestimmung der Standard-Niveau-Grenzen basierend auf Expertenmeinungen angewendet.<sup>17</sup>

## Finalisierung

Die Rückmeldungen während der verschiedenen Phasen der Validierung, Konsultation und Pilotierung zwischen Februar 2015 und Juni 2017 waren sehr hilfreich bei der Identifizierung und Beseitigung von weniger erfolgreichen Deskriptoren und Skalen und bei der Überarbeitung von Formulierungen. Dieser Prozess ist in einem Archiv dokumentiert, das der Forschung auf der Europarats-Website zur Verfügung steht. Die endgültige Version der Deskriptoren hat allen Rückmeldungen Rechnung getragen.

Weil sehr viele Deskriptoren für bestimmte Niveaus mancher Skalen validiert wurden, besonders für B2, wurden einige von der ausführlichen Version der Beispielskalen ausgeschlossen, obgleich es sich um erfolgreich validierte Deskriptoren handelt. Sie sind in Anhang 8 aufgeführt. An sich ist diese Redundanz gut, weil sie die Kohärenz der Kalibrierung zu verschiedenen Niveaus unterstreicht, dennoch ist es nicht notwendig, alle betreffenden Deskriptoren in die endgültigen GeR-Beispielskalen aufzunehmen. Zu einem späteren Zeitpunkt werden sie als ergänzende Deskriptoren in der Deskriptoren-Datenbank für den GeR zu finden sein, die auf der Website des Europarats zur Verfügung steht.

16 Obgleich sie klein sind, entsprechen diese Werte der minimalen a-priori-Anforderung für 95%-Konfidenzintervalle bei Schwierigkeitsparametern, nämlich einer Bandbreite von +/- 1 logit (Linacre, 1994: „Sample Size and Item Calibration Stability“. *Rasch Measurement Transactions* vol. 7, no 4, p. 328). Der Standardmessfehler für die Gebärdensprache-Deskriptoren ist größer als der für andere Deskriptoren, aber eine Kalibrierung auf der Skala ist intuitiv sinnvoll. In wenigen Fällen wurden Deskriptoren, die im Irrtumsbereich zur nächsten Kompetenzstufe lagen, auf Grundlage eines gemeinsamen Expertenurteils auf das angrenzende Niveau gesetzt.

17 Die angewandte Methode war eine Variante der *Bookmark Method*, die in *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) – A Manual* (Council of Europe 2009) erläutert wird.