

ANBINDUNG DES SPRACH- UNTERRICHTS AN DEN GER: EIN HANDBUCH

April 2022

Sprachen lehren, lernen, prüfen und der GER: Ein Handbuch

Dieses Handbuch richtet sich an alle, die unterrichten, testen und Materialien für den Sprachunterricht entwickeln, sowie an alle Beteiligten, die mit Bildungsrichtlinien zu tun haben und zur Entscheidungsfindung beitragen.

Die Verfasserinnen und Verfasser gehen davon aus, dass viele Mitwirkende im Bereich des Sprachunterrichts diese erste Version des Handbuchs im Rahmen von Projekten zur Anbindung an den GeR bald verwenden werden. Dies liegt an der großen Aufmerksamkeit, die sowohl der GeR 2001 als auch der GeR-Begleitband 2020 erzielt haben. Alle an solchen Projekten Beteiligte sind eingeladen, sowohl ihre Erfahrungen zu teilen als auch Rückmeldung zu geben, wie hilfreich das Handbuch ist. Für 2024 ist eine Veranstaltung geplant, bei der Fallstudien zur Verwendung des Handbuchs präsentiert werden sollen. Diese Ergebnisse und weitere Vorschläge werden in einer überarbeiteten Ausgabe des Handbuchs berücksichtigt.

Zur deutschen Übersetzung

Dieses Handbuch befasst sich mit den wichtigsten Schritten der Anbindung (engl. *alignment*) von Materialien wie Curricula, Lehrwerke und Sprachprüfungen an den Gemeinsamen europäischen Referenzrahmen für Sprachen (GeR) von 2001 und an den 2020 auf Deutsch erschienenen Begleitband. Das Handbuch liefert sowohl theoretische Hintergründe als auch praktische Hinweise für den Prozess der Anbindung.

Die Übersetzung des Handbuchs war ein spannendes, aber auch aufwendiges Projekt. Ziel war es, bei der Übersetzung von Fachbegriffen die schon etablierten Begriffe aus dem GeR bzw. dem Begleitband zu übernehmen. Auch die deutsche Übersetzung der von der Association of Language Testers in Europe (ALTE) im Auftrag des Europarats erstellten Publikation *Manual for Language Test Development and Examining* (dt. *Handbuch zur Entwicklung und Durchführung von Sprachtests*) diente als Orientierung.

Im Übersetzungsprozess standen durchgängig die Verständlichkeit und der Lesefluss für das deutschsprachige Publikum im Fokus. Dabei stellte sich die Übersetzung mancher Textstellen oder Begriffe als komplex heraus. Zu diesen Wörtern zählten beispielsweise die zentralen Begriffe *alignment*, *claim* und *resource*. Wir haben uns bei diesen Wörtern je nach Kontext für folgende Übersetzungsvarianten entschieden:

- *Alignment* – Anbindung, Verlinkung, Ausrichtung
- *Claim* – Anspruch, Behauptung, aber auch in bestimmten Kontexten Argumentation
- *Resource* – Ressource, wenn es um Mittel wie Finanzen oder Personal geht. Meist jedoch mit dem Oberbegriff Material, da es um die Projekte oder Produkte geht, die an den GeR angebunden werden sollen.

Bei Termini, für die zwar ein deutsches Wort existiert, jedoch im fachlichen Kontext meist das englische Wort verwendet wird, haben wir uns für die etablierte englische Variante entschieden. Als Beispiel sei hier *Standard Setting* genannt – im deutschen Text also Standard-Setting und nicht Standardsetzung oder Ähnliches.

Bei einer Übersetzung ins Deutsche spielt zudem der Aspekt der geschlechtergerechten Sprache eine Rolle. Der Text ist weitestmöglich neutral formuliert. Wenn es notwendig war, Personen oder Berufe zu benennen, sind beide Formen aufgeführt.

Das Handbuch wird in Zusammenarbeit mit der European Association for Language Testing and Assessment (EALTA), der UK Association for Language Testing and Assessment (UKALTA), dem British Council und der Association for Language Testers in Europe (ALTE) herausgegeben.

ISBN: 978-1-7397544-1-9

© British Council, UKALTA, EALTA und ALTE. 2022.

INHALT

Vorwort	6
KAPITEL 1: DER GER	8
1.1 Was ist der Europarat?	9
1.2 Was ist der Gemeinsame europäische Referenzrahmen für Sprachen (GeR)?.....	10
1.3 Worum geht es bei der Anbindung an den GeR und warum ist sie sinnvoll?.....	11
1.4 Warum ist es wichtig, die verschiedenen Komponenten eines Lernsystems aneinander anzupassen?.....	12
1.5 Was umfasst die Anbindung an den GeR?	13
1.6 Welche Publikationen stehen für die Anbindung an den GeR zur Verfügung?.....	15
1.7 Welche Orientierungshilfen bietet das Handbuch?	16
Hinweise zur eigenen Implementierung	17
KAPITEL 2: FAMILIARISIERUNG	18
2.1 Was versteht man unter Familiarisierung?	19
2.2 Allgemeine Aktivitäten zur Familiarisierung	20
2.3 Spezifische Aktivitäten zur Familiarisierung	21
2.3.1 Rollen und Entscheidungen von Koordinatorinnen und Koordinatoren	21
2.3.2 Vorbereitung des Materials	23
2.4 Leitfaden für Berichte über die Familiarisierungsphase	26
2.4.1 Berichte zur internen Verwendung	26
2.4.2 Berichte zur externen Verwendung	27
Hinweise zur eigenen Implementierung	27
KAPITEL 3: SPEZIFIKATION	28
3.1 Was versteht man unter Spezifikation?	29
3.2 Allgemeine Aktivitäten zur Spezifikation	30
3.3 Spezifische Aktivitäten zur Spezifikation.....	32
3.4 Leitfaden für Berichte über die Spezifikationsphase	35
Hinweise zur eigenen Implementierung	36
KAPITEL 4: STANDARDISIERUNG	37
4.1 Was versteht man unter Standardisierung?	38
4.2 Organisation des Standardisierungsprozesses.....	39
4.3 Schritte bei der Standardisierung	40
4.3.1 Aktivitäten.....	42
4.3.2 Tools (Hilfsmittel)	44
4.4 Leitfaden für Berichte über die Standardisierungsphase	44
Hinweise zur eigenen Implementierung	45

KAPITEL 5: STANDARD-SETTING	46
5.1 Was versteht man unter Standard-Setting?	47
5.1.1 Warum setzen wir Standards?	47
5.1.2 Wie wird das Standard-Setting durchgeführt?	48
5.1.3 Praktisches zum Standard-Setting: Planung, Organisation, Datenerhebung und Analyse	50
5.2 Beispiele zum Standard-Setting für Rezeption und Produktion	52
5.2.1 Beispiel für das Standard-Setting für Rezeption	52
5.2.2 Beispiel zum Standard-Setting für Produktion	54
5.3 Herausforderungen, Lösungen und Alternativen	58
5.3.1 Anpassung des Umfangs	58
5.3.2 Anpassungen auf vertretbare Weise vornehmen	59
5.3.3 Online-Standard-Setting	59
5.3.4 Weitere Methoden zum Standard-Setting	60
5.3.5 Sprachenübergreifendes Standard-Setting	61
5.4 Anpassung der Methoden zum Standard-Setting bei anderen Materialien als Prüfungen	61
5.5 Leitfaden für Berichte über die Standard-Settingphase	62
5.6 Weiterführende Literatur	63
Hinweise zur eigenen Implementierung	65
KAPITEL 6: VALIDIERUNG DER ARGUMENTATION	66
6.1 Definition von Validierung	67
6.2 Wesentliche Einschränkungen der Validierungsargumente zur Anbindung	67
6.3 Aufbau einer Evidenzbasis für die Validierung	68
6.4 Hinweis zur Struktur des Berichts	69
6.5 Leitfaden für Berichte über die Validierungsphase	70
6.6 Hinweis zur Anbindung von Tests, die KI-unterstützte Bewertungssysteme verwenden	74
Hinweise zur eigenen Implementierung	75
ANHANG: DER PROZESS DER ANBINDUNG UND ARGUMENTATION – FORMULARE	76
Dokumentationsbogen zu Kapitel 2: Familiarisierung	77
Dokumentationsbogen 2.1: Allgemeine Familiarisierung	77
Dokumentationsbogen 2.2: Spezifische Familiarisierung – Sitzung 1	79
Dokumentationsbogen 2.3: Allgemeine Zusammenfassung des Familiarisierungsprozesses	81
Dokumentationsbogen zu Kapitel 3: Spezifikation	83
Dokumentationsbogen 3.1: Beschreibung der Verlinkung und des Materials	83
Dokumentationsbogen 3.2: Sprachaktivitäten, Strategien und Kompetenzen	84
Dokumentationsbogen 3.3: Grafische Profilierung der Verlinkung	85
Dokumentationsbogen 3.3: Grafische Profilierung der Verlinkung	85
Dokumentationsbogen 3.4: Illustrative Deskriptoren von Aktivitäten und Strategien beim Spracherwerb	86
Dokumentationsbogen 3.5: Detaillierte Zuordnung	87
Dokumentationsbogen zu Kapitel 4: Standardisierung	88
Dokumentationsbogen 4.1: Zusammenfassung der Standardisierung – Teilnehmende	88
Dokumentationsbogen 4.2: Zusammenfassung der Standardisierung – Koordinatorinnen/Koordinatoren	90
Dokumentationsbogen 4.3: Analyse illustrativer Beispiele – Holistisches Formular	92
Dokumentationsbogen zu Kapitel 5: Standard-Setting	93

Liste der Abbildungen

Abbildung 1.1: Mitgliedsstaaten des Europarats.....	9
Abbildung 1.2: Das „Comprehensive Learning System“ (CLS) von O’Sullivan.....	12
Abbildung 1.3: Schritte im Anbindungsprozess	14
Abbildung 2.1: Allgemeine und spezifische Familiarisierung	19
Abbildung 2.2: Fünf allgemeine Aktivitäten zur Familiarisierung	20
Abbildung 2.3: Aktivitäten für die erste Sitzung zur Familiarisierung.....	23
Abbildung 2.4: Aktivitäten für weitere Sitzungen zur spezifischen Familiarisierung.....	26
Abbildung 3.1: Spezifikationsprozess	29
Abbildung 3.2: Schritte im allgemeinen Spezifikationsprozess.....	30
Abbildung 3.3: Vier Modi der Kommunikation.....	31
Abbildung 3.4: Übersicht über den spezifischen Spezifikationsprozess	32
Abbildung 3.5: Ausgewählte Kannbeschreibungen der Skala „Kreatives Schreiben“ - Niveau A1, A2 und A2+	34
Abbildung 6.1: Übersicht zur Validierung der Verlinkungsbehauptung [C – Curriculum, D – Durchführung, A – Prüfung (engl. <i>Assessment</i>)]	68
Abbildung 6.2: Hin zu einer starken Behauptung.....	69

Liste der Tabellen

Tabelle 4.1: Schritte im Standardisierungsprozess	40
Tabelle 5.1: Wichtige Schritte bei der Planung und Vorbereitung einer Aktivität zum Standard-Setting.....	50
Tabelle 5.2: Beispielzeitplan für Aktivitäten zu einem Test der Fertigkeit Hören	53
Tabelle 5.3: Zeitplan für Aktivitäten zum Prüfungsteil Sprechen und Schreiben.....	55
Tabelle 6.1: Vorschlag für die Herangehensweise beim Verfassen des Validierungsberichts.....	70
Tabelle 6.2: Übersicht über die Phasen und möglichen Behauptungen.....	71
Tabelle 6.3: Übersicht über die Schritte, die für die verschiedenen Arten von Materialien erforderlich sind	72

Vorwort

Der [Gemeinsame europäische Referenzrahmen für Sprachen](#) (GeR) wurde 2001 herausgegeben.

Insbesondere Sprachinstitute begrüßten die gemeinsamen Referenzstufen und illustrativen Skalen, wenngleich sich Expertinnen und Experten für Sprachunterricht und Stakeholder bald fragten, woher man beispielsweise wissen könne, dass alle Mitwirkenden unter der Niveaustufe B1 dasselbe verstünden. Der Europarat forderte daraufhin, ein Validierungssystem zu entwickeln, das die Anbindung von Sprachprüfungen an eine oder mehrere Stufen des GeR bestätigen würde. Ein solches System lag jedoch nicht im Zuständigkeitsbereich des Europarats, der stattdessen zwei Leitfäden veröffentlichte – einen zur Verlinkung von Sprachtests mit den GeR¹ und einen zur Erstellung von Tests unter Berücksichtigung des GeR.²

2013 gab es ein Projekt zur Überarbeitung der illustrativen Skalen des GeR. Eine [vorläufige Version](#) des GeR-Begleitbands erschien 2017 und die [endgültige Version](#) ist seit 2020 verfügbar. Ein EALTA-Symposium in Dublin im Januar 2018 sowie ein vom British Council durchgeführtes EALTA-UKALTA-Symposium in London im Februar 2020 bestätigten jeweils den Bedarf nach einem neuen Handbuch zur Unterstützung der Anbindung an das erweiterte deskriptive Schema des GeR-Begleitbands. Das Steuerungsgremium, das sich bei der Veranstaltung im Februar 2020 konstituierte, erkannte während der Produktion des Handbuchs, dass die Anbindung nicht nur für Sprachprüfungen, sondern auch für politische Richtlinien, Bildungs- und Lehrpläne, Lehrwerke und andere Lehr- und Lernmaterialien relevant ist.

Die Gruppe beschloss zudem, dass das Handbuch der Information von Entscheidungsträgerinnen und -trägern, in der Lehrkräfteausbildung, von Lehrkräften und anderen Stakeholdern im Sprachunterricht dienen soll. Außerdem hat es die Aufgabe, die mehr oder weniger technischen Prozesse, von denen die Verlinkung abhängt, zu unterstützen. Das Handbuch soll den Nutzerinnen und Nutzern auch dabei helfen, sich einen Überblick über die zahlreichen inzwischen verfügbaren Referenzdokumente zum GeR zu verschaffen.

Dabei richtet sich das Handbuch an **zwei Zielgruppen**, die jeweils mehrere spezialisierte Untergruppen umfassen (siehe unten). Die Aufzählung erhebt keinen Anspruch auf Vollständigkeit.

ZIELGRUPPE 1	ZIELGRUPPE 2
<p>Alle, die eine praktische oder angewandte Aufgabe zur Anbindung des Sprachunterrichts an den GeR in einem bestimmten Kontext und zu einem bestimmten Zweck durchführen wollen (z. B. um die Anbindung an den GeR nachzuweisen oder zu prüfen).</p> <p>Dazu gehören:</p> <ul style="list-style-type: none">▪ Lehrkräfte im Dienst▪ Entwicklerinnen/Entwickler von Lernmaterialien▪ Erstellerinnen/Ersteller von Lehrplänen▪ Lehrkräfte in der Ausbildung▪ Entwicklerinnen/Entwicklerinnen von Prüfungen▪ Autorinnen/Autoren von Lehrbüchern	<p>Andere Stakeholder im Bildungswesen oder in der Gesellschaft insgesamt, die sich vorwiegend mit politischen Fragen und der Entscheidungsfindung bezüglich der Sprachvermittlung im Allgemeinen befassen (z. B. um die Aussagen von Lehrbuchverlagen und Testinstitutionen hinsichtlich des GeR-Niveaus zu überprüfen).</p> <p>Dazu gehören Personen aus den Bereichen:</p> <ul style="list-style-type: none">▪ Administration▪ Curriculumserstellung▪ Bildungsministerium▪ Arbeitgeber▪ Management▪ Politik▪ Verlage▪ Ausbildung von Lehrkräften

1 [Relating Language Examinations to the 'Common European Framework of Reference for Languages: Learning, Teaching, Assessment' \(CEFR\). A Manual](#), Straßburg: Europarat, überarbeitete Version 2009.

2 [Manual for Language Test Development and Examining](#), Straßburg: Europarat, 2011.

Die Kapitel 1 und 2 des Handbuchs sind allgemein gehalten und für beide Zielgruppen von unmittelbarer Relevanz. Sie bieten eine grundlegende Einführung in den GeR und den GeR-Verlinkungsprozess sowie einen grundlegenden Überblick über das Thema. Die Kapitel 3 bis 6 bieten sowohl allgemeine als auch detaillierte Informationen zu den aufeinanderfolgenden Schritten im Rahmen eines Verlinkungsprozesses. Zudem werden die verfügbaren Tools und Ansätze bei der Berichterstellung zu den ergriffenen Maßnahmen vorgestellt.

Das Handbuch wurde folgendermaßen erstellt: Das Steuerungsgremium einigte sich zunächst auf den Umfang, erstellte einen Entwurf der Kapitel 1 und 2 und betraute daraufhin ausgewählte Fachleute im Bereich des Sprachunterrichts mit einer Begutachtung. Anschließend wurden Expertinnen und Experten mit umfassender Erfahrung in der Arbeit mit dem GeR gebeten, einen Entwurf der Kapitel 3 bis 6 zu erstellen. David Little überarbeitete dann den Entwurf des Handbuchs. Diejenigen Fachleute, die bereits die ersten Kapitelentwürfe 1 und 2 begutachtet hatten, lektorierten im nächsten Schritt Littles überarbeitete Version. Zum Schluss wurden ihre Kommentare und Vorschläge in das Handbuch eingearbeitet.

<p>Steuerungsgremium</p> <ul style="list-style-type: none"> ▪ Neus Figueras, European Association for Language Testing and Assessment (EALTA) ▪ David Little, Trinity College Dublin ▪ Barry O’Sullivan, British Council ▪ Nick Saville, Association of Language Testers in Europe (ALTE) ▪ Lynda Taylor, UK Association of Language Testing and Assessment (UKALTA) 	<p>Redaktion</p> <ul style="list-style-type: none"> ▪ Neus Figueras, European Association for Language Testing and Assessment (EALTA) ▪ David Little, Trinity College Dublin ▪ Barry O’Sullivan, British Council
<p>Autorenteam</p> <ul style="list-style-type: none"> ▪ Glyn Jones, freiberufliche Beraterin, GBR ▪ Elif Kantarcioğlu, Bilkent-Universität, Türkei ▪ Noriko Nagai, Ibaraki-Universität, Japan ▪ Graham Seed, Cambridge Assessment English, GBR ▪ Jamie Dunlea, British Council, GBR ▪ Neus Figueras, European Association for Language Testing and Assessment (EALTA) ▪ David Little, Trinity College Dublin, Irland ▪ Barry O’Sullivan, British Council, GBR ▪ Nick Saville, Association of Language Testers in Europe (ALTE) ▪ Lynda Taylor, UK Association of Language Testing and Assessment (UKALTA) 	<p>Gutachterinnen und Gutachter</p> <ul style="list-style-type: none"> ▪ Elaine Boyd, freiberufliche Beraterin, GBR ▪ Bronagh Ćatibušić, Trinity College Dublin, Irland ▪ Karen Dunn, British Council, GBR ▪ Peter Lenz, Institut für Mehrsprachigkeit, Universität Freiburg, Schweiz ▪ Vincent Folny, France Éducation International, Frankreich ▪ Jane Lloyd, Cambridge University Press & Assessment, GBR ▪ Waldemar Martyniuk, Uniwersytet Jagiellonski w Krakowie, Polen ▪ Josep Ma. Montcada, Generalitat de Catalunya, Spanien ▪ Felicity O’Dell, LT123, GBR ▪ Edgar Rene Pacheco, Centro de Enseñanza para Extranjeros (CEPE), UNAM, Mexiko ▪ Lorenzo Rocca, Società Dante Alighieri, Italien ▪ Cristina Rodriguez, Xunta de Galicia, Spanien ▪ Wu Sha, National Education Examination Authority (NEEA), China ▪ Joe Sheils, ehem. Leiter des Bereichs für Sprachpolitik im Europarat ▪ Richard Spiby, British Council, GBR ▪ Dina Vilcu, Universität Babeş-Bolyai, Rumänien ▪ Kateřina Vodičková, Karls-Universität Prag, Tschechien ▪ Caroline Westbrook, British Council, GBR ▪ Russell Whitehead, LT123, GBR ▪ Jessica Wu, Language Training and Testing Center (LTTC), Taiwan

KAPITEL 1: DER GER

DER EUROPARAT, DER GEMEINSAME REFERENZRAHMEN FÜR SPRACHEN SOWIE DIE ANBINDUNG AN DEN GER

In Kapitel 1 wird der Hintergrund des *Gemeinsamen europäischen Referenzrahmens für Sprachen (GeR)* beleuchtet und erklärt, wie er für Verlinkungsprozesse beim Unterrichten einer Zweitsprache verwendet werden kann. Folgenden Fragen werden beleuchtet:

- 1.1 Was ist der Europarat?
- 1.2 Was ist der Gemeinsame europäische Referenzrahmen für Sprachen (GeR)?
- 1.3 Worum geht es bei der Anbindung an den GeR und warum ist sie sinnvoll?
- 1.4 Warum ist es wichtig, die verschiedenen Komponenten eines Lernsystems aneinander anzupassen?
- 1.5 Was umfasst die Anbindung an den GeR?
- 1.6 Welche Publikationen stehen für die Anbindung an den GeR zur Verfügung?
- 1.7 Welche Orientierungshilfen bietet das Handbuch?

Am Ende des Kapitels sollten Sie die folgenden Fragen beantworten können:

- Was ist der Europarat und welche Bildungsziele verfolgt er?
- Welchem Zweck dient der GeR und welche Inhalte umfasst er?
- Warum ist die Anbindung an den GeR für Sie relevant und hilfreich?
- Warum und wie können Sie eine Anbindung an den GeR in Ihrem Kontext vornehmen?

1.1 Was ist der Europarat?

Der **Europarat** mit seinem Hauptsitz in Strasbourg, Frankreich, wurde 1949 gegründet, um sich für Menschenrechte, Demokratie und Rechtsstaatlichkeit einzusetzen. Heute umfasst er 46 Mitgliedsstaaten – von Island im Westen bis hin zu Aserbaidschan im Osten. (Der Europarat sollte nicht mit dem Europäischen Rat verwechselt werden, bei dem es sich um das Gremium der Staats- und Regierungschefs der Europäischen Union handelt. Alle Mitgliedsstaaten der Europäischen Union sind auch Mitglieder des Europarats.)

Die **wichtigsten Instrumente** des Europarats sind die Europäische Menschenrechtskonvention (1950), das Europäische Kulturabkommen (1954) und die Europäische Sozialcharta (1961, Überarbeitung 1996).



Abbildung 1.1: Mitgliedsstaaten des Europarats

Die **Bildungsprogramme des Europarats** basieren auf den geltenden Menschenrechten und Rechtsnormen und zielen darauf ab, die demokratischen Kompetenzen der Bürgerinnen und Bürger zu stärken, den Respekt für die Würde und Vielfalt der Menschen zu fördern und Diskriminierung zu bekämpfen.

Seit den 1970er Jahren setzt sich der Europarat für das **Lernen und Lehren von Sprachen zum Zweck der Kommunikation** ein, um

1. die sprachliche und kulturelle Vielfalt Europas zu wahren und Bürgerinnen und Bürger dazu zu ermutigen, diese Vielfalt wertzuschätzen.
2. Kommunikation und Austausch auf internationaler Ebene zu erleichtern.

Das Konzept des Europarats der **plurilingualen und interkulturellen Bildung** verbindet das Lehren und Erlernen einer Sprache explizit mit den Grundwerten und politischen Zielen des Rats, insbesondere mit der sozialen Integration, dem Respekt für Vielfalt und dem interkulturellen Dialog.

1.2 Was ist der Gemeinsame europäische Referenzrahmen für Sprachen (GeR)?

Der Gemeinsame europäische Referenzrahmen für Sprachen (GeR) wurde 2001 erstmalig veröffentlicht. Er ermöglicht eine kritische Betrachtung der verschiedenen Ebenen des Sprachunterrichts – Bildungs- und Lehrpläne, Lehrwerke und Prüfungen – und stellt eine gemeinsame Basis für die Bildungssysteme der Mitgliedsstaaten des Europarats dar.

Die zentrale Komponente des GeR von 2001 ist eine detaillierte taxonomische Beschreibung (Klassifikation) der Sprachkompetenz hinsichtlich des Sprachgebrauchs. Sie ist in vier Bereiche unterteilt:

- Rezeption (Hören und Lesen)
- Produktion (Sprechen und Schreiben)
- Interaktion (in Wort und Schrift)
- Mediation (d. h. Ermöglichung einer Kommunikation zwischen Einzelpersonen oder Gruppen, die aus irgendeinem Grund nicht direkt miteinander kommunizieren können)

Das **nicht sprachspezifische, deskriptive Schema** des GeR hat zwei zusätzliche Dimensionen:

(1) Die kommunikativen Aufgaben, die Sprachnutzende und -lernende ausführen müssen, und (2) die Kompetenzen, von denen die erfolgreiche Erfüllung der Aufgabe abhängt. Beide Dimensionen umfassen illustrative Skalen mit Kann-Beschreibungen („can do“-Statements), um die Kenntnisse auf sechs Ebenen zu beschreiben: A1 und A2 (elementare Sprachverwendung), B1 und B2 (selbständige Sprachverwendung) sowie C1 und C2 (kompetente Sprachverwendung).

Während sich der Europarat mit den Rechten und Pflichten der einzelnen Bürgerinnen und Bürger befasst, sieht der GeR die Sprachnutzenden und -lernenden als soziale Akteurinnen und Akteure – d.h. als Mitglieder einer Gesellschaft, die unter bestimmten Umständen, in einer bestimmten Umgebung und innerhalb eines bestimmten Handlungsfelds Aufgaben erfüllen müssen (nicht zwingend sprachbezogen).

Die endgültige Version des **GeR-Begleitbands**, 2020 vom Europarat veröffentlicht, umfasst Aktualisierungen und Erweiterungen der illustrativen Deskriptoren. Er enthält neue Deskriptoren zu den Skalen von 2001, einige neue Skalen und eine neue Stufe vor A1. Zusätzlich wurden Skalen für plurilinguale und plurikulturelle Kompetenzen hinzugefügt und Mediation anhand von 24 illustrativen Skalen ausführlich definiert. Die Skala zur Phonologie von 2001 wurde durch drei neue Skalen ersetzt. Die Formulierung der Deskriptoren ist geschlechtsneutral sowie hinsichtlich der Modalität inklusiv (und beziehen sich somit sowohl auf Gebärden- als auch auf Lautsprache).

Im gesamten Handbuch sollten Hinweise auf den GeR als Referenz zu den Dokumenten von 2001 und 2020 verstanden werden.

1.3 Worum geht es bei der Anbindung an den GeR und warum ist sie sinnvoll?

Der GeR bietet eine **umfassende, einheitliche und transparente Beschreibung von Sprachkenntnissen** hinsichtlich des Sprachgebrauchs. Mithilfe der einheitlichen Referenzstufen und illustrativen Skalen kann diese Beschreibung für Folgendes verwendet werden:

- Bildungs- und Lehrpläne, Lehrwerke und andere Lehr-/Lernmaterialien, Prüfungen sowie andere Formen der Leistungsmessung erstellen bzw. entwickeln
- Aufzeigen, dass zwei oder mehrere dieser Aspekte des Sprachunterrichts aufeinander abgestimmt sind
- Zeigen, dass bestehende Bildungs- oder Lehrpläne, Lehrwerke, Prüfungen oder andere Formen der Leistungsmessung an relevante Teile des beschreibenden GeR-Schemas und mindestens eine seiner Kompetenzstufen angepasst wurden
- Bildungs- und Lehrpläne, Lehrwerke und andere Lehr-/Lernmaterialien, Prüfungen sowie andere Formen der Leistungsmessung miteinander vergleichen

Seit der Veröffentlichung im Jahr 2001 haben Behörden, Agenturen und Institutionen im Bildungsbereich erkannt, dass sich durch die **Verlinkung mit dem GeR** gewisse Vorteile hinsichtlich verschiedener Aspekte ihres Bildungssystems ergeben.

Zu den wahrgenommenen Vorteilen gehören:

- Erreichen von systemischer Kohärenz und Transparenz
- Schaffung einer Basis für prinzipielle Vergleiche
- Monitoring zum Zweck der Qualitätssicherung

Die GeR-Verlinkung fördert ein gemeinsames Verständnis und Kohärenz hinsichtlich verschiedener Aspekte eines Bildungssystems, erleichtert deren Vergleich und unterstützt die Qualitätssteigerung. Aus diesen Gründen ist sie von Interesse für:

- Institutionen mit dem Auftrag zur Erstellung nationaler/regionaler Bildungs- und Lehrpläne
- Institutionen mit dem Auftrag zur Entwicklung und Verwaltung nationaler/regionaler Sprachprüfungen und anderer Formen der Leistungsmessung
- Autonome Bildungseinrichtungen, insbesondere Universitäten und private Sprachschulen
- Lehrkräfte, die ihre eigenen Lehrmaterialien und/oder Prüfungen vorbereiten müssen
- Ausbilderinnen und Ausbilder von Lehrkräften
- Sprachlehrkräfte, die in einem Kontext arbeiten, der teilweise oder vollständig von den Vorgaben des GeR und des GeR-Begleitbands geprägt ist
- Nationale und internationale Institute für Sprachprüfungen
- Nationale und regionale Bildungsministerien
- Berufsverbände für Qualitätssicherung und -optimierung in der Sprachausbildung
- Autorinnen, Autoren und Herausgebende von Lehrwerken

Für alle, die mit der Arbeit des Europarats und/oder mit dem GeR noch nicht vertraut sind, ist es wichtig darauf hinzuweisen, dass eine Anbindung an das GeR-Niveau **nicht** als obligatorisch oder verpflichtend angesehen wird. Der GeR stellt somit eine Orientierungshilfe dar und kein fertiges Paket zur Implementierung.

1.4 Warum ist es wichtig, die verschiedenen Komponenten eines Lernsystems aneinander anzupassen?

Laut O’Sullivan (2020) hängt der Erfolg jedes Lernsystems von der engen Abstimmung von Bereichen ab, die traditionell als unabhängig voneinander gesehen wurden: Bildungspläne, Lehr-/Lernmaterialien, Lehransätze, Ausbildung von Lehrkräften und Prüfungen. Auch wenn diese Bereiche sich sowohl deutlich gegenseitig als auch das Lernen beeinflussen, arbeiten Entwicklerinnen und Entwickler von Bildungsplänen und Lernmaterialien, Ausbilderinnen und Ausbilder von Lehrkräften und Prüfungsspezialistinnen und -spezialisten meistens unabhängig voneinander. Die Notwendigkeit, diese Bereiche aus einer einheitlichen Perspektive zu betrachten, bildet die Grundlage für O’Sullivans (2020) Konzept des „[Comprehensive Learning System](#)“ (CLS).

Abbildung 1.2 verdeutlicht die drei Kernbereiche des CLS:

1. **Curriculum** – informell und formell
2. **Durchführung** – umfasst die Auswahl von Lehrkräften, Akkreditierung, berufliche Entwicklung und Führungsqualitäten, Lehr-/Lernmaterialien sowie die räumliche Umgebung, in der die Durchführung stattfindet
3. **Beurteilung** – umfasst die entwicklungsbezogene Beurteilung (Diagnostik, Fortschrittskriterien, formative Beurteilung usw.) sowie die bewertende Beurteilung (Einstufung, Fortschrittskriterien, Leistung, Kenntnisse usw.)

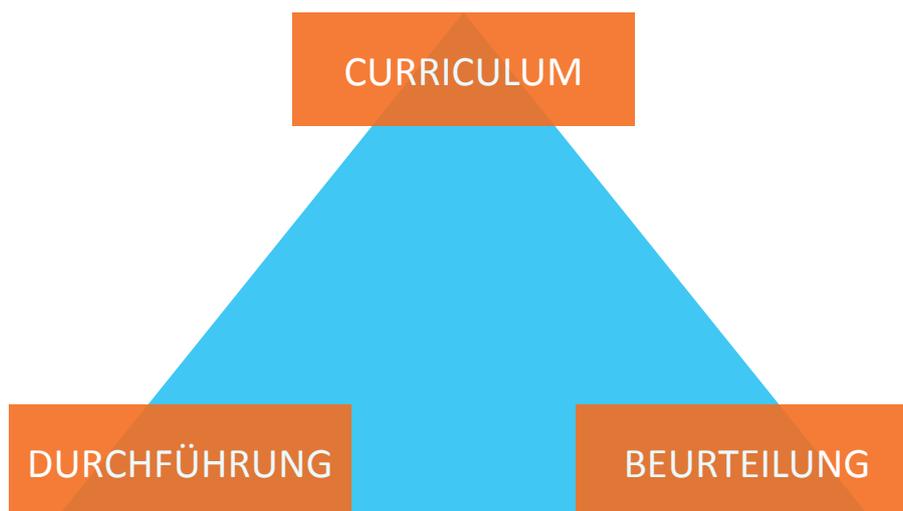


Abbildung 1.2: Das „Comprehensive Learning System“ (CLS) von O’Sullivan

Wenn einer dieser Bereiche in irgendeiner Weise von den anderen abgetrennt ist, ist das System bedroht. Nachfolgend einige Beispiele für die Abkopplung eines oder mehrerer dieser Bereiche:

Beispiel 1: Die Lehrkräfte werden für ein neues oder überarbeitetes Curriculum nicht ausreichend geschult und sind somit höchstwahrscheinlich nicht in der Lage, das Curriculum angemessen umzusetzen. Die Änderung des Curriculums verfehlt folglich seine beabsichtigte positive Wirkung.

Beispiel 2: Die Prüfung passt nicht zu den Anforderungen des Curriculums, beispielsweise bezieht sich das Curriculum auf die vier Sprachfertigkeiten, überprüft wird jedoch mit einem Grammatik- und Lesetest in Form von Multiple-Choice-Fragen. Oder das Curriculum basiert auf einem bestimmten Bildungskontext, die Prüfung jedoch nicht.

Beispiel 3: Die Lehrwerke und Materialien, die im Unterricht verwendet werden, entsprechen nicht den Anforderungen des Curriculums (beispielsweise verfolgt das Curriculum einen kommunikativen, die Materialien jedoch einen grammatikbasierten Ansatz oder das Curriculum wurde für einen bestimmten Bildungskontext entwickelt, die Materialien jedoch nicht).

Beispiel 4: Das Curriculum, die Ausbildung der Lehrkräfte und die Leistungsmessung sind gut aufeinander abgestimmt, die baulichen Gegebenheiten im Klassenzimmer machen eine effektive Umsetzung allerdings schwierig oder gar unmöglich (beispielsweise verhindern fest verbaute Möbel Paar- oder Gruppenaktivitäten).

Wenn in einem Lernsystem also eine maßgebliche Änderung eingeführt wird, sollte aus den obigen Ausführungen deutlich geworden sein, dass alle Bereiche berücksichtigt werden müssen. Wenn also **ein neues oder überarbeitetes Curriculum eingeführt wird, müssen Lehrkräfte angemessen geschult werden, damit sie das Curriculum erwartungsgemäß umsetzen können**. Auch die Tests müssen aktualisiert oder neu geschrieben werden, damit sie den revidierten Bildungsanforderungen entsprechen.

Bevor eine Verlinkung zwischen einem Curriculum, einem Teil der Testauspielung/-bereitstellung oder einem Test vorgenommen wird, kann es zwar hilfreich (und angemessen) sein zu argumentieren, dass das, was im Fokus steht, in einer Weise mit dem GeR verbunden ist die Ihre Verlinkungsbehauptung stützt. Das bedeutet jedoch nicht, dass es für die Verwendung in einem bestimmten Lernkontext geeignet ist. Beispielsweise funktioniert ein Kursbuch für das B1-Niveau, das zur Verwendung in einem regionalen spanischen Schulsystem entwickelt wurde, nicht zwangsläufig auch in einem entsprechenden System in Malaysia oder Indien. Es reicht nicht zu argumentieren, dass beide auf das GeR-Niveau B1 ausgerichtet sind und das Kursbuch deshalb für beide Kontexte gleichermaßen angemessen ist.

1.5 Was umfasst die Anbindung an den GeR?

Die Anbindung an den GeR umfasst einen von zwei Prozessen:



Prozess 1 und Prozess 2 umfassen jeweils **eine Reihe etablierter und größtenteils aufeinanderfolgender Schritte** oder Abläufe (siehe Abbildung 1.3).



Der erste Schritt im Anbindungsprozess ist die **FAMILIARISIERUNG** (weitere Details siehe Kapitel 2). Dabei handelt es sich um eine entscheidende Phase zu Beginn jeder Verlinkung. Erfahrungen aus früheren Untersuchungen zur Verlinkung haben gezeigt, dass sie in reduziertem Umfang auch eine sehr hilfreiche und manchmal auch notwendige Vorbereitung auf die anderen Phasen (z. B. Standardisierung) darstellt. Sie kann zum Beispiel eine nützliche Gelegenheit zur Auffrischung oder zur Neuausrichtung auf bestimmte Bereiche bieten, die für den Kontext des Verlinkungsprojekts relevant sind.

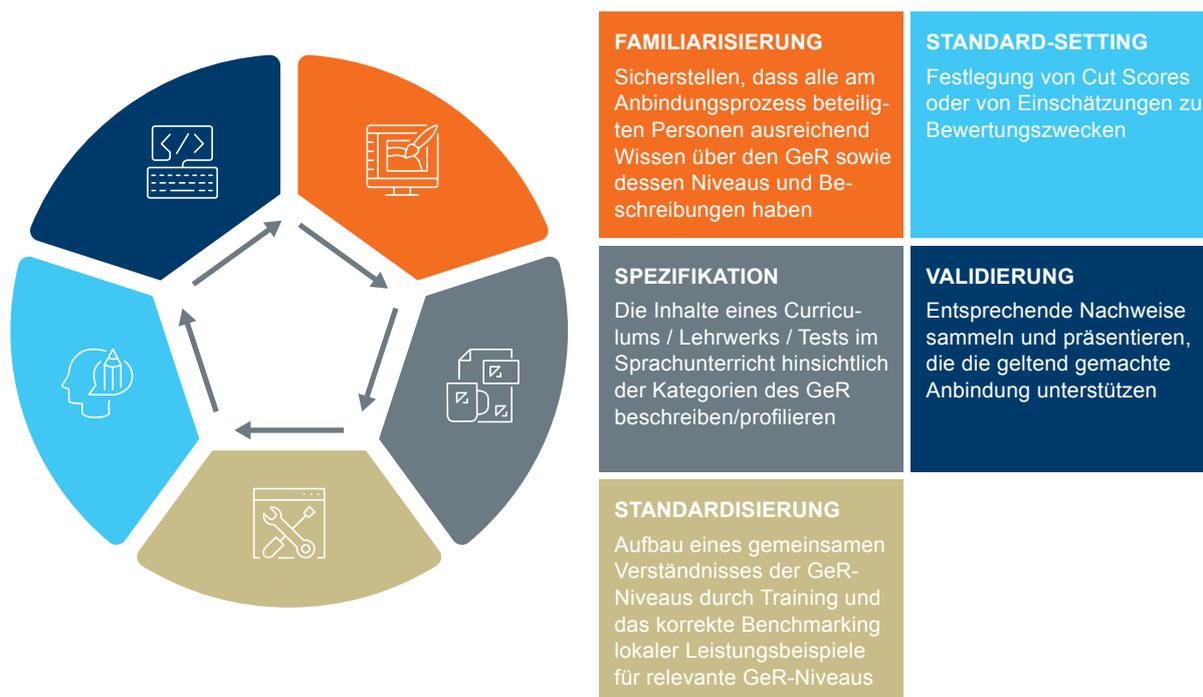


Abbildung 1.3: Schritte im Anbindungsprozess



Die nächsten Schritte sind die **SPEZIFIKATION** und **STANDARDISIERUNG** (Kapitel 3 und 4), die bei Projekten zur Anbindung an den GeR in den meisten oder sogar allen Kontexten erforderlich sind.



Für Mitwirkenden, die an allen der Entwicklungsaspekten einer empirisch fundierten Verbindung zwischen einem Curriculum, Lehrmaterialien (z. B. Lehrwerk oder Onlinekurs) oder einer Leistungsmessung bzw. einem Test beteiligt sind, stellen die Verfahren zum **STANDARD-SETTING** (Kapitel 5) eine Grundvoraussetzung dar. Bei einigen dieser Kontexte sind wahrscheinlich weniger Maßnahmen zum Standard-Setting erforderlich als bei einem Test. Die spezifische Nutzung eines Tests bestimmt die Auswahl und Angemessenheit der Verfahren zur Standardsetzung.



VALIDIERUNG (Kapitel 6) ist die kontinuierliche Qualitätskontrolle zur Sammlung von Nachweisen für die Gültigkeit der GeR-Verlinkungsbehauptungen. Wie die Familiarisierung ist auch die Validierung in gewisser Weise für alle anderen Schritte im Verlinkungsprozess relevant: Durch das korrekte Absolvieren sämtlicher Phasen erbringt man den Validitätsnachweis für die nachfolgenden Behauptungen über die Anbindung an den GeR.

1.6 Welche Publikationen stehen für die Anbindung an den GeR zur Verfügung?

Der [GeR 2001](#) und der [Begleitband 2020](#) stellen jeweils eine *einzelne* Publikation dar, die bei einer Anbindung an den GeR als Referenz dient. Hier ist eine kurze Zusammenfassung dazu, was in dem jeweiligen Dokument enthalten ist:

GeR 2001

- Die Richtlinien des Europarats zum Sprachunterricht (Kapitel 1)
- Die Begründung zum handlungsorientierten Ansatz (Kapitel 2)
- Die gemeinsamen Referenzniveaus (Kapitel 3)
- Kommunikative Sprachaktivitäten ohne die Berücksichtigung der illustrativen Skalen (Kapitel 4)
- Die Kompetenzen der Sprachverwendenden/Lernenden ohne die Berücksichtigung der illustrativen Skalen (Kapitel 5)
- Fremdsprachenlernen und -lehren (Kapitel 6)
- Aufgaben und ihre Rolle beim Fremdsprachenlernen (Kapitel 7)
- Sprachenvielfalt und das Curriculum (Kapitel 8)
- Beurteilen und Bewerten (Kapitel 9)



Begleitband 2020

- Schlüsselaspekte des Lehrens und Lernens im GeR
- Aktualisierte Skalen und Beispieldeskriptoren des GeR einschließlich
 - Ergänzungen der Skalen von 2001
 - Neue Skalen für Rezeption, Produktion und Interaktion
 - Neue Skalen für Phonologie
 - Skalen für Mediation
 - Skalen für Gebärdenkompetenzen
 - Skalen für plurilinguale und plurikulturelle Kompetenzen
- Kommunikative Sprachaktivitäten und Strategien (ersetzt Abschnitt 4.4 in Kapitel 4 des GeR 2001 und liefert signifikantes neues Material)
- Kommunikative Sprachkompetenzen (ersetzt Abschnitt 5.2 in Kapitel 5 des GeR 2001)

Die [Datenbank der Deskriptoren zum GeR-Begleitband](#) mit Suchfunktion kann während des Verlinkungsprozesses hilfreich sein. Für alle an der Erstellung oder Produktion von Sprachtests, Prüfungen und anderen Formen der Leistungsmessung beteiligt sind, die sich am GeR orientieren, ist das [Handbuch zur Anbindung von Sprachprüfungen an den GeR](#) (in englischer Sprache, Europarat 2009) und das [Handbuch zur Entwicklung und Durchführung von Sprachtests](#) (in englischer Sprache Europarat 2011, in deutscher Sprache telc 2012) relevant.

Unterstützende Publikationen mit Fallstudien, die den Prozess und die Erfahrungen der Anbindung von Tests beschreiben, sind beispielsweise von [Figueras und Noijons \(Hrsg.\), 2008](#) oder Martyniuk (Hrsg.), 2010.

1.7 Welche Orientierungshilfen bietet das Handbuch?

Der Rest dieses Handbuchs bietet **praktische und zugängliche Tipps**, um die Verlinkung an den GeR in verschiedenen spezifischen Kontexten zu unterstützen. In jedem der folgenden Kapitel geht es um die fünf oben beschriebenen Prozesse:

- Kapitel 2 – Familiarisierung
- Kapitel 3 – Spezifikation
- Kapitel 4 – Standardisierung
- Kapitel 5 – Standard-Setting
- Kapitel 6 – Validierung

Jedes Kapitel beginnt mit einer **Erläuterung des Prozesses und dessen Wichtigkeit**, gefolgt von **allgemeinen Tipps und praktischen Übungen, die für alle Nutzerinnen und Nutzer relevant sind**, unabhängig vom Kontext des Sprachunterrichts (Bildungsplan, Lehre, Prüfung).

Das Ende jedes Kapitels liefert **gezielte Anleitungen und praktische Übungen für verschiedene Kontexte und die jeweiligen Untergruppen**, z. B. für Sprachlehrkräfte, Erstellerinnen und Ersteller von Curricula, Verfasserinnen und Verfasser von Lehrwerken sowie für Entwicklerinnen und Entwickler von Prüfungen. Sie können diese nach Bedarf auswählen und berücksichtigen.

Uns ist bewusst, dass Sie Ihr Verlinkungsvorhaben möglicherweise innerhalb einer Gruppe durchführen, die eine oder mehrere Koordinierende und Teilnehmende umfasst. Aus diesem Grund bieten wir **praktische Tipps und Vorschläge zu diesen unterschiedlichen Rollen** einschließlich der zu erledigenden Aufgaben mit der voraussichtlichen Zeitplanung.

Im Anhang finden Sie **fotokopierfähige** Dokumentationsbogen, die verwendet und vervollständigt werden können. Sie stehen auch online im bearbeitbaren Format zur Verfügung. Dieses praktische Zusatztool hilft Ihnen bei der regelmäßigen Qualitätskontrolle und Validierung im Verlinkungsprozess. Es empfiehlt sich, die Formulare so zu verwenden, wie sie hier präsentiert werden. Es ist durchaus möglich, dass situationsbedingt Änderungen erforderlich sind, um die Anforderungen eines bestimmten Ansatzes oder von einem bestimmten Material im Rahmen des Verlinkungsprozesses zu erfüllen. Sämtliche Änderungen des in diesem Handbuch vorgeschlagenen Ansatzes sollten sorgfältig abgewogen und in Ihrem finalen Bericht begründet werden. Weitere Informationen finden Sie im Abschnitt **Hinweise zur eigenen Implementierung** in Kapitel 6.



Hinweise zur eigenen Implementierung ...

Sie können am besten beurteilen, was innerhalb Ihres bestimmten Kontextes möglich oder realistisch ist und was nicht.

Wir empfehlen Ihnen deshalb dringend, **die Maßnahmen auf Ihren bestimmten Kontext auszurichten** und dabei Ihre Ressourcen und (evtl. eingeschränkten) Möglichkeiten zu berücksichtigen. **Sie sollten nicht mehr tun**, als für Ihren Zweck **tatsächlich erforderlich ist**. Gleichzeitig sollten Sie aber auch **das, was durchgeführt werden muss, nicht unterschätzen**.

Wichtige Fragen, die Sie während des Verlinkungsprozesses stets im Auge behalten sollten:

1. **Warum machen wir das?**
2. **Welche Schritte sind für uns in unserem Kontext und für unseren Zweck entscheidend?**
3. **Welche Schritte haben Priorität?**
4. **Welche Schritte sind unter Umständen nicht notwendig?**
5. **Wie lange dauert der Prozess?**
6. **Wie hoch werden die Kosten sein?**
7. **Welches Fachwissen benötigen wir bzw. zu welchem haben wir Zugang?**
8. **Welche Ressourcen (z. B. Zeit, Finanzierung, Fachwissen) müssen wir einplanen?**



KAPITEL 2: FAMILIARISIERUNG

Sich mit einem Thema familiarisieren heißt, sich damit vertraut machen – ein wichtiger erster Schritt bei jedem Anbindungsprojekt. Kapitel 2 beschäftigt sich mit folgenden Punkten:

- 2.1 Was versteht man unter Familiarisierung?
- 2.2 Allgemeine Aktivitäten zur Familiarisierung
- 2.3 Spezifische Aktivitäten zur Familiarisierung
 - 2.3.1 Rollen und Entscheidungen von Koordinatorinnen und Koordinatoren
 - 2.3.2 Vorbereitung des Materials
- 2.4 Leitfaden für Berichte über die Familiarisierungsphase

Die Abschnitte 2.1, 2.2 und 2.4 richten sich an alle, die an einem Anbindungsprojekt mitwirken, unabhängig von ihrer Rolle oder ihrem Aufgabenbereich. Der Abschnitt 2.3 hingegen ist für Personen gedacht, die für die Organisation und Durchführung bestimmter Aktivitäten zur Familiarisierung zuständig sind.

Am Ende des Kapitels sollten Sie eine klare Vorstellung davon haben,

- wie wichtig die Familiarisierung ist und was sie umfasst,
- was durchgeführt werden muss, um den GeR-Ansatz, seine Skalen und Deskriptoren zu verstehen und
- wie die Durchführung der Aktivitäten zu Berichtszwecken dokumentiert wird.

2.1 Was versteht man unter Familiarisierung?

Die Familiarisierung soll sicherstellen, dass die Beteiligten in einem Verlinkungsvorhaben über angemessene Kenntnisse der GeR verfügen und ein gemeinsames Verständnis des Projektziels haben. In diesem Handbuch werden zwei Arten der Familiarisierung erläutert: die allgemeine und die spezifische Familiarisierung (siehe Abbildung 2.1). Die allgemeine Familiarisierung kann in einer Sitzung erfolgen und individuell erarbeitet werden, während die spezifische Familiarisierung in der Regel unter der Anleitung einer Koordinatorin bzw. eines Koordinators in der Gruppe stattfindet und mehrere Sitzungen erfordert.

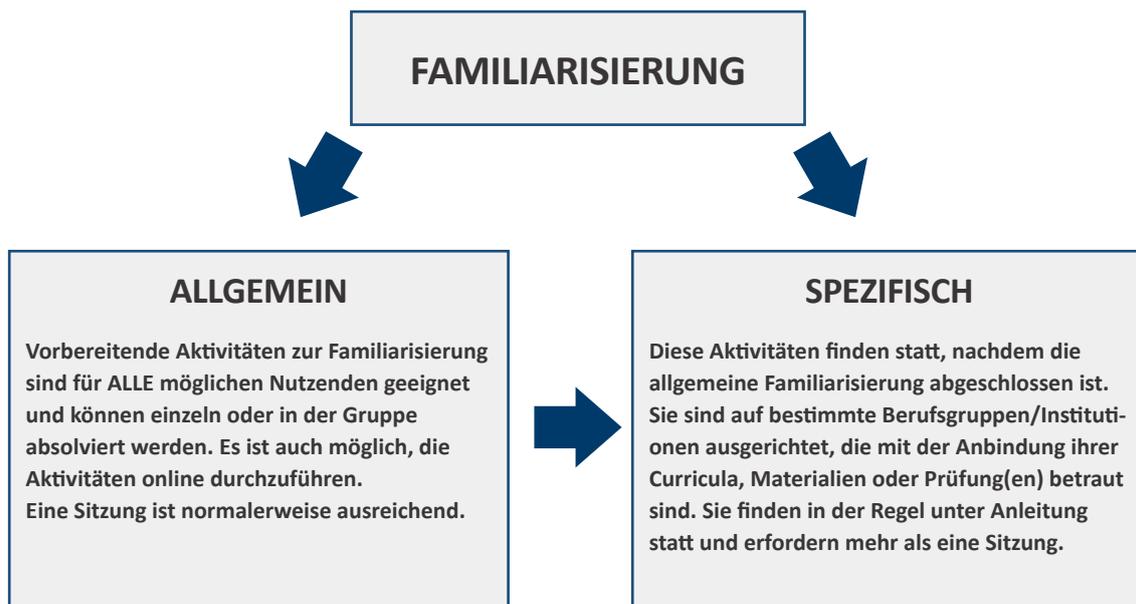


Abbildung 2.1: Allgemeine und spezifische Familiarisierung

Wie lange die Aktivitäten in der jeweiligen Veranstaltung zur Familiarisierung dauern, hängt zum einen davon ab, wie gut die Teilnehmenden mit dem GeR (2001) und dem GeR-Begleitband (2020) vertraut sind und zum anderen vom Ziel und Umfang des jeweiligen Verlinkungsprojekts. Wir empfehlen, einen kompletten Vor- oder Nachmittag für jede Veranstaltung einzuplanen.

Einzelne Beteiligte möchten sich vielleicht aus persönlichem Interesse oder aufgrund ihrer kontextspezifischen Bedürfnisse die Links zu weiterführenden Informationen im GeR-Begleitband ansehen. Bei der Erstellung von Curricula könnte beispielsweise der Leitfaden [From Linguistic Diversity to Plurilingual Education: Guide for the Development of Language Education Policies in Europe](#) von Beacco und Byram (2007) interessant sein, für Prüfende [Relating Examinations to the CEFR. A Manual](#) von North et al. (2009) oder das [Manual for Language Test Development and Examining](#) der ALTE (2011).

2.2 Allgemeine Aktivitäten zur Familiarisierung

Die elektronische oder gedruckte Version des GeR-Begleitbands 2020 unterstützt dabei, die in Abbildung 2.2 gezeigten Aktivitäten in der vorgegebenen Reihenfolge durchzuführen. Die Bearbeitung kann einzeln oder in der Gruppe, in Präsenz oder online erfolgen. Notieren Sie Ihre Antworten auf jede Frage. Das wird Ihnen beim Ausfüllen des Dokumentationsbogens 2.1 in Aktivität 5 helfen.



Abbildung 2.2: Fünf allgemeine Aktivitäten zur Familiarisierung

- 1. Wichtige Aspekte:** In den Kapiteln 1 und 2 des GeR-Begleitbands 2020 erfahren Sie mehr über die Ergänzungen und Änderungen des GeR 2001. Stellen Sie sich die folgenden Fragen:
 - Was bedeutet eine Anbindung an den GeR?
 - Welche Auswirkungen hat das auf mich und meinen Kontext?
- 2. Beschreibungssystem:** Konzentrieren Sie sich auf Abschnitt 2.4 des GeR-Begleitbands (*Das Beschreibungssystem des GeR*) und stellen Sie sich die folgenden Fragen:
 - Welche Aktivitäten, Kompetenzen und Strategien sind für meinen Kontext am relevantesten?
 - Sind sie alle gleich relevant? Wenn nicht, wie kann ich sie priorisieren?
- 3. Zentrale Merkmale:** Lesen Sie Anhang 1 des GeR-Begleitbands (*Zentrale Merkmale der GeR-Niveaus*), in dem die Entwicklung von Niveaustufe zu Niveaustufe beschrieben wird. Stellen Sie sich folgende Fragen:
 - Mit welcher Niveaustufe bzw. welchen Niveaustufen werde ich hauptsächlich arbeiten?
 - Was sind die zentralen Merkmale dieser Niveaustufe bzw. der angrenzenden Stufen?

Sehen Sie sich Anhang 2 an (*Raster zur Selbstbeurteilung*) und definieren Sie die wichtigsten kommunikativen Sprachaktivitäten (Rezeption, Produktion, Interaktion, Mediation). Stellen Sie sich folgende Frage:

 - Welche Aspekte sind in meinem Kontext am relevantesten?
- 4. Relevante Skalen:** Sehen Sie sich die Skalen mit den verschiedenen illustrativen Deskriptoren an und ermitteln Sie, welche Skalen für Ihren Kontext am relevantesten sind:
 - Kommunikative Sprachaktivitäten und -strategien:
 - Rezeption (Abbildung 11, S. 57)
 - Produktion (Abbildung 12, S. 74)
 - Interaktion (Abbildung 13, S. 87)
 - Mediation (Abbildung 14, S. 112)
 - Plurilinguale und plurikulturelle Kompetenz (Abbildung 15, S. 144)
 - Kommunikative Sprachkompetenzen (Abbildung 16, S. 152)
 - Gebärdenkompetenzen (Abbildung 17, S. 169).

Der Einführungstext vor jeder Skala kann hilfreich sein. Hinweis: Unter Umständen müssen Sie sich einige der Skalen im Detail ansehen, um entscheiden zu können, wie relevant sie für Ihre Zwecke sind. Notieren Sie sich die Titel der Skalen zur späteren Verwendung. Achten Sie darauf, dass Sie bei der Vielzahl der verfügbaren Skalen *nur diejenigen* auswählen, die für Ihre Arbeit am wichtigsten sind. Stellen Sie sich die folgenden Fragen:

 - Welche Skalen sind in meinem Kontext am relevantesten? Warum?
 - Wie begründe ich meine Entscheidungen? (Die Antwort auf diese Frage ist maßgebend, da sie die von Ihnen aufgestellten Behauptungen untermauert.)
- 5. Zusammenfassung:** Füllen Sie abschließend den Dokumentationsbogen 2.1 im Anhang aus.

2.3 Spezifische Aktivitäten zur Familiarisierung

Spezifische Aktivitäten zur Familiarisierung sind auf die bestimmten Bedürfnisse von Expertengruppen abgestimmt, die bereits die allgemeinen Aktivitäten zur Familiarisierung absolviert haben. Die spezifische Familiarisierung sollte nach Möglichkeit als Gruppenaktivität durchgeführt werden, da Diskussionen dabei helfen, ein gemeinsames Verständnis dafür zu entwickeln, wie die verschiedenen Deskriptoren interpretiert werden können. Das gilt insbesondere, wenn diese nicht vollständig transparent sind. Beim Deskriptor *Analyse und Kritik kreativer Texte (einschließlich Literatur)* für das B1-Niveau finden wir beispielsweise die folgende Aussage zum Kenntnisstand: „Kann die Schlüsselthemen und die zentralen literarischen Figuren kurzer Erzählungen beschreiben, die vertraute Situationen beinhalten und in gängiger Alltagssprache verfasst sind.“ Gruppendiskussionen zu bestimmten Wörtern oder Formulierungen wie „kurz“, „vertraut“ oder „gängige Alltagssprache“ können den Teilnehmenden dabei helfen, sich möglichst präzise darüber zu verständigen, wie diese Aussage auf B1-Niveau interpretiert werden kann.

Auf Anfrage der Koordinatorin oder des Koordinators sollen die Teilnehmenden ihre ausgefüllten Dokumentationsbogen 2.1 zur Verfügung stellen.

Umfasst ein Projekt alle Modi der Kommunikation (Rezeption, Produktion, Interaktion, Mediation), sind für die Aktivitäten mindestens zwei Sitzungen erforderlich. Der konkrete Zeitbedarf hängt davon ab, wie sehr die Teilnehmenden mit dem GeR vertraut sind und welche allgemeinen Aktivitäten zur Familiarisierung als relevant erachtet wurden.

Die Unterabschnitte 2.3.1 und 2.3.2 richten sich an die koordinierende Person. Noch bevor die Aktivitäten vorbereitet werden, kann es hilfreich sein, sich Kapitel 3 aus dem [Manual for Relating Examinations to the CEFR](#) durchzulesen, insbesondere dann, wenn die Teilnehmenden eine High-Stakes-Prüfung an den GeR anpassen wollen.

Auch die [Datenbank der Deskriptoren zum GeR-Begleitbands mit Suchfunktion](#) kann für Koordinierende relevant sein.

2.3.1 Rollen und Entscheidungen von Koordinatorinnen und Koordinatoren

Koordinierende spielen eine entscheidende Rolle bei der Organisation spezifischer Aktivitäten zur Familiarisierung. Sie sollten über umfassende Kenntnisse des GeR und der Tragweite eines Anbindungsprozesses verfügen. Außerdem sollten sie mit dem Kontext und den Anforderungen der organisierenden Gruppe/Institution vertraut sein. Nur auf Grundlage dieser Kenntnisse sind sie in der Lage,

- die Logistik und Inhalte der Veranstaltung(en) festzulegen,
- das Material für die aufeinanderfolgenden Aktivitäten vorzubereiten und
- einen Bericht über den Familiarisierungsprozess zu verfassen.

Die Zeit, die die Koordinierenden in die Vorbereitung, Bereitstellung und Dokumentation jeder Veranstaltung zur Familiarisierung investieren müssen, hängt von verschiedenen Faktoren ab: von der Erfahrung mit der Durchführung von Veranstaltungen zur Familiarisierung zum GeR 2001, von den geforderten Aufgaben (Art des erforderlichen Berichts; Menge an Informationen, die zusammengestellt werden müssen; Anzahl und Art der Veranstaltungen usw.), vom Umfang der (personellen) Unterstützung zur Durchführung (Vorbereitung des Materials, Organisation der Veranstaltungsorte und des Caterings, Verwaltung der Onlineplattform usw.). Auch die Anzahl der Teilnehmenden sowie ihr Vorwissen hat Auswirkungen darauf, wie viel Zeit die Planung in Anspruch nimmt. Als Faustregel gilt: Die Vorbereitungsphase der spezifischen Familiarisierung dauert mindestens doppelt so lang wie die Veranstaltungen mit den Teilnehmenden selbst.

Vor der Vorbereitung der spezifischen Aktivitäten zur Familiarisierung sollten Koordinierende Folgendes tun:

1. Entscheiden, ob Aktivitäten als Präsenzveranstaltung, online oder als Kombination stattfinden sollen. Außerdem ist sicherzustellen, dass die Anzahl der angemeldeten Teilnehmenden eine adäquate Diskussion und eine erfolgreiche Durchführung der Aktivitäten ermöglicht. Präsenzveranstaltungen sollten auf 18 bis maximal 20 Teilnehmende begrenzt werden. Bei Onlinesitzungen sollten die verfügbaren Ressourcen (Software, IT-Support usw.) überprüft werden, um sowohl Diskussion als auch Durchführung der Aktivitäten zu ermöglichen.
2. Die Antworten der Teilnehmenden in den Dokumentationsbögen 2.1 zur allgemeinen Familiarisierung lesen und evaluieren. Die Bögen wurden ggf. von den Teilnehmenden individuell bearbeitet. So kann entschieden werden,
 - a) auf welchen Kommunikationsmodi/-fertigkeiten, Skalen, Niveaus und Deskriptoren der Fokus liegen sollte (mit Begründung),
 - b) wie viele Sitzungen erforderlich sind,
 - c) wie viel Wiederholung nötig ist, bevor die Arbeit zu den Kommunikationsmodi – Rezeption, Interaktionen, Produktion und Mediation – begonnen werden kann (dies hängt auch davon ab, wie lange die allgemeine Familiarisierung zurückliegt),
 - d) wie der Fokus, die Struktur und die Inhalte jeder Sitzung aussehen sollen,
 - e) wie viele Aktivitäten welcher Art, Diskussionszeiten, Pausen usw. benötigt werden,
 - f) welche Materialien, die für den entsprechenden Kontext relevant sind, verwendet werden sollen (Prüfungsaufgabe, Aktivitäten zur Bewertung im Kurs, Ausschnitte aus einem Curriculum, Stundenplan usw.) und wie sie den Teilnehmenden zur Verfügung gestellt werden.

Falls das Projekt alle vier Modi der Kommunikation umfasst, ist es empfehlenswert, sich in Sitzung 1 auf die Rezeption, in Sitzung 2 auf die Interaktion und Produktion und in Sitzung 3 auf die Mediation – mit der die Teilnehmenden unter Umständen weniger vertraut sind – zu konzentrieren. Wenn für die Mediation eine ganze Sitzung eingeplant wird, können Mediationsstrategien und -aktivitäten behandelt und vielleicht auch Aktivitäten zur Familiarisierung mit den Skalen zur plurilingualen und plurikulturellen Kompetenz durchgeführt werden.

Die Koordinierenden müssen entscheiden, wann sie die Skalen für kommunikative Sprachkompetenzen behandeln (wir empfehlen Sitzung 2) und ob die Teilnehmenden eine Familiarisierung zu den Skalen der Gebärdenkompetenz benötigen oder nicht.

3. Auf Grundlage der mit dem Projekt verknüpften Aufgaben und des erhaltenen Auftrags entscheiden, welche Berichtsart und welche Nachweise erforderlich sind.

2.3.2 Vorbereitung des Materials

Die Vorbereitung des Materials und die Veranstaltungsorganisation sollte an die angestrebte Durchführungsmethode (in Präsenz vor Ort oder online) sowie an die Zeitplanung (synchron oder asynchron) angepasst werden.

Spezifische Familiarisierung – Sitzung 1

Sitzung 1 besteht aus zwei Teilen. Teil 1 ist ein Überblick über die Aktivitäten zur allgemeinen Familiarisierung und sollte wie unten dargestellt durchgeführt werden. Teil 2 widmet sich der Rezeption, Produktion und Interaktion oder Mediation, wobei es empfehlenswert ist, mit der Rezeption zu beginnen. Die Aktivitäten in Teil 1 und Teil 2 werden in Abbildung 2.3 grafisch dargestellt.

Bei jeder Aktivität und vor einer Diskussion sollten sich die Teilnehmenden verschiedene Fragen stellen. Ihre Antworten sind beim Ausfüllen des Dokumentationsbogens 2.2 am Ende der Sitzung hilfreich.

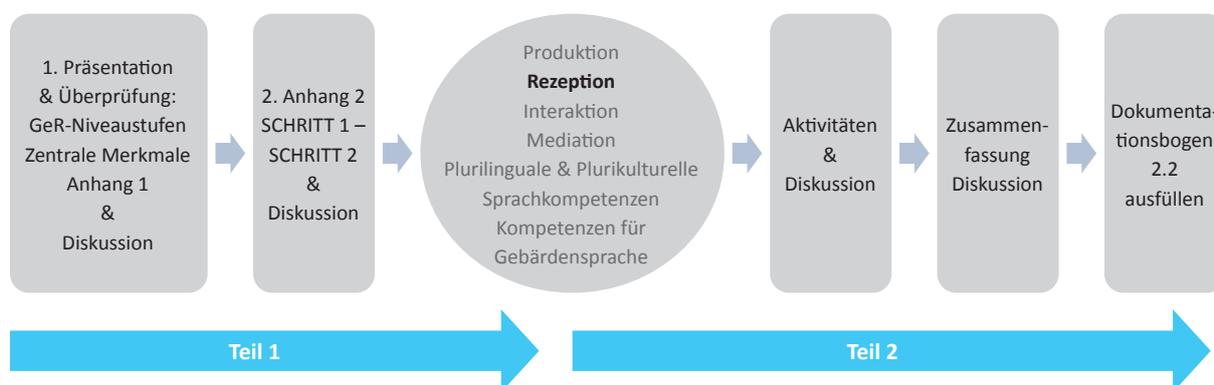


Abbildung 2.3: Aktivitäten für die erste Sitzung zur Familiarisierung

1. Präsentation und Übersicht

Bereiten Sie das Material (PowerPoint-Präsentation, Quiz usw.) vor, um die Ziele der spezifischen Familiarisierung sowie die Struktur und Inhalte der Sitzung(en) darzulegen.

Weisen Sie die Teilnehmenden nach der Präsentation auf Anhang 1 des GeR-Begleitbands (stellen Sie bei Bedarf Kopien bereit) hin und bitten Sie sie, das Dokument durchzulesen und jeweils die prägnantesten Merkmale der GeR-Niveaus zu unterstreichen/hervorzuheben. Bitten Sie die Teilnehmenden, den ausgefüllten Dokumentationsbogen 1 zu prüfen und ihre Antwort(en) auf die folgende Frage zu notieren:

- Welche Eigenschaften der GeR-Niveaus sind in Ihrem Kontext am relevantesten?

Die Teilnehmenden diskutieren ihre Antworten paarweise oder in Kleingruppen.

2. Anhang 2 des GeR-Begleitbands: Raster zur Selbstevaluation

Bereiten Sie das Raster zur Selbstevaluation für eine Aktivität in zwei Schritten vor.

Schritt 1: Da der Fokus in diesem ersten Schritt auf Rezeption, Interaktion und Produktion liegt, ist es erforderlich, die Seiten 3 bis 5 des Anhangs 2 des GeR-Begleitbands zu kopieren (zu finden unter [676999_GER_Begleitband_ANHANG_02_Raster zur Selbstbeurteilung \(1\).pdf](#)) und

- a) einige Felder im Raster auszublenden bzw. zu entfernen (die für das Projekt besonders relevant erachtet werden),
- b) Schnipsel aus entfernten Deskriptoren vorzubereiten, die die Teilnehmenden dann wieder einsetzen können,
- c) das vollständige Raster für Rezeption, Interaktion und Produktion zur Kontrolle und Besprechung zu kopieren.

Bitten Sie die Teilnehmenden, ihre individuellen Antworten auf die folgenden Fragen zu notieren:

- Welche Stichworte (Handlungen, Informationsarten, Bedingungen und Einschränkungen) beschreiben die folgenden Aspekte am besten?
 - Hörverständnis
 - Leseverständnis
 - Mündliche Produktion und mündliche Interaktion
 - Schriftliche Produktion und schriftliche / Online-Interaktion

Bitten Sie die Teilnehmenden, paarweise oder in Kleingruppen die leeren Felder mit den ausgeschnittenen Schnipseln der Deskriptoren, die Sie in Schritt 1 vorbereitet haben, zu füllen sowie die Hauptfunktionen von Rezeption, Interaktion und Produktion aller Niveaus zu identifizieren.

Schritt 2: Da der Fokus in diesem Schritt auf Mediation liegt, drucken bzw. kopieren Sie die Seiten 6 und 7 von Anhang 2 des GeR-Begleitbands und

- a) schneiden Sie die Deskriptoren für die drei Arten von Mediation aus: Mediation eines Texts, eines Konzepts und von Kommunikation (alle Niveaus);
- b) legen Sie sie vermischt in einen Umschlag;
- c) bereiten Sie für jede Art von Mediation ein Blanko-Raster vor, damit die Teilnehmenden auswählen können, zu welcher Skala und zu welchem Niveau die ausgeschnittenen Deskriptoren jeweils gehören;
- d) kopieren Sie nach Abschluss der Aufgabe die vervollständigten Raster, um sie zu korrigieren und zu besprechen.

Bitten Sie die Teilnehmenden – paarweise oder in Kleingruppen –, die Deskriptoren der jeweiligen Mediationsart und den Niveaus zuzuordnen. Verteilen Sie zur Korrektur und Besprechung das vollständige Raster. Lassen Sie die Teilnehmenden in Einzelarbeit die Hauptfunktionen auf allen Niveaustufen ermitteln. Bitten Sie sie, ihre Antworten auf die folgenden Fragen zu notieren, bevor Sie mit der Diskussion beginnen:

- Welche Stichwörter (Handlungen, Informationsarten, Bedingungen und Einschränkungen) beschreiben die folgenden Aspekte am besten?
 - Mediation von Texten
 - Mediation von Konzepten
 - Mediation von Kommunikation

3. Aktivitäten mit Deskriptoren und Skalen und Diskussion

Bereiten Sie Aktivitäten vor, die zum jeweiligen Kontext und Niveau passen. Als Grundlage dafür dient der Dokumentationsbogen 2.1. In der verbleibenden Zeit von Sitzung 1 kann nur ein Modus von Kommunikation besprochen werden. Falls das Projekt alle Kommunikationsmodi betrifft, empfehlen wir, sich in dieser ersten Sitzung auf die Rezeption zu konzentrieren.

Als Aktivitäten eignen sich die qualitative Analyse relevanter Deskriptoren, das Zuordnen ausgewählter, durcheinandergewürfelter Deskriptoren zu den entsprechenden Niveaustufen, das Ranking von ausgewählten Deskriptoren und das Platzieren auf einem Raster oder das Zuweisen verschiedener Deskriptoren zu den entsprechenden Skalen und Niveaus usw. (weitere Details zur Vorbereitung dieser Aktivitäten siehe Kapitel 3 des [„Manual for Relating Examinations to the CEFR“](#)).

Nachdem sie alle Aktivitäten abgeschlossen haben und bevor sie in die Gruppendiskussion starten, bitten Sie die Teilnehmenden, die folgenden Fragen zu beantworten:

- Warum ist diese Skala in meinem Kontext relevant?
- Welche Stichwörter beschreiben meine(n) Zweck(e) am besten?

An dieser Stelle sollten alle Teilnehmenden den Dokumentationsbogen 2.2 ausfüllen.

4. Eine Abschlussveranstaltung vorbereiten

Bereiten Sie **eine Abschlussveranstaltung** vor, damit die Teilnehmenden einen Überblick darüber erhalten, was erarbeitet und erreicht wurde.

Wenn Sie im finalen Bericht Kennziffern zur Übereinstimmung³ unter den Teilnehmenden hinzufügen möchten, sollten Sie sicherstellen, dass die Aktivitäten so präsentiert werden, dass Sie die Bewertungen/Rankings der Teilnehmenden zusammenführen und analysieren können.

5. Dokumentationsbögen einsammeln

Sammeln Sie die Dokumentationsbögen 2.2 für Sitzung 1 ein.

6. Zusammenfassung der Antworten schreiben

Schreiben Sie eine Zusammenfassung der in den Dokumentationsbögen 2.2 festgehaltenen Antworten.

³ Das sind Angaben zum Grad der Übereinstimmungen unter den Teilnehmenden – sie können quantitativ (in der [Referenzergänzung zum Manual](#) beschrieben) oder qualitativ (Vergleich von Kommentaren/Beobachtungen) sein. Letzteres ist besonders relevant für die Anbindung von Curricula oder Material.

Material für weitere Sitzungen

Ob Sie eine oder zwei weitere Sitzungen zur Familiarisierung vorzubereiten haben, ergibt sich daraus, wie Sie in Sitzung 1 bezüglich der Kommunikationsmodi, auf denen der Fokus liegen sollte, entschieden haben und wie vertraut die Teilnehmenden mit dem GeR sind. Folgen Sie dabei den oben beschriebenen Schritten 3, 4, 5 und 6.

Auch weitere Sitzungen müssen analog zu Sitzung 1 Folgendes umfassen (siehe Abbildung 2.4):

- Eine einführende Aktivität, in der die Arbeit aus den vorherigen Sitzungen rekapituliert wird, um zurück und auf die weitere Vorgehensweise zu blicken
- Aktivitäten zu den ausgewählten Skalen und Diskussion im Anschluss (in der Regel mit einer Pause dazwischen)
- Eine zusammenfassende Sitzung und abschließende Diskussion, die von der koordinierenden Person angeleitet wird
- Ausfüllen des zugehörigen Dokumentationsbogens durch die Teilnehmenden



Abbildung 2.4: Aktivitäten für weitere Sitzungen zur spezifischen Familiarisierung

2.4 Leitfaden für Berichte über die Familiarisierungsphase

Um sämtliche Nachweise zusammenzutragen, die während der Familiarisierungsphase erarbeitet wurden, sollte die koordinierende Person die Aktivitäten und Antworten der Teilnehmenden mithilfe des Bogens 2.3 zusammenfassen. Dies bildet die Grundlage für spätere Berichte – sowohl zu dieser Sitzung als auch für den finalen Validierungsbericht, in dem die finalen Annahmen hinsichtlich der Anbindung an den GeR aufgestellt und belegt werden.

Berichte zu jedem Verlinkungsschritt erfüllen zweierlei Zweck. Erstens sind sie eine Art institutionelles Gedächtnis und somit als interne Aufzeichnungen notwendig, und zweitens werden die im Bericht erfassten Nachweise Teil des finalen Berichts, der den Verlinkungsprozess belegt (Kapitel 6). Berichte können sowohl zur ausschließlich internen Verwendung als auch für ein externes Publikum erstellt werden.

Die Hauptautorinnen und -autoren des Berichts werden zu Koordinatorinnen und Koordinatoren ernannt, die sicherstellen, dass alle Aktivitäten durchgeführt wurden, und die die Ergebnisse dokumentieren. Sie leiten den Bericht an die beauftragende Institution weiter, die sich wiederum um dessen Weitergabe kümmert.

2.4.1 Berichte zur internen Verwendung

Für die interne Berichterstellung sollten administrative Prozesse, die für die Institution spezifisch sind, in den Unterlagen dokumentiert werden:

- Berücksichtigte Protokolle und erforderliche Genehmigungen (Auswahl der koordinierenden Person(en), Auswahl der Teilnehmenden usw.)
- Ressourcen und notwendige Vorbereitung (Budgets, Technologie, Material, Unterstützung usw.)
- Logistik (Veranstaltungsort, Catering usw.)
- Gewonnene Erkenntnisse und Empfehlungen

Der Bericht sollte auch Kopien der ausgefüllten Familiarisierungsbögen sowie ein kurzes zusammenfassendes Dokument enthalten, das die Koordinierenden unterzeichnen und folgende Angaben enthält:

- Anzahl der Teilnehmenden,
- die in den Abschnitten behandelten Aktivitäten/Inhalte sowie Zeitpläne und Kalender,
- Grad der Zufriedenheit der Teilnehmenden und
- eine allgemeine Evaluierung der Koordinierenden.

2.4.2 Berichte zur externen Verwendung

Die Art eines Berichts zur externen Verwendung variiert je nach Projekt und Zielgruppe, die entweder breiter gefasst (z. B. Schulen, Verlage, Nutzerinnen und Nutzer oder Testteilnehmende) oder spezieller (z. B. Lehrkräfte oder Prüfteams) sein kann.

Richtet sich ein Bericht an ein breites Publikum, sollten die im Bericht zur Verfügung gestellten Informationen zur internen Verwendung ausreichen (ausgenommen nicht relevante Informationen, die speziell die Institution betreffen).

Sollte ein Bericht für ein spezifisches Publikum gedacht sein, sind weitere Details erforderlich, beispielsweise Informationen zu relevanten Eckdaten der Teilnehmenden oder deren Leistung(en) in den verschiedenen Aktivitäten (z. B. Hinweise zur Übereinstimmung beim Ranking von Deskriptoren oder bei der Beurteilung der Zuordnungsgenauigkeit von Deskriptoren zu Skalen oder Niveaus).



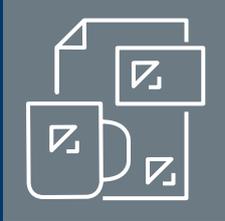
Hinweise zur eigenen Implementierung ...

Wenn Sie eigenständig arbeiten:

- Das Erarbeiten der Familiarisierung ist eine intensive Phase. Regelmäßige Pausen tragen zum Erfolg bei.
- Vermeiden Sie zu ambitioniertes Verhalten und zu hoch gesteckte Ziele. Sorgfalt ist besser als Schnelligkeit.
- Nehmen Sie sich ausreichend Zeit für jede Aktivität.
- Dokumentieren Sie regelmäßig Ihre Entscheidungen. Das erleichtert das Ausfüllen der Dokumentationsbögen.

Koordinatorinnen und Koordinatoren sollten zusätzlich Folgendes beachten:

- Planen Sie ausreichend Zeit für Diskussionen ein.
- Hetzen Sie die Teilnehmenden nicht.
- Variieren Sie das Format der Aktivitäten innerhalb von Sitzungen und von Sitzung zu Sitzung.
- Achten Sie darauf, dass Sie alle relevanten Informationen zusammentragen, die Sie für den finalen Bericht benötigen, unter anderem die Dokumentationsbögen am Ende jeder Sitzung.



KAPITEL 3: SPEZIFIKATION

Bei der Spezifikation werden die Inhalte bestehender oder neuer Materialien hinsichtlich ihres Ansatzes und der Abdeckung der im GeR dargestellten Kategorien analysiert. Der Prozess der Spezifikation dient sowohl dazu, die Anbindung an den GeR festzustellen als auch dafür zu sensibilisieren, was zu einer weiteren Optimierung führen kann. Bei der Spezifikation soll die folgende Frage beantwortet werden:

„Inwieweit decken meine Materialien die Inhalte des GeR ab?“

Wie detailliert die Spezifikation ausgeführt wird, hängt von den Anprüchen ab, die Sie an die Verlinkung stellen wollen.

Der Spezifikationsprozess erfordert eine Familiarisierung mit dem GeR (wie in Kapitel 2 beschrieben) sowie ein fundiertes Verständnis des angestrebten Sprachniveaus der Zielsprache, des Kontextes sowie der Art des Materials (z. B. Lehrplan oder Lehrwerk), die an den GeR angepasst werden soll. Die für die Spezifikation zuständige(n) Person(en) können die relevantesten GeR-Deskriptorskalen und Deskriptoren auswählen und sie mit den Zielen und Inhalten des Materials in Beziehung setzen, um eine Anbindung an den GeR zu erreichen.

Dieses Kapitel besteht aus folgenden Unterpunkten:

- 3.1 Was versteht man unter Spezifikation?
- 3.2 Allgemeine Aktivitäten zur Spezifikation
- 3.3 Spezifische Aktivitäten zur Spezifikation
- 3.4 Leitfaden für Berichte über die Spezifikationsphase

Am Ende des Kapitels sollten Sie die folgenden Fragen präzise beantworten können:

- Was umfasst die Spezifikation?
- Welcher Ansatz ist im jeweiligen Kontext am relevantesten/hilfreichsten?
- Wie analysieren und beschreiben Sie Ihre eigene Materialien?
- Wie erstellen Sie ein GeR-Profil Ihres Materials?
- Wie dokumentieren Sie den Grad der Anbindung an den GeR?

3.1 Was versteht man unter Spezifikation?

Spezifikation ist der Analyseprozess der Materialien und deren Zuordnung zum GeR. Wie detailliert spezifiziert wird, hängt von den individuellen Anforderungen ab. Bei einem nationalen Curriculum wäre beispielsweise ein hohes Maß an Details erforderlich, während auf institutioneller Ebene eine allgemeinere Zuordnung ausreichen könnte. Die Spezifikation ist der erste Schritt bei der Zusammenstellung und des Vorlegens von Nachweisen der Verlinkung mit dem GeR.

Für die Spezifikation ist es erforderlich, mindestens zwei der allgemeinen Familiarisierungsaktivitäten in Kapitel 2 durchzuführen. Wir empfehlen auch, diese noch einmal anzusehen und explizit auf das Material, das an den GeR angepasst werden soll, abzustimmen. Für die Spezifikation ist eine sorgfältige und detaillierte Einarbeitung in die GeR-Deskriptoren notwendig.

In diesem Handbuch werden zwei Arten von Spezifikationsprozessen vorgeschlagen: **allgemein und spezifisch**. Der allgemeine Prozess umfasst eine weitreichende, generelle Verlinkung von Struktur und Inhalten der Materialien (Curriculum, Material oder Prüfungen) mit den Inhalten und Niveau-Deskriptoren des GeR. Verschiedene allgemeine Aktivitäten zur Spezifikation sind in Abschnitt 3.2 zu finden. Beim spezifischen Prozess werden dieselben Prinzipien angewendet, allerdings viel detaillierter und nuancierter, wie in Abschnitt 3.3 beschrieben.

Jeder Prozess ist auf eine oder zwei Richtungen ausgelegt: Top-down oder Bottom-up. Beim Top-down-Prozess dient der GeR als Basis für die Erstellung komplett neuer Bildungspläne, Materialien oder Prüfungen, während der Bottom-up-Prozess mit der Analyse bestehender Materialien und Praktiken beginnt, die an den GeR angepasst werden müssen. In beiden Fällen muss der Verwendungszweck des Materials sowohl deutlich gemacht werden als auch klar im Kontext verankert sein und auf den Anforderungen der Lernenden basieren. Zudem wird in beiden Fällen ein tiefgreifendes Verständnis des GeR und des GeR-Begleitbands (siehe Abschnitt 2.1 oben) vorausgesetzt.



Abbildung 3.1: Spezifikationsprozess

Beim allgemeinen Prozess werden die zentralen Merkmale jeder Niveaustufe (Anhang 1 im GeR-Begleitband) und das Raster zur Selbstbeurteilung (Anhang 2 im GeR-Begleitband) von allen Nutzerinnen und Nutzern zur Kontextualisierung und Anbindung verwendet. Dieses Vorgehen entspricht Abschnitt 2.1 zur allgemeinen Familiarisierung. Beim spezifischen Prozess der Spezifikation wiederum sollten die illustrativen Deskriptorskalen für die bestimmten kommunikativen Sprachaktivitäten, -strategien und -kompetenzen, die für den jeweiligen Kontext und Zweck relevant sind, ermittelt und ausgewählt werden. Detaillierte Schritte für jeden Prozess werden in den folgenden Abschnitten vorgeschlagen.

3.2 Allgemeine Aktivitäten zur Spezifikation

Die in diesem Abschnitt vorgestellten Aktivitäten können den Grundstein für Ihren Nachweis der Anbindung Ihres Materials an den GeR legen. Mit dem Durchlaufen dieser Schritte erhalten Sie eine allgemeine Beschreibung Ihres Materials bezüglich der GeR-Kategorien und Sprachniveaus.

Die fünf allgemeinen Aktivitäten zur Spezifikation (Abbildung 3.2) können von einer Person einzeln oder von einer Gruppe gemeinsam erarbeitet werden. Währenddessen sollten die Antworten auf die Fragen zu jeder Aktivität notiert werden, da sie beim Ausfüllen der zusammenfassenden Dokumentationsbögen 3.1 bis 3.3 Aktivität 5 hilfreich sein können.



Abbildung 3.2: Schritte im allgemeinen Spezifikationsprozess

Aktivität 1: Relevantes Sprachniveau und zentrale Merkmale

Rufen Sie Anhang 1 des GeR-Begleitbands auf, der für die Aktivität 3 im Familiarisierungsprozess in Kapitel 2 bereits verwendet wurde.

- a) Sehen Sie sich die Tabelle mit der Zusammenfassung der gemeinsamen Referenzniveaus genau an. Lesen Sie sich die Deskriptoren für jedes der sechs Sprachniveaus durch. Unterstreichen Sie die Verben im Zusammenhang mit den verschiedenen Kommunikationsaktivitäten: Die Verben „verstehen“ und „zusammenfassen“ beziehen sich auf die rezeptiven Fähigkeiten (mündlich und schriftlich), „ausdrücken“ und „äußern“ auf die produktiven Fähigkeiten (mündlich und schriftlich); „verständigen“ und „austauschen“ beziehen sich auf die interaktiven Fähigkeiten.

Stellen Sie sich die folgenden Fragen:

- 1) Welche(s) Sprachniveau(s) ist/sind für den Kontext und Zweck meines Vorhabens am relevantesten?
- 2) Decken verschiedene Aktivitäten meines Vorhabens mehr als eine Niveaustufe ab oder beabsichtigen dies?

- b) Lesen Sie sich den Text zu zentralen Merkmalen für jede Niveaustufe sorgfältig durch und ermitteln Sie Wörter und Formulierungen, die jede Niveaustufe definieren und ein Differenzierungskriterium darstellen.

Stellen Sie sich die folgenden Fragen:

- 1) Hat die Lektüre des Textes zu den zentralen Merkmalen meine Antworten auf die Fragen in Abschnitt 1 oben bestätigt?
- 2) Kann ich entscheidende Unterschiede zwischen der Niveaustufe, die ich als für mein Vorhaben und die zugehörigen Niveaustufen am relevantesten definiert habe, erklären?

Aktivität 2: Sprachaktivitäten und -strategien

Beim GeR werden die beschreibenden Schemata gemäß den vier Arten von Kommunikation organisiert. Diese werden in Abbildung 3.3 zusammen mit den zugehörigen Sprachaktivitäten und -strategien gezeigt. Da es keine allgemeinen Skalen für Strategien gibt, wurden die spezifischen Strategieskalen für jede Art der Kommunikation einbezogen.

Der Dokumentationsbogen 3.4 sollte von der Entwicklerin bzw. dem Entwickler des Materials vor der Datenerhebung ausgefüllt werden. Dies stellt eine Übersicht der beabsichtigten Phasenergebnisse dar und kann später mit den tatsächlichen Ergebnissen verglichen werden, um die Verlinkungsanforderungen zu validieren, die nach den Aktivitäten definiert wurden.

Stellen Sie sich die folgenden Fragen:

- 1) Welche Art(en) der Kommunikation, Sprachaktivitäten und Sprachstrategien sind für mein Material am relevantesten?

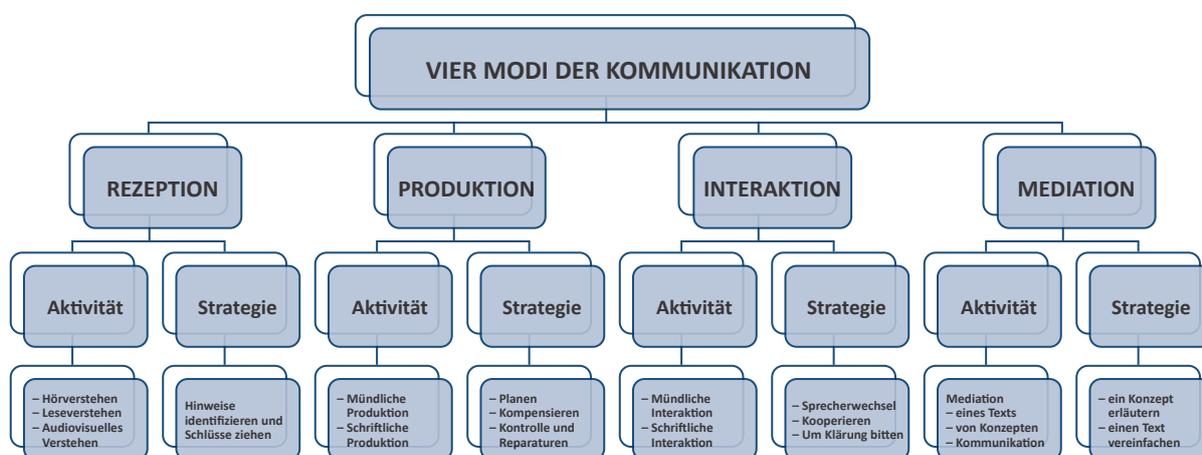


Abbildung 3.3: Vier Modi der Kommunikation

Aktivität 3: Kompetenzen

Sehen Sie sich die folgenden Abbildungen im GeR-Begleitband an:

- Abbildung 15: *Plurilinguale und plurikulturelle Kompetenzen* (S. 144)
- Abbildung 16: *Kommunikative Sprachkompetenzen* (S. 152)
- Abbildung 17: *Gebärdenkompetenzen* (S. 168).

Stellen Sie sich zu jeder Abbildung die folgende Frage:

- 1) Auf welche Kompetenz(en) ist mein Vorhaben ausgerichtet, sofern zutreffend?
(In den meisten Fällen ist es unwahrscheinlich, dass das Vorhaben auf Kompetenzen in allen drei Bereichen (s. Abbildungen 15-17) abzielt.)

Aktivität 4: Grafische Darstellung der Vereinheitlichung

Rufen Sie das Raster zur Selbstbeurteilung in Anhang 2 des GeR-Begleitbands auf.

Stellen Sie sich die folgenden Fragen:

- 1) Welche Deskriptoren sind für mein Vorhaben bezüglich Sprachniveau und kommunikativen Sprachaktivitäten am relevantesten?
- 2) Wie kann ich ein grafisches Profil erstellen, das die Beziehung zwischen meinen Materialien und den Deskriptoren der GeR-Niveaus im Raster zur Selbstbeurteilung abbildet?

Aktivität 5: Zusammenfassung

Füllen Sie basierend auf den Antworten auf die Fragen zu den Aktivitäten 1 bis 4 die Dokumentationsbögen 3.1, 3.2 und 3.3 im Anhang aus.

3.3 Spezifische Aktivitäten zur Spezifikation

Bei der spezifischen Spezifikation werden die Inhalte des Materials Abschnitt für Abschnitt den bestimmten GeR-Deskriptoren zugeordnet. Nach diesem Prozess sollten Sie in der Lage sein zu zeigen, welche der illustrativen Skalen des GeR damit in Beziehung stehen, und für jede davon Folgendes ermitteln:

- Auf welches Sprachniveau zielt sie ab? (Und welche(r) bestimmte(n) Deskriptor(en) gilt/gelten?)
- Wo wird im Material darauf eingegangen?

Der Prozess beginnt mit der Analyse des Materials. Wie Sie zu Ihrer Analyse gelangen, hängt von zwei Faktoren ab:

- Ob Sie neues Material erstellen (Top-down) oder bestehendes anpassen (Bottom-up).
- Von der Art des Materials: Handelt es sich um ein Curriculum, eine Kursplanung, ein Kursbuch, eine Materialsammlung oder eine Prüfung?

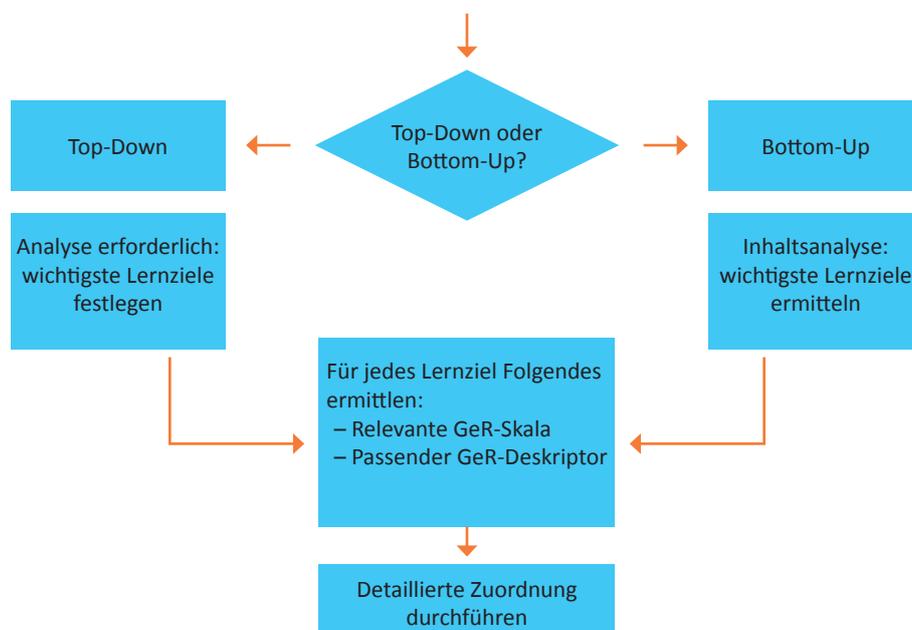


Abbildung 3.4: Übersicht über den spezifischen Spezifikationsprozess

Aktivität 1: Das eigene Material analysieren

Wenn Sie neues Material erstellen (Top-down), sollten Sie mit einer Bedarfsanalyse beginnen. Wozu sollten erfolgreiche Lernende (oder Prüfungsteilnehmende) nach dem Absolvieren des Programms (oder dem Bestehen der Prüfung) in der Lage sein? Wenn Sie die allgemeine Spezifikation (Abschnitt 3.2) durchgeführt haben, sollten Sie bereits eine allgemeine Antwort auf diese Frage haben. An dieser Stelle gehen Sie mehr ins Detail. Ziel ist es, eine Liste einzelner Lernziele zusammenzustellen.

Wenn Sie bestehendes Material verlinken, beginnen Sie mit einer Inhaltsanalyse. Das wurde beim Erstellen des Materials unter Umständen schon durchgeführt. Möglicherweise existiert eine Übersicht, ein Inhaltsverzeichnis oder eine Testspezifikation, die zeigt, welche einzelnen Lernziele es gibt (oder wie die Prüfung aufgebaut ist). Das kann die Analyse erleichtern. Sie sollten die angegebenen Ziele genau prüfen, um – so gut es geht – herauszufinden, was von den Lernenden explizit erwartet wird. Bei einem Curriculum oder einer Kursplanung sollten Sie auf Deskriptoren oder Beispiele von Aktivitäten achten, bei Lehr-/Lernmaterialien auf die Aktivitäten und bei einem Programm zur Leistungsmessung oder einer Prüfung auf die geplanten Aktivitäten und/oder die Testelemente und -aufgaben. Passen sie zu den angegebenen Zielen?

Unabhängig davon, ob Sie neues Material erstellen oder eine bestehende anpassen, sollten Sie entscheiden, wie detailliert Ihre Analyse sein soll. Ein Lehrwerk besteht beispielsweise aus mehreren Kapiteln, die jeweils mehrere Einheiten enthalten, und jede der Einheiten umfasst unter Umständen verschiedene Lernaktivitäten. Sollten Sie jedes Kapitel an einen GeR-Deskriptor anpassen, jede Unterrichtseinheit oder – auf unterster Ebene – jede einzelne Aktivität? Auf diese Frage gibt es keine einfache Antwort, aber generell lässt sich sagen, dass Sie auf klar definierbare Lernziele achten sollten, die angeben, was erfolgreiche Lernende *können müssen*. Diese Lernziele sollen letztlich den Kannbeschreibungen des GeR zugeordnet werden. Hier ein Beispiel: Bei einem Lehrwerk könnte eine bestimmte Einheit eine zehnminütige Aufgabe im Kurs sein, in deren Rahmen zehn Vokabeln vermittelt werden sollen. Dieses Lernziel kann nicht leicht einem GeR-Deskriptor zugeordnet werden, aber als *Gerüst* für die Hauptaktivität in der Einheit dienen, die aus der Lektüre eines Zeitschriftenartikels besteht. Das ist wiederum ein Ziel, das einer GeR-Skala zugeordnet werden kann (z. B. „Leseverstehen: Information und Argumentation verstehen“). In diesem Fall ließe sich das Material am besten auf Ebene der gesamten Einheit an den GeR anbinden.

Das andere Problem ist die Anpassung oder Lokalisierung der GeR-Deskriptoren. Das geschieht, wenn bestehende Deskriptoren nicht richtig zum Kontext passen, für den das Material vorbereitet wird. Eine tiefgreifende Diskussion geht zwar über den Umfang dieses Handbuchs hinaus, Nutzerinnen und Nutzer sollten aber wissen, dass die GeR-Deskriptoren nicht in Stein gemeißelt sind. Kannbeschreibungen können:

- mit den aktuellen Formulierungen übernommen werden,
- gestrichen werden, wenn sie nicht zum Kontext passen,
- an einen bestimmten Kontext oder eine bestimmte Zielgruppe angepasst oder
- komplett umgeschrieben werden (für einen bestimmten Kontext oder eine bestimmte Zielgruppe).

In den letzten beiden Fällen kann die Durchführung einer Studie zur Validierung des alternativen Wordings oder des Fokus erforderlich sein. Dies würde sich auf das Urteil von Expertinnen und Experten stützen, ähnlich wie die Methodik, die zur Erstellung des GeR selbst verwendet wurde.

Aktivität 2: Die zugehörigen GeR-Skalen ermitteln

Sehen Sie sich nun die Liste der Lernziele aus Aktivität 1 an. Stellen Sie sich bei jedem Ziel die folgende Frage:

- Welche illustrative Skala ist am zutreffendsten?

Je nach Art des Ziels kann es mehr als eine Antwort auf diese Frage geben. Wenn das Ziel beispielsweise darin besteht zu prüfen, ob Prüfungsteilnehmende eine kurze, informative Präsentation planen und durchführen können, dann könnten Sie „Vor Publikum sprechen“ (unter *Produktionsaktivitäten*) als relevanteste Skala auswählen. Sie könnten aber auch entscheiden, dass „Themenentwicklung“ (unter *Pragmatische Kompetenzen*) relevant ist. Wenn Sie sich die tatsächlichen Deskriptoren in diesen Skalen ansehen, erkennen Sie vielleicht, dass eine besser zum Lernziel passt als die andere. Andersherum kann es aber auch sein, dass Sie zu dem Ergebnis gelangen, dass beide in gleichem Maße zutreffen. In diesem Fall könnten Sie im nächsten Schritt beide hinzufügen.

Aktivität 3: Das GeR-Niveau ermitteln

Sehen Sie sich innerhalb der identifizierten Skalen den Deskriptor an, der am besten zum Ziel passt. Beginnen Sie mit dem Deskriptor für das erwartete Ziel oder Zielniveau (das Niveau, das Sie für diese Kategorie von Sprachaktivität, -strategie oder -kompetenz im Kurs zur Skizzierung Ihrer allgemeinen Spezifikation ermittelt haben).

Wenn ein Deskriptor mehrere Kannbeschreibungen umfasst (was oft der Fall ist), welche ist dann am relevantesten? Sehen Sie sich dann die Deskriptoren für die umliegenden Niveaus an. So können Sie prüfen, ob das aktuelle Ziel tatsächlich über oder unter dem erwarteten Ziel oder Zielniveau liegt. Identifizieren Sie bei jedem Deskriptor die Formulierungen, die sie von den angrenzenden Niveaus unterscheiden. Wenn das Lernziel beispielsweise das Schreiben kurzer zusammenhängender Texte über die eigene Person mit dem Zielniveau A2 ist, sollten Sie sich die Deskriptoren für A1, A2 und A2+ in der Skala für „Kreatives Schreiben“ ansehen. Identifizieren Sie die relevantesten Kannbeschreibungen in diesen Deskriptoren und legen Sie dabei den Fokus vor allem auf die Formulierungen zu den Themen, über die die Lernenden schreiben können sollten (siehe Abbildung 3.5), sowie die Art des Textes, der produziert werden sollte (in Rot angezeigt). In diesem Beispiel zeigt sich, dass der relevanteste Deskriptor der für A2+ ist (da das Lernziel ein *zusammenhängender* Text ist).

A2+	Kann in Form verbundener Sätze etwas über alltägliche Aspekte des eigenen Umfelds schreiben, wie z. B. über Menschen, Orte, einen Job oder Studienerfahrung .
A2	Kann in einer Reihe einfacher Sätze über die eigene Familie, die Lebensumstände, den Bildungshintergrund oder die momentane oder vorige berufliche Tätigkeit schreiben.
A1	Kann einfache Wendungen und Sätze über sich selbst und fiktive Personen verfassen, wo sie leben und was sie tun .

Abbildung 3.5: Ausgewählte Kannbeschreibungen der Skala „Kreatives Schreiben“ - Niveau A1, A2 und A2+

Aktivität 4: Detaillierte Zuordnung

Notieren Sie für jeden passenden Deskriptor, den Sie ermitteln, das entsprechende Lernziel sowie den Verweis (ein Hinweis darauf, wo es in Ihrem Material vorkommt, z. B. eine Seiten-, Einheits- oder Abschnittsnummer), die GeR-Skala, das GeR-Niveau und die Kannbeschreibung selbst. Verwenden Sie dafür den Dokumentationsbogen 3.5 im Anhang.

Neben der Dokumentation einer detaillierten Zuordnung der Anbindung an den GeR lassen sich mit dem ausgefüllten Dokumentationsbogen auch bestimmte allgemeine Aspekte auf einen Blick erkennen oder bestätigen. Wenn Sie sich beispielsweise die Spalte für das „GeR-Niveau“ ansehen, können Sie feststellen, inwieweit die detaillierte Zuordnung das Niveauprofil bestätigt, das Sie im Dokumentationsbogen 3.3 erarbeitet haben, oder ob die verschiedenen Komponenten einer Leistungsmessung in Bezug auf das Zielniveau ausgewogen sind. Beachten Sie, dass der Dokumentationsbogen 3.5 auch Platz für die Domänen, Themen und Textarten hat, die mit jedem Lernziel verbunden sind. Wenn Sie diese Spalten ausgefüllt haben, können Sie besser beurteilen, inwiefern die detaillierte Zuordnung mit der im Dokumentationsbogen 3.4 erstellten Übersicht übereinstimmt.

3.4 Leitfaden für Berichte über die Spezifikationsphase

Die Richtlinien für die Erstellung von Berichten zum Familiarisierungsprozess (Kapitel 2) gelten auch für die Spezifikation, sofern die folgenden zusätzlichen Details einbezogen werden:

- Die Fachkenntnis der für die Analyse des Materials zuständigen Expertinnen und Experten
- Ob die Ergebnisse der in den Dokumentationsbogen eingetragenen Aktivitäten noch einmal überprüft wurden, von wem und wie.

Die Dokumentationsbogen können auch dabei unterstützen, die durchgeführten Aktivitäten sowie deren Ergebnisse zu protokollieren. Ausgefüllte Dokumentationsbögen können verwendet werden um:

- Diskussionen und Reflexionen zu den möglichen Anforderungen und Herausforderungen der verschiedenen Phasen des Verlinkungsprojekts anzuregen
- zu illustrieren, welche Aktivitäten zur Spezifikation durchgeführt wurden und wie und von wem dies durchgeführt wurde (dient als Validitätsnachweis für den Verlinkungsprozess)
- als Teil des Validitätsnachweises für die Anbindung an den GeR zu dienen, die man nachweisen möchte



Hinweise zur eigenen Implementierung ...

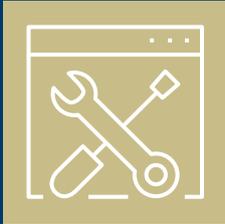
Wenn Sie eigenständig arbeiten:

Die Durchführung der Spezifikationsaktivitäten ist sehr zeitintensiv und Sie müssen entscheiden, ob Ihr Verlinkungsprojekt nur einen allgemeinen Ansatz erfordert (und die allgemeine Abdeckung priorisiert wird) oder ob auch spezifische Aktivitäten durchgeführt werden sollen (und der Fokus auf einer allgemeinen Abdeckung und detaillierten Anbindung liegt).

- Vermeiden Sie zu ambitioniertes Verhalten und zu hoch gesteckte Ziele.
- Achten Sie darauf, dass Sie ausreichend Ressourcen haben, um den ausgewählten Spezifikationsansatz zu verfolgen (Zeit, Personen, die sich mit der an den GeR anzupassendem Material auskennen und bei einer Überprüfung der Analyse usw. unterstützen können).
- Erleichtern Sie das Ausfüllen der Dokumentationsbögen, indem Sie Entscheidungen regelmäßig dokumentieren.

Koordinatorinnen und Koordinatoren sollten zusätzlich Folgendes beachten:

- Organisieren Sie das Team und definieren Sie klar, wer sich auf was konzentrieren soll.
- Planen Sie ausreichend Zeit für Diskussionen ein.
- Hetzen Sie nicht bei der Analyse.
- Sorgen Sie dafür, dass zusätzliche Überprüfungen stattfinden.
- Achten Sie darauf, dass Sie alle relevanten Informationen zusammentragen, die Sie für den finalen Bericht benötigen, unter anderem die Dokumentationsbögen am Ende jeder Sitzung.



KAPITEL 4: STANDARDISIERUNG

Bei der Standardisierung geht es darum, dafür zu sorgen, dass die Hauptmerkmale eines bestimmten Materials ein klares Verständnis der relevanten GeR-Niveaustufen und GeR-Deskriptoren widerspiegeln. Wenn zwei oder mehr Personen beteiligt sind, besteht das Ziel darin, einen Konsens darüber zu erzielen, was Lernende oder Prüfungsteilnehmende auf einem bestimmten Niveau können müssen und ob dies dem im Material angestrebten Niveau entspricht. Wie die Familiarisierung und Spezifikation variiert auch die Standardisierung je nach Material, Zweck der Anbindung und dem Anspruch, den Sie erheben wollen/müssen.

Was der Prozess der Standardisierung umfasst, wie er ausgeführt werden sollte und welche Instrumente bei der Anbindung von Sprachprüfungen oder Lernleistungen an den GeR verwendet werden müssen, wird in Kapitel 5 des *Manual for Relating Examinations to the CEFR (Europarat, 2009)* beschrieben und ist Voraussetzung für die Standardisierung. Das hier vorliegende Kapitel enthält einen Vorschlag für eine Adaption des Standardisierungsprozesses des *Manuals*, um den Anforderungen bei einem Curriculum oder Lehrwerk gerecht zu werden.

Dieses Kapitel gliedert sich in folgende Abschnitte:

- 4.1 Was versteht man unter Standardisierung?
- 4.2 Organisation des Standardisierungsprozesses
- 4.3 Schritte bei der Standardisierung
 - 4.3.1 Aktivitäten
 - 4.3.2 Tools (Hilfsmittel)
- 4.4 Leitfaden für Berichte über die Standardisierungsphase

Am Ende des Kapitels sollten Sie eine klare Vorstellung davon haben,

- was die Standardisierung umfasst und wie sie mit den anderen Aktivitäten im Handbuch zusammenhängt,
- welcher Ansatz für den Kontext und die Teilnehmenden am relevantesten/hilfreichsten ist,
- wie Sie eine Aktivität zur Standardisierung in Ihrem Kontext durchführen,
- wie Sie die Aktivität dokumentieren.

4.1 Was versteht man unter Standardisierung?

Durch die Standardisierung wird bestätigt, dass die wichtigsten Merkmale eines neuen oder bestehenden Materials – das kann ein Curriculum, ein Lehrwerk oder eine Prüfung sein – ein klares Verständnis des/der relevanten GeR-Niveaus widerspiegeln.

Die Teilnehmenden verwenden die GeR-Deskriptoren (oder lokale Versionen davon) sowie erklärende Beispiele (nach Möglichkeit bereits an den GeR angepasst), um zu einem gemeinsamen Verständnis zu gelangen, was Lernende oder Prüfungsteilnehmende auf den jeweiligen GeR-Niveaustufen können sollten und wie sich die Niveaustufe(n) für das betreffende Material operationalisieren lässt/lassen. Personen, die an einem Standardisierungsprozess beteiligt sind, sollten sich selbst darauf vorbereiten, indem sie die Familiarisierungs- und Spezifikationsphasen absolvieren.

Die Standardisierung ist Voraussetzung für das Standard Setting, sie kann aber auch zu einem anderen Zeitpunkt durchgeführt werden, beispielsweise vor dem Erstellen von Lern- oder Prüfungsmaterialien. So gelangen diejenigen, die für das Verfassen und Editieren von Material und Items zuständig sind sowie Content-Managerinnen und -Manager zu einem gemeinsamen Verständnis der GeR-Niveaustufen, das ihr Material anzielt. Alternativ ist die Standardisierung auch im Rahmen einer nachträglichen Validierung oder Evaluierung möglich. In diesem Fall wird ein **Bottom-up-Verfahren** gewählt, da das Material bereits vorhanden ist und eine Anbindung an den GeR angestrebt wird. Wir beginnen also mit dem Material und vergleichen es mit bestehenden exemplarischen Materialien (z. B. Prüfungs- oder Lernaufgaben, die vom Europarat zur Verfügung gestellt werden).

Wenn ein Material gerade erst erstellt werden soll oder sich am Anfang des Entwicklungsprozesses befindet, ist die Situation eine andere. Hierbei empfehlen wir, die Standardisierung früh im Entwicklungsprozess durchzuführen. In diesem Fall (bei einem **Top-down-Ansatz**) muss mindestens eine Komponente des Materials vollständig entwickelt werden, bevor der Standardisierungsprozess ernsthaft beginnen kann.

Die Koordinatorin oder der Koordinator entscheidet, welcher Teil des Materials entwickelt werden soll (wie in den Bögen 3.4 und 3.5 beschrieben). Die Teilnehmenden entwickeln diesen Teil in Kleingruppen (zwei bis vier Personen). Anschließend tauschen sich die Gruppen über ihre Arbeit aus und gleichen die Ergebnisse mit den Dokumentationsbögen 3.4 und 3.5 ab, um die folgende Frage zu beantworten: **Spiegelt diese Komponente die GeR-Niveaustufe(n), auf die sie ausgerichtet ist, präzise wider?**

Bei den Ansätzen Bottom-up und Top-down besprechen die Teilnehmenden zuerst ihre Antworten auf die Frage in Kleingruppen. Anschließend gibt die Koordinatorin oder der Koordinator die Diskussion für die gesamte Gruppe frei und ermutigt die Teilnehmenden, sich daran zu beteiligen. Ziel ist es, einen Konsens darüber zu erreichen, welche der entwickelten Komponenten die GeR-Niveaustufe(n) am besten operationalisieren. Das Wissen, das aus der Entwicklung der Komponente und der Konsensbildung resultiert, dient als Basis für die weitere Arbeit an dieser und weiteren Materialien.

Der Standardisierungsprozess konzentriert sich normalerweise auf eine bestimmte GeR-Niveaustufe. Um ein Verständnis für die oberen und unteren Bereiche der Niveaustufe zu erhalten, ist es jedoch erforderlich, auch die umliegenden Niveaus zu berücksichtigen. Wenn der Kontext einen breiter gefassten Kompetenzbereich erfordert, kann sich die Standardisierung auf mehr als ein Niveau beziehen, was auch ein oder mehrere „Plus“-Niveaus umfassen kann (A2+, B1+, B2+). Es ist wichtig, den Profilansatz des GeR bezüglich der Sprachfertigkeit zu beachten, da verschiedene Sprachaktivitäten und Strategien manchmal miteinander verbunden sind. So kann sich beispielsweise eine hohe Sprechfertigkeit zusammen mit einem ausgezeichneten Hörverständnis entwickeln.

Wenn Sie mit einer seltener verwendeten Sprache arbeiten oder für einen anderen Zweck als Testen, steht Ihnen eine limitiertere Zahl an Quellen zur Verfügung. Die Standardisierung kann aber dennoch effektiv durchgeführt werden, sofern eine Anpassung an den entsprechenden Kontext erfolgt.

Die Standardisierung kann auch individuell durchgeführt werden. In diesem Fall gibt es allerdings weder eine Diskussion noch eine Konsensbildung. Falls die Arbeit eigenständig erfolgen soll, sollten im Rahmen der Standardisierung ein oder zwei Kolleginnen/Kollegen oder Expertinnen/Experten hinzugezogen werden, um den Prozess kritisch zu hinterfragen und die Rückschlüsse zu validieren. Wir raten jedoch dringend davon ab, dies ohne ein kleines Team zu tun, da die Aussagekraft der Nachweise, die für eine einzelne Person geltend gemacht werden können, wahrscheinlich äußerst gering ist. Für Low-Stakes-Unterrichtsmaterialien oder -Prüfungen mag dies zwar genügen, jedoch nicht für High-Stakes-Situationen, in denen die Lernenden benachteiligt werden könnten (z. B. eine Abschlussprüfung am Ende des Unterrichtsjahres, deren Ergebnis für das Erreichen der nächsten Stufe ausschlaggebend ist, oder die Vorbereitung auf eine wichtige Prüfung, in deren Rahmen die Materialien nicht auf das korrekte Niveau ausgerichtet sind). Aus diesem Grund gehen wir im weiteren Verlauf des Kapitels von einer größeren Teilnehmergruppe aus.

4.2 Organisation des Standardisierungsprozesses

Anleitende Informationen und Logistik spielen in allen Phasen des Verlinkungsprozesses eine entscheidende Rolle, wie in den Abschnitten 2.3.1 und 2.3.2 dieses Handbuchs im Kontext der Familiarisierung und in Kapitel 5 des *Manual for Relating Examinations to the CEFR* im Kontext der Standardisierung deutlich wurde.

Die für den Koordinierungsprozess zuständige Person hat folgende Aufgaben:

- die Logistik und Inhalte der Sitzung(en) festlegen,
- das Material für die aufeinanderfolgenden Aktivitäten vorbereiten,
- einen Bericht über den Prozess verfassen

Dabei schenkt sie den folgenden Punkten besondere Beachtung:

- die Anzahl der Teilnehmenden, deren Kenntnisstand und deren Verhältnis zu Ihrem Material,
- die Arbeit, die im Rahmen der Familiarisierung und Spezifikation bereits erledigt wurde,
- die Identifikation/Ortung von Beispielen, die dem Material ähneln und die nach Möglichkeit bereits an das/die GeR-Niveau(s) angebunden wurden, auf das/die Ihr Material ausgerichtet ist/sind,
- die Analyse der wichtigsten Merkmale der ermittelten Beispiele in Bezug auf die Hauptmerkmale des Materials,
- die Diskussionen, die während der Standard-Setting-Sitzungen stattfinden. Diese sind entscheidend, um ein gemeinsames Verständnis des/der GeR-Niveaustufe(n) in Ihrem Material zu schaffen,
- das Zusammentragen von Nachweisen, wenn die Standardisierungsaktivitäten ausgeführt wurden, und
- die Dokumentation der Ergebnisse der Diskussion(en).

4.3 Schritte bei der Standardisierung

Im *Manual for Relating Language Examinations to the CEFR* (Council of Europe 2009, S. 37) werden vier Schritte im Standardisierungsprozess aufgeführt:

1. GeR-Familiarisierungsaktivitäten ausführen.
2. Mit erläuternden Beispielen von Materialien, Aufgaben oder Leistungen arbeiten, die bereits an den GeR angepasst wurden, um ein adäquates Verständnis der GeR-Niveaus zu erlangen. Diese Aufgaben oder Leistungen sind entscheidend, da alle späteren Entscheidungen und Argumentationen davon abhängen, wie präzise sie das Niveau reflektieren.
3. Die Fähigkeit entwickeln, lokale/regionale Materialien, Aufgaben und Leistungen an die GeR-Niveaus anzupassen.
4. Sicherstellen, dass alle beteiligten Parteien dasselbe Verständnis haben und dass dies konsistent umgesetzt wird.

Vor Beginn des Standardisierungsprozesses gehen wir davon aus, dass die Familiarisierungs- und Spezifikationsphasen bereits abgeschlossen sind. In diesem Sinne schlagen wir vor, den in der Tabelle 4.1 beschriebenen Ansatz zu verfolgen. Diese Tabelle kann den Organisatorinnen und Organisatoren des Standardisierungsprozesses als Checkliste der Aktivitäten dienen.

Tabelle 4.1: Schritte im Standardisierungsprozess

Schritt	Aktivität	Anmerkungen
1	Wahl eines/einer Koordinierenden für die Aktivität (Stage lead =Leitung der Phase)	Zu den Aufgaben gehören: <ul style="list-style-type: none"> ▪ die Logistik und Inhalte der Sitzung(en) festlegen, ▪ das Material für die aufeinanderfolgenden Aktivitäten vorbereiten, ▪ einen Bericht über den Prozess verfassen.
2	Auswahl der Teilnehmenden	Dabei sollte besonders geachtet werden auf: <ul style="list-style-type: none"> ▪ die Art und das Niveau ihrer Expertise, ▪ ihre Erfahrung mit ähnlichen Veranstaltungen bzw. Aktivitäten, ▪ die gesamte Personenanzahl (bis zu 20 für einen High-Stakes-Test, 2 oder 3 für einen Low-Stakes-Test [z. B. formativ], 5–10 für die Beurteilung von Curricula oder Material).
3	Identifikation der erwarteten GeR-Niveaus	Entwickelte Curricula oder Materialien müssen ein bestimmtes GeR-Niveau oder mehrere Niveaus abbilden. Bei einer einzelnen Niveaustufe müssen die Standardisierungsmaterialien auf ein Niveau darüber und darunter ausgeweitet werden. Falls das Curriculum ohne Berücksichtigung des GeR entwickelt wurde, ist es ratsam, diesen Schritt in zwei Phasen zu unterteilen: <ul style="list-style-type: none"> ▪ Feststellung (der ungefähren Bandbreite der Niveaus) ▪ Bestätigung (das festgestellte Niveau bestätigen oder ablehnen – die ursprünglich beabsichtigte Aktivität)

4	Ermittlung von erläuternden GeR-Beispielen	<p>Einige davon sind verfügbar (siehe Abschnitt 4.3.2). Achten Sie jedoch auf Folgendes:</p> <ul style="list-style-type: none"> ▪ Ist der Bezug zum entsprechenden GeR-Niveau für diese Beispiele hinreichend nachgewiesen? ▪ Passt die Aktivität zum Kontext? (Hinweis: Falls unwahrscheinlich, sollte diese Aktivität hier nicht verwendet werden, da dies zu einem falschen Verständnis des Niveaus führen kann.) Diese Aufgabe kann von internen Mitarbeitenden übernommen werden. ▪ Wenn keine passenden Beispiele gefunden werden, sollten eigene entwickelt werden. Dazu kann eine Untergruppe der erfahrensten Teilnehmenden (intern und/oder extern) gebeten werden, sich die Bandbreite der Materialien anzusehen, um passende Beispiele zu finden. ▪ Sind die in den Beispielen operationalisierten Skalen relevant für den jeweiligen Kontext?
5	Planung des Events	<p>Das Standardisierungsevent sollte vorab sorgfältig geplant werden. Sie sollten sicherstellen,</p> <ul style="list-style-type: none"> ▪ dass alle Materialien vollständig vorbereitet sind, ▪ der Veranstaltungsort verfügbar und angemessen ist und ▪ die verfügbare Zeit ausreichend bemessen ist.
6	Durchführung der Aktivitäten	<p>Wie im nächsten Abschnitt zu sehen, liegt der Fokus auf Folgendem:</p> <ul style="list-style-type: none"> ▪ Erneute Familiarisierung mit dem GeR (kann je nach Teilnehmenden etwas Zeit erfordern) ▪ Die Ergebnisse der Spezifikationsphase noch einmal prüfen ▪ Eine Erläuterung oder Beschreibung der erwarteten Ergebnisse der Lernenden basierend auf den GeR-Deskriptoren konzipieren ▪ Erläuternde Beispiele prüfen <p>Die wichtigste Frage ist: <i>Spiegelt die Komponente das/die GeR-Niveau(s) präzise wider, auf das/die sie ausgerichtet ist/sind?</i></p> <p>Lautet die Antwort „Ja“ – wie geplant fortfahren</p> <p>Lautet die Antwort „Nein“ – entweder das Material entsprechend anpassen oder den Anbindungsprozess vorerst einstellen und die bis dato vorhandene, wenn auch eingeschränkte Argumentation akzeptieren</p>
7	Analyse der Ergebnisse	<p>Vor dem Ausfüllen der Formulare in diesem Abschnitt ist es wichtig, den Prozess noch einmal zu reflektieren und Folgendes in Betracht zu ziehen:</p> <ul style="list-style-type: none"> ▪ Bericht der Koordinatorin bzw. des Koordinators über die Aktivitäten ▪ Feedback der Teilnehmenden
8	Ausfüllen der Dokumentationsbogen	<p>Dies sollte zeitnah nach Abschluss der Aktivitäten geschehen (manche können aber auch zusammen mit den einzelnen Aktivitäten ausgefüllt werden). Diese Aufgabe kann, wenn nicht anders möglich, auch von der Koordinatorin oder vom Koordinator allein ausgeführt werden – idealerweise sollte sich jedoch ein kleines Team mit internen und externen Teilnehmenden darum kümmern.</p> <p>Nach Abschluss ist eine finale Besprechung mit allen Teilnehmenden hilfreich, um Feedback zu erhalten und die Validität des Berichts und der gesammelten Erkenntnisse zu erhöhen.</p>

4.3.1 Aktivitäten

Die Aktivitäten 1 bis 5 in diesem Abschnitt richten sich an Personen, die an einem Standardisierungsprozess beteiligt sind. Sie sollten in der dargestellten Reihenfolge durchgeführt werden. Jede im Material behandelte Kommunikationsart erfordert mindestens einen Tag Zeitaufwand.

Wir empfehlen den Teilnehmenden, ihre individuellen Antworten für jede Aktivität im Dokumentationsbogen zur Standardisierung 4.1 zu notieren. Es enthält eine Frage dazu, wie hilfreich die Standardisierungsaktivitäten sind und wie sie noch verbessert werden könnten. Außerdem ist es ratsam, die Gruppendiskussionen im Audio- oder Videoformat festzuhalten. Sowohl die Formulare der Teilnehmenden als auch die Aufzeichnungen sind für die Koordinatorin oder den Koordinator sehr hilfreich, wenn sie den Dokumentationsbogen 4.2 ausfüllen und den Standardisierungsbericht schreiben.

Aktivitäten 1 und 2: Familiarisierung und Spezifikation noch einmal ansehen

Alle am Standardisierungsprozess beteiligten Personen sollten ihr Wissen rund um den GeR auffrischen. Das kann über die bereits ausgefüllten Dokumentationsbogen geschehen (Anhang, Dokumentationsbogen 2.1–3). Vor allem dann, wenn eine größere Gruppe involviert ist, können zusätzliche Überprüfungen sinnvoll sein. So könnten der Gruppe beispielsweise zusätzliche Informationen bereitgestellt oder ein Quiz erarbeitet werden (siehe auch Abschnitt 5.6.1 des *Manual for Relating Examinations to the CEFR*).

Die ausgefüllten Dokumentationsbogen zur Spezifikation (Anhang, Dokumentationsbogen 3.1–5) können verwendet werden, um die Teilnehmenden daran zu erinnern, welche GeR-Skalen für das Material, an dem sie arbeiten, relevant sind. Diese Skalen können weitere Familiarisierungsaktivitäten bieten. Beispielsweise haben die Teilnehmenden die Möglichkeit, Deskriptoren den Niveaus zuzuordnen. Es kann auch hilfreich sein, sich darauf zu einigen, welche vier oder fünf Skalen am wichtigsten sind und dass bestimmte Skalen für bestimmte Abschnitte des Curriculums, für das Lehrwerk oder für die Prüfung relevanter sind. So haben die Teilnehmenden die Gelegenheit, sich auf die Bereiche der Sprache zu konzentrieren, die unmittelbar betroffen sind. In späteren Aktivitäten können sie sich dann speziell auf diese GeR-Skalen beziehen.

Die Teilnehmenden sollten noch einmal die Arbeit überprüfen, die sie in der Spezifikationsphase ausgeführt haben. Dazu könnte zum Beispiel eine gemeinsam vereinbarte Erläuterung zum Kontext erstellt werden, die die Zielgruppe und das Zielniveau enthält. Diese Beschreibung sollte Antworten auf die folgenden Fragen geben:

1. **Worum handelt es sich bei dem Material?**
2. **An wen richtet sich das Material? (An wen richtet sich das Material nicht?)**
3. **Auf welche(s) GeR-Niveau(s) ist das Material ausgerichtet?**
4. **Auf welche GeR-Skalen ist das Material ausgerichtet?**
5. **Wo und wann wird das Material verwendet?**
6. **Welche Merkmale des Materials geben exemplarische Antworten auf die Fragen 1 bis 5?**

Aktivität 3: Eine standardisierte Beschreibung der erwarteten Lernergebnisse entwerfen

Die Teilnehmenden sollten anhand der Erläuterung zum Kontext und den relevanten GeR-Skalen eine Beschreibung entwerfen, wozu die Lernenden der Zielgruppe in der Lage sein sollten. Dabei sollten sie Formulierungen aus den GeR-Deskriptoren verwenden und anpassen. Bei der Vorbereitung einiger Methoden zum Standard Setting können „Lernende der Zielgruppe“ durch „minimal kompetente Lernende“ oder „minimal kompetente Prüfungsteilnehmende“ ersetzt werden. Nach der individuellen Ausführung dieser Aufgabe besprechen die Teilnehmenden ihre Beschreibungen zuerst paarweise und dann in größeren Gruppen, bis die gesamte Gruppe eine Einigung über die Beschreibung der Standardisierung erzielt. Weitere Informationen dazu, wie eine Einigung erzielt und bestätigt werden kann, sind in den Abschnitten 5.4.1 und 5.7.2 des *Manual for Relating Examinations to the CEFR* zu finden.

Aktivität 4a: Erläuternde Beispiele analysieren

Ziel dieser Prozessphase ist es, die Merkmale der erläuternden Beispiele im Vergleich zu den Merkmalen des Materials zu analysieren und zu besprechen. So soll sichergestellt werden, dass die Materialien das/die Zielniveau(s) korrekt widerspiegeln und dass alle Sitzungsteilnehmenden ein gemeinsames Verständnis davon haben.

Vor der Analyse der erläuternden Beispiele können die Teilnehmenden gebeten werden, die folgenden Fragen einzeln zu beantworten und ihre Antworten und Kommentare für eine spätere Besprechung zu notieren:

- 1. Sind die erklärenden Beispiele für das/die Leistungsniveau(s) relevant, um das/die es mir geht?**
- 2. Sind die Beispiele relevant für Kontext und Zweck meines Materials?**
- 3. Sind die Beispiele relevant für die Kommunikationsmethode(n), Sprachaktivitäten und Sprachstrategien in meinem Material?**

Anschließend kann eine detailliertere Analyse der Dokumentationsbogen erfolgen, die im Rahmen der Spezifikationsaktivität bereits ausgefüllt wurden (Anhang, Dokumentationsbogen 3.1–5). Die Teilnehmenden ordnen die Beispiele den Inhalten der Dokumentationsbogen zu und notieren ihre Analysen, indem sie die Kästchen in den Dokumentationsbogen mit einem Häkchen versehen oder einen kurzen Kommentar verfassen. Das eignet sich am besten als Einzelaufgabe, bevor das Ergebnis paarweise und in größeren Gruppen besprochen wird, um auf einen gemeinsamen Nenner zu kommen. Es kann auch ein umfassenderer Ansatz gewählt werden, indem die Teilnehmenden eine benutzerdefinierte Version des Formulars C3 im *Manual for Relating Examinations to the CEFR* ausfüllen (siehe Dokumentationsbogen 4.3 im Anhang dieses Handbuchs).

Weitere Informationen zur Standardisierung der Leistungsbeispiele von Lernenden sind in Abschnitt 5.5.1 (Seite 42–44) des *Manual for Relating Examinations to the CEFR* für Sprechen bzw. in Abschnitt 5.5.2 (Seite 44–46) für Schreiben zu finden.

Falls keine erläuternden Beispiele vorhanden sind, bietet die Aktivität 4b eine alternative Möglichkeit, zu einem gemeinsamen Verständnis zu gelangen.

Aktivität 4b: Benchmarking des Materials

Falls erläuternde Beispiele fehlen, müssen die Teilnehmenden bei dieser Aktivität ihr Material analysieren und entscheiden, ob sie bezüglich der GeR-Skalen und -Deskriptoren, die sie widerspiegeln soll, korrekt operationalisiert wurde.

Bei einem bereits bestehenden Material (Bottom-up-Ansatz für die Standardisierung) wird eine kleine Gruppe Teilnehmender mit dem größten Know-how (idealerweise eine Mischung aus internen und externen Personen) gebeten, die in Aktivität 3 vereinbarte Standardisierungsbeschreibung mithilfe der Dokumentationsbogen 3.4 (Übersicht) und 3.5 (detailliertes Mapping) zu analysieren. Sie können diese Aufgaben einzeln durchführen und die folgende Frage auf Basis ihrer Analyse beantworten:

- 1. Spiegelt mein Material das/die GeR-Niveau(s) präzise wider, auf das/die sie ausgerichtet ist?**

Die anderen Teilnehmenden erhalten Zeit, um ihre Antworten paarweise oder in Kleingruppen zu besprechen, in der Regel geführt von den Personen aus der kleineren Gruppe, die sich mit der Beurteilung befasst haben. Anschließend eröffnet die Koordinatorin oder der Koordinator die Diskussion für die gesamte Gruppe. Die Teilnehmenden sollen sich dazu äußern, wie gut das/die angestrebte(n) GeR-Niveau(s) im Material operationalisiert wird/werden und welche Änderungen, wenn nötig, erforderlich sind.

Aktivität 5: Dokumentationsbogen ausfüllen und Standardisierungsbericht schreiben

Diese Aktivität wird selbstverständlich erst dann durchgeführt, nachdem alle vorherigen Aktivitäten abgeschlossen, dokumentiert und untersucht wurden. Sie wird normalerweise von der Koordinatorin bzw. dem Koordinator angeleitet und umfasst eine Gruppe von Co-Autorinnen und -Autoren, bei denen es sich meist um interne Teilnehmende einer Gruppe einer vorherigen Aktivität handelt. Die Durchführung sollte so bald wie möglich nach Aktivität 4 erfolgen. Es ist ratsam, zuerst die Dokumentationsbogen auszufüllen und dann eine relativ kurze Aktivität zur Reflexion des Prozesses durchzuführen. Dabei wäre es sinnvoll, alle Aspekte der Aktivitäten (Teilnahme, Diskussion, Ergebnisse) sowie die mögliche Stärke der Nachweise zu berücksichtigen. Wir empfehlen, die ausgefüllten Dokumentationsbogen und den Bericht von einer separaten Gruppe prüfen zu lassen – entweder einzeln oder im Rahmen eines Gruppenevents. So kann die Koordinatorin bzw. der Koordinator die Aussagen und die Entscheidung, mit der nächsten Phase fortzufahren, bestätigen (oder nicht).

4.3.2 Tools (Hilfsmittel)

Die folgenden Webseiten bieten Zugang zu erläuternden Beispielen, die an den GeR angepasst wurden. Dabei muss natürlich geprüft werden, ob die angebotenen Beispiele zum jeweiligen Kontext passen:

- [Beispielvideos](#) von Cambridge Assessment English:
- [YouTube-Videos](#) von Cambridge English (Sprechbeispiele auf Englisch)
- Mündliche Prüfung, die die sechs GeR-Niveaus veranschaulicht – Beispiele auf [Französisch](#), [Deutsch](#), [Englisch](#), [Spanisch](#) und [Italienisch](#):
- Aufgaben zum [Sprechen, Lesen, Hören und Schreiben](#) des Europarats
- [Aufgaben für das Schreiben des Europarats: Beispiele](#) – Beispiele auf Englisch, Französisch, Deutsch, Italienisch und Portugiesisch
- [English Profile](#) (English Vocabulary Profile, English Grammar Profile und Text Inspector) (um das GeR-Niveau Texte auf Englisch zu verstehen)
- Merlin [Platform](#)
- [Plan Curriculas Instituto Cervantes](#) (Beispiele für Spanisch)
- [Profile Deutsch](#) (Beispiele für Deutsch)
- [Profilo Lingua Italiana](#) (Beispiele für Italienisch)

4.4 Leitfaden für Berichte über die Standardisierungsphase

Der Leitfaden für Berichte zum Familiarisierungsprozess (Kapitel 2) gilt auch für die Standardisierung. Wie in den vorherigen Kapiteln bieten die während des Prozesses ausgefüllten Dokumentationsbogen auch hier eine Übersicht über die durchgeführten Aktivitäten und Informationen in der Reihenfolge, in der sie durchgeführt wurden. Die Dokumentationsbogen mit den Berichten zur durchgeführten Arbeit bilden die Basis der Argumentation zur Anbindung an den GeR. Es ist deshalb wichtig, sich sorgfältig Gedanken über die Zielgruppe und die zu vermittelnde Botschaft zu machen, wenn der Bericht für diese und andere Phasen zusammengestellt wird.

Um zu zeigen, dass die Ergebnisse und die Interpretation der Standardisierung valide sind, muss dokumentiert werden, **wie und mit welchen Vorkommnissen** der Standardisierungsprozess **verlaufen ist**. Jeder Schritt im Prozess sollte beschrieben und jede Entscheidung begründet werden. Folgende Punkte sollten enthalten sein:

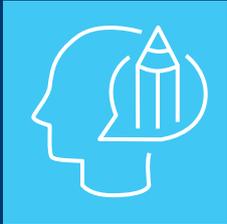
- Liste der Teilnehmenden mit Angaben zur Person
- Namen und Aufgaben der an der Organisation der Sitzung(en) beteiligten Personen
- Zeitplan der Sitzung(en)



Hinweise zur eigenen Implementierung ...

Die Standardisierungsaktivitäten sind sehr zeitintensiv. Bevor Sie beginnen, sollten Sie also den Umfang der Übung festlegen und wissen, wo Sie die wichtigen erläuternden Beispiele erhalten. Hier sind die wichtigsten Grundlagen noch einmal zusammengefasst.

- Entscheiden Sie sich je nach Anforderungen des Kontexts und der verfügbaren Ressourcen für eine Methode und Intensität der Standardisierung. Seien Sie nicht zu ambitioniert, aber versuchen Sie sicherzustellen, dass Sie genug tun, um die Ansprüche, die Sie an die Anbindung stellen wollen/müssen, zu unterstützen
- Achten Sie darauf, dass Sie auf ausreichende Mittel zugreifen können, insbesondere auf Zeit und den Zugang zu Expertinnen/Experten, die mit dem Material vertraut sind, das an den GeR angepasst werden soll, und die bei der Prüfung der Analyse helfen können.
- Organisieren Sie das Team und definieren Sie klar, wer sich worauf konzentrieren soll.
- Planen Sie ausreichend Zeit für Diskussionen ein.
- Ermöglichen Sie den Austausch mit externen Expertinnen/Experten und Stakeholder/-innen.
- Achten Sie darauf, dass Sie alle Informationen zusammentragen, die Sie für den finalen Bericht benötigen, u. a. die Dokumentationsbögen am Ende jeder Sitzung.
- Vergewissern Sie sich, dass Daten sicher und im Einklang mit den datenschutzrechtlichen Vorschriften erhoben und gespeichert werden (z. B. gemäß der DSGVO in der EU). Dabei kann es erforderlich sein, dass diejenigen, die die Daten bereitstellen, und die potenziellen Nutzenden über das vollständige Spektrum der Verwendungszwecke transparent informiert werden müssen.



KAPITEL 5: STANDARD-SETTING

Das Standard-Setting basiert auf dem empirischen Prozess, quantitative Nachweise zu sammeln, um Entscheidungen *über geeignete Bestehensgrenzen auf Punkteskalen* zu treffen. Dieses Vorgehen wird in erster Linie mit Tests und Prüfungen assoziiert. Die Methoden rund um das Standard-Setting wurden entwickelt, um Tests und Prüfungen an Standards, Qualifikationsstufen und Referenzrahmen anzubinden, die entscheidende Leistungsmerkmale sowie Kannbeschreibungen definieren. Das Standard-Setting hilft dabei, die folgende Frage zu beantworten: *An welcher Stelle auf der Punkteskala eines bestimmten Tests können wir zu dem Ergebnis kommen, dass die geprüfte Person einen ausreichenden Leistungsnachweis für eine bestimmte Stufe erbracht hat?*

Der entsprechende Leistungsstand der betreffenden Stufe wird durch eine Reihe von Standards, Qualifikationsstufen oder Referenzrahmen mit Beschreibung der Leistungsniveaus definiert.

In diesem Kapitel liegt der Fokus im Vergleich zu den vorausgegangenen Kapiteln also mehr auf Tests und Prüfungen. Das bedeutet aber nicht, dass das Standard-Setting für die anderen in diesem Handbuch behandelten Materialien nicht relevant ist. Es wird jedoch nicht immer praktikabel sein zu versuchen, die für Tests und Prüfungen relevanten Methoden vollständig auf andere Materialien anzuwenden.

Dieses Kapitel ist folgendermaßen aufgeteilt:

- 5.1 Was versteht man unter Standard-Setting?
- 5.2 Beispiele zum Standard-Setting für Rezeption und Produktion
- 5.3 Herausforderungen, Lösungen und Alternativen
- 5.4 Anpassung der Methoden zum Standard-Setting bei anderen Materialien als Prüfungen
- 5.5 Leitfaden für Berichte über die Standard-Settingphase
- 5.6 Weiterführende Literatur

Am Ende des Kapitels sollten Sie von den folgenden Punkten eine klare Vorstellung haben:

- Das Konzept des Standard-Settings
- Inwiefern das Standard-Setting mit den anderen Phasen des Verlinkungsprozesses zusammenhängt
- Die praktischen Schritte bei der Planung und Ausführung des Standard-Settings
- Die logistischen und ressourcenbezogenen Anforderungen an das Standard-Setting
- Typische Datenerhebungsaktivitäten im Rahmen des Standard-Settings
- Wie das Standard-Setting mit den verschiedenen Arten von Materialien, die in diesem Handbuch beschrieben werden, zusammenhängt (Prüfungen, Curricula, Lehrwerke usw.)
- Die Wichtigkeit der Dokumentation und Berichterstellung
- Wo detailliertere technische Informationen zu den Methoden des Standard-Settings sowie Beispiele für deren Anwendung zu finden sind

Hinweis: Alle Methoden zum Standard-Setting umfassen statistische Analysen quantitativer Daten. Wir werden diese Prozesse jedoch nur umreißen. Abschnitt 5.6 enthält eine Liste von Literaturangaben, um diese technischen Aspekte des Standard-Settings detaillierter zu erfassen.



Ein Hinweis zu diesem Kapitel: Erwartungen definieren

Wie bereits erwähnt, wurde das Standard-Setting zur Verwendung für Tests und Prüfungen konzipiert. Das erklärt, warum Kapitel 5 einen etwas anderen Ansatz als die restlichen Kapitel verfolgt. Das Handbuch konzentriert sich auf die Nachweise, die gesammelt werden müssen, um die Anbindung verschiedener Materialien an den GeR zu unterstützen: Curricula, Lehr- und Lernmaterialien und Prüfungen. In Kapitel 1 werden diese Materialien als Komponenten eines ganzheitlichen Lernsystems präsentiert (siehe Abbildung 1.2 in Abschnitt 1.4). Da sie entwickelt wurden, um verschiedene Funktionen zu erfüllen, haben diese Komponenten folglich auch nicht dieselben Merkmale. Es wird auch nicht erwartet, dass die fünf Verlinkungsphasen, die in Kapitel 1 beschrieben wurden, bei jeder Komponente des Dreiecks zur Beschreibung des ganzheitlichen Lernsystems auf exakt dieselbe Weise implementiert werden können. Prüfungen sind oft die „harten Fakten“, nach denen beurteilt wird, ob ein/-e Lernende/-r das Lernziel erreicht hat und wie wirksam die auf dem Curriculum basierende Vermittlung von Wissen tatsächlich war. Daher ist es entscheidend, beim Standard-Setting den Fokus auf Prüfungen zu legen. So kann sichergestellt werden, dass die Beurteilung der Lernenden und die Evaluation in anderen Bereichen des ganzheitlichen Lernsystems durch die Nutzung der an den GeR angebotenen Bewertungskriterien fair ablaufen. In den Abschnitten 5.1–3 liegt der Fokus deshalb auf das Standard-Setting hinsichtlich der Bewertungsinstrumente. In Abschnitt 5.4 wird wiederum thematisiert, wie evidenzbasierte Ansätze zum Nachweis der Verlinkung von Prüfungen anhand eines Standard-Setting modifiziert werden können, um die Verlinkungsnachweise auch von anderen Materialien zu stützen.

5.1 Was versteht man unter Standard-Setting?

5.1.1 Warum setzen wir Standards?

Cizek beschreibt das Standard-Setting als „die ordnungsgemäße Vorgehensweise beim Befolgen eines vorgegebenen, rationalen Systems mit Regeln oder Verfahren für die Zuweisung einer Zahl, um zwischen zwei oder mehr Leistungsphasen oder -graden zu unterscheiden“ (siehe Abschnitt 5.6, Nr. 3). Diese Definition spiegelt die Ursprünge des Standard-Settings in den USA wider, insbesondere im Rahmen großer standardisierter Prüfungen. Der GeR hat seit seiner Veröffentlichung im Jahr 2001 zahlreiche innovative Herangehensweisen bei der Anwendung und Anpassung der Methoden zum Standard-Setting inspiriert. Das Ziel einer solchen Aufgabe besteht darin, die Stelle auf einer Punkteskala zu ermitteln, an der die getestete Person ausreichend Nachweise darüber erbracht hat, die in einer der GeR-Niveaus beschriebenen Kompetenzen zu besitzen. Wenn bei einer Prüfung die Ergebnisse der Prüfungsteilnehmenden in Bezug zu den GeR-Niveaus ermittelt werden sollen, ist das Standard-Setting laut Kaftandjieva (2004) „der zentrale Aspekt des Verlinkungsprozesses“.

Wie in Kapitel 3 erwähnt, handelt es sich bei der Verlinkung um einen Top-down-Prozess, wenn eine neue Prüfung basierend auf dem GeR erstellt wird, und um einen Bottom-up-Prozess, wenn der Versuch unternommen wird, eine bestehende Prüfung an den GeR anzupassen. In beiden Fällen müssen vor der Planung des Standard-Settings die in Kapitel 3 und 4 erklärten Schritte ausgeführt werden. Auf diese Weise erhält man nicht nur ein detailliertes Verständnis der GeR-Niveaustufe(n), auf die sich die Bewertung bezieht, sondern auch der spezifischen Skalen und Deskriptoren, an denen die Aufgaben und Items ausgerichtet werden sollten. Außerdem spielt die Entwicklung klarer Testspezifikationen sowie der Nachweis der (technisch-handwerklichen) Qualität der Prüfung, beispielsweise der Reliabilität usw., eine zentrale Rolle. Wie im *Manual for Relating Examinations to the CEFR* (Abschnitt 5.6, Nr. 4) beschrieben, ist die „Verlinkung einer qualitativ niedrigen Prüfung mit dem GeR ein sinnloses Unterfangen, da dies durch ein sorgfältiges Standard-Setting weder gerettet noch in Ordnung gebracht werden kann“. Es ist wichtig zu beachten, dass das Standard-Setting selbst keine Validierung der Prüfung darstellt, unabhängig davon, ob das Standard-Setting erfolgreich ist oder nicht.

Es ist auch wichtig hervorzuheben, dass Standard-Setting und Validierung auch für kleinere Prüfungen in Institutionen und Schulen nicht minder relevant sind als für umfangreiche Tests im großen Stil.

Wenn der Verlinkungsprozess kein sinnloses Unterfangen sein soll, sollte die Beurteilung auf allen Niveaus den Prinzipien einer guten Testentwicklung folgen: Testspezifikationen, klar definierte Bewertungsskalen und Analyse und Evaluation der Testergebnisse. Es gibt viele hilfreiche Diskussionen darüber, wie effiziente Prinzipien für die Testentwicklung und -validierung bei der Durchführung von Prüfungen in Schulen angewendet werden können.

Auf Basis der GeR-bezogenen Nachweise, die in den Spezifikations- und Standardisierungsphasen (Kapitel 3 und 4) gesammelt wurden, und der technisch-handwerklichen Nachweise, die die Qualität des Tests bestätigen, erhält man bereits eine ziemlich klare Vorstellung davon, wo die Bestehensgrenze(n)⁴ für die gewünschten GeR-Niveaus liegen. Im Standard-Setting sollte sich nichts völlig Neues über die Bedeutung der Testergebnisse herausstellen. Tests sind jedoch nie absolut korrekt und keine präzisen Instrumente – auch wenn das für Nichtspezialistinnen und Nichtspezialisten manchmal schwer zu akzeptieren ist – und enthalten immer den einen oder anderen Messfehler⁵. Das Standard-Setting ist daher aus zweierlei Gründen erforderlich: Zunächst muss ermittelt werden, wo sich die Bestehensgrenzen auf der Skala tatsächlich befinden. Dazu ist es notwendig, die erste Einschätzung zu verfeinern und anzupassen. Zudem müssen Stakeholder empirische Nachweise erhalten, damit begründet werden kann, wie die Testergebnisse hinsichtlich der GeR-Niveaus zustande kommen. Auch wenn oben erwähnt wurde, dass das Standard-Setting selbst keine Validierung der Prüfung darstellt, trägt sie doch maßgeblich zur Validierung bei. Falls durch den Nachweis im Rahmen des Standard-Settings *nicht* bestätigt wird, dass eine Anbindung der Prüfung an den GeR erfolgte, wurden im Prozess trotzdem Informationen gesammelt, mit deren Hilfe das Design und die Durchführung der Prüfung überarbeitet werden können.

5.1.2 Wie wird das Standard-Setting durchgeführt?

Das Standard-Setting ist eine Möglichkeit, quantitative, empirische Nachweise für die Entscheidungen über die Bewertung in Bezug auf den GeR zu erhalten. Die quantitative Natur des Standard-Settings sollte jedoch nicht zu unkritischem Vertrauen in die „Zahlen“ führen. Es ist wichtig hervorzuheben, dass alle gängigen Methoden zu einem gewissen Maß menschliche Beurteilungen erfordern und dass es keine allgemein gültigen Methoden oder Ansätze gibt. Verschiedene Methoden zum Standard-Setting stehen zur Auswahl und auch zur Implementierung gibt es nun zahlreiche Werke über die Anbindung von Prüfungen an den GeR (Fallstudien siehe Abschnitt 5.6, Weiterführende Literatur und Links 5, 6, 7, 9, 10 und 11).

Für welche Methode man sich entscheiden sollte, hängt von verschiedenen Faktoren ab. Manche Methoden wurden für bestimmte Prüfungsarten entwickelt oder haben sich als passendste Option dafür erwiesen. Bei der Auswahl der Methode ist es daher notwendig, die jeweiligen Merkmale einer Prüfung zu berücksichtigen, also die rezeptiven oder produktiven Fähigkeiten, Antwortformate usw. Zudem müssen weitere Faktoren berücksichtigt werden:

- Verfügbare Ressourcen
- Die Erfahrung oder das Know-how der Person oder des Teams, die bzw. das für das Standard-Setting zuständig ist
- Die Verfügbarkeit der Expertinnen und Experten, die am Standard-Setting teilnehmen
- Zugang zu Testinhalten, Testdaten und Leistungsbeispielen zur Prüfung

4 Die Bestehensgrenze (cutscore) ist die niedrigste mögliche Punktzahl bei einem Test oder einer Prüfung, mit der man besteht. In diesem Fall ist es die niedrigste Punktzahl, mit der noch bestätigt werden kann, dass ein/-e Prüfungsteilnehmende/-r das spezifische GeR-Niveau aufzeigt.

5 Ein Messfehler bezieht sich auf die Differenz zwischen dem beobachteten Wert einer Variablen (hier das Testergebnis eines Lernenden) und dem „wahren Wert“ (true value) dieser Variablen (der tatsächlichen Fähigkeit der/des Prüfungsteilnehmenden).

Das Standard-Setting ist von Natur aus ein praxisorientiertes Vorhaben. Vor, während und nach der Anwendung der ausgewählten Methode müssen die logistischen Vorkehrungen sorgfältig getroffen werden. Dies gilt unabhängig von der gewählten Methode darüber hinaus auch für die mehr technisch ausgerichteten Aspekte beim Sammeln und Analysieren der Daten für das Standard-Setting selbst.

In diesem Kapitel wird nicht im Detail erklärt, wie die verschiedenen verfügbaren Methoden angewendet werden sollten. Wir fassen jedoch einige der zentralen Merkmale der gängigsten Methoden zusammen. Diese Merkmale sind für die Prozesse und logischen Überlegungen relevant, die in Abschnitt 5.1.3 behandelt werden. Bei der Planung des Standard-Settings sollten die folgenden Überlegungen in Betracht gezogen werden:

- Es ist üblich, zwischen zwei Methodenarten zu unterscheiden. Einerseits die **testzentrierten** Methoden, die Items, Aufgaben usw. beurteilen und andererseits die **prüfungsteilnehmerzentrierten** Methoden, die den Fokus entweder auf die Beurteilung von Prüfungsteilnehmerleistungen legen oder auf Hinweise auf das Leistungsniveau von Lernenden, bevor diese die Prüfung ablegen.
- Die Beurteilung erfolgt durch ein Gremium aus Expertinnen und Experten auf der Grundlage einer durch die ausgewählte Methode definierten Aufgabe. Die Expertinnen und Experten müssen in der Verwendung dieser Methode geschult sein.
- Die Testinhalte der entsprechenden Prüfung, die an den GeR angepasst werden sollen, müssen dem Expertengremium zugänglich gemacht werden. Bei prüfungsteilnehmerzentrierten Ansätzen kann es auch notwendig sein, Leistungsbeispiele der Prüfungsteilnehmenden einschließlich der Punkte, die sie nach den für das Prüfungsformat verwendeten Bewertungskriterien erhalten haben, zu sammeln.
- Es müssen Verfahren zur Analyse der erhobenen Daten festgelegt werden.
- Für die Festlegung der Bestehensgrenze muss basierend auf der Analyse der gesammelten Daten und entsprechend der gewählten Methode eine Regel definiert werden.

Bei der Diskussion, **wie** das Standard-Setting funktioniert, sollte beachtet werden, dass die anderen Verlinkungsphasen, die in Kapitel 1 zusammengefasst wurden – Familiarisierung, Testspezifikation, Standardisierung und Validierung –, entscheidende Schritte bei der Entwicklung und Anbindung des betreffenden Materials sind, bevor das Standard-Setting erfolgt. Mehr oder weniger stark sind sie auch im Mikrokosmos des Standard-Settingsprozesses enthalten:

Familiarisierung ist ein entscheidendes Schulungselement, um sicherzustellen, dass die Expertengremien die ausschlaggebenden Merkmale der Aspekte (Deskriptoren) des GeR, die beim Standard-Setting im Fokus stehen, vollständig erfassen.

Spezifikation ist etwas, das die zuständige Person für den Verlinkungsprozess vor dem Beginn des Standard-Settings selbst durchführen sollte. Dieser Schritt ist entscheidend, um die beabsichtigte Anbindung der Prüfung oder eines anderen Materials an den GeR zu verstehen. Bei der Entwicklung einer neuen Prüfung oder Materials ist die Spezifikation ein entscheidender Teil des Entwicklungsprozesses. Dabei kann eine zusätzliche Aktivität zur Spezifikation aussagekräftige Nachweise liefern, um die im Rahmen des quantitativen Standard-Settingsprozesses gesammelten Daten zu triangulieren. Die Durchführung einer Spezifikation als Ergänzung zum Standard-Setting ist besonders wichtig, wenn bestehendes Material angeglichen werden soll, das möglicherweise nicht mit Blick auf die GeR-Anbindung entwickelt wurde.

Standardisierung ist auch ein entscheidender Schritt bei der Testentwicklung vor dem Standard-Setting, insbesondere für die Autorinnen und Autoren von Items und für andere wichtige Beteiligte am Entwicklungsprozess. Einige der in Kapitel 4 beschriebenen Standardisierungsaktivitäten können in Schulungen zum Standard-Setting verwendet werden, insbesondere beim Sammeln von lokal geeigneten Leistungsbeispielen, die mit den GeR-Niveaus verglichen werden. Schließlich ist es auch wichtig, Nachweise zu sammeln, die helfen, den Prozess des Standard-Settings zu analysieren und festzustellen, ob die Entscheidungen bezüglich der Bestehensgrenze plausibel und vertretbar sind.

5.1.3 Praktisches zum Standard-Setting: Planung, Organisation, Datenerhebung und Analyse

Wie oben erwähnt, ist das Standard-Setting eine essenzielle praktische Aktivität mit erheblichen logistischen Anforderungen. Wir betonen jedoch, dass viele Methoden zum Standard-Setting zur Nutzung in spezifischen Kontexten und im kleinen Rahmen angepasst werden können, beispielsweise bei internen Prüfungen, die innerhalb eines bestimmten Programms oder einer bestimmten Institution entwickelt wurden. Abschnitt 5.2 umfasst einige Beispiele der Methoden zum Standard-Setting hinsichtlich einfacherer Formate von Sprachtests basierend auf Szenarien, die für mittlere bis große Organisationen typisch sind. Während diese Beispiele relativ ressourcenintensiv sein können, werden in Abschnitt 5.3 einige mögliche Alternativen und Anpassungen an Kontexte für kleinere Maßstäbe.

Es kann nicht oft genug darauf hingewiesen werden, wie wichtig es ist, sich bei der Planung des Standard-Settings auf die praktischen Aspekte des Prozesses zu konzentrieren. In Tabelle 5.1 sind einige der wichtigsten Schritte und Entscheidungen aufgelistet, die bei der Planung und Vorbereitung des Standard-Settings erforderlich sind. Diese Liste mag auf den ersten Blick abschreckend erscheinen, in Abschnitt 5.3 wird jedoch erörtert, was diese Aspekte für kleinere, ressourcenbeschränkte Kontexte bedeuten könnten.

Tabelle 5.1: Wichtige Schritte bei der Planung und Vorbereitung einer Aktivität zum Standard-Setting

Methode auswählen
Machen Sie sich mit der Literatur zum Standard-Setting vertraut. Vielleicht haben Sie die finanziellen Ressourcen, um eine Expertin bzw. einen Experten zu beauftragen, um die notwendigen Aktivitäten durchzuführen. Nichtsdestotrotz müssen Sie wissen, welche Optionen verfügbar sind, wenn Sie die Empfehlungen Ihrer Expertin bzw. Ihres Experten kritisch evaluieren und Ihre Entscheidungen vor wichtigen Entscheidungsträgerinnen und -trägern rechtfertigen möchten. In vielen Fällen kann es erforderlich sein, das Standard-Setting selbst zu planen, zu leiten und durchzuführen.
Die erforderlichen Ressourcen zur Durchführung des Standard-Settings ermitteln
Physisch: Ausrüstung (z. B. Laptops, spezielle Software, Tontechnik für das Abspielen von Hör-Dateien); Räume oder Onlineplattformen für Meetings zum Standard-Setting. Finanziell: Vergütung des Expertengremiums; Kosten für die Vorbereitung und den Druck von Materialien, Leihgebühr für Laptops für Expertinnen, Experten usw.; Anmietung von Räumen für präsentische Aktivitäten des Standard-Settings usw. Personell: administrative Unterstützung bei den Aktivitäten; Schulung von Expertinnen und Experten; Durchführung des Verfahrens (z. B. Aushändigen und Einsammeln von Materialien, Unterstützung des Expertengremiums bei technischen Problemen usw.); Analyse der Daten.
Relevante Expertinnen/Experten für das Gremium oder die Gremien finden
Sie müssen Kriterien für die Auswahl von Expertinnen und Experten festlegen. Es kann eine Herausforderung sein, ausreichend qualifizierte Personen zu finden, die sowohl an einer Teilnahme interessiert als auch verfügbar sind. Für die in Abschnitt 5.2 beschriebenen Methoden sind 12 bis 15 Expertinnen und Experten angemessen. Wenn Sie jedoch mehrere Tests verlinken oder das Standard-Setting in den Bereichen Lesen, Hören, Sprechen und Schreiben durchführen möchten, benötigen Sie unter Umständen mehr als ein Gremium. Sehen Sie sich dazu den Kasten 3 zur Methodik in diesem Kapitel an.

Zugriff auf die erforderlichen Testmaterialien gewährleisten

Bei testzentrierten Methoden kann es sich dabei um die eigentlichen Testinhalte handeln, also um die Testaufgaben. Dies mag zwar einfach erscheinen, doch die Verwendung echter Testaufgaben birgt einige Sicherheitsrisiken. Bei Tests zum Hören ist es notwendig, Zugang zu den tatsächlichen Audioaufnahmen bereitzustellen. Bei computergestützten Tests wiederum müssen Gremiumsmitglieder in der Lage sein, die Testinhalte auf die gleiche Weise wie die Teilnehmenden zu sehen. Die Methode kann die Verwendung tatsächlicher mündlicher oder schriftlicher Leistungsbeispiele der Testteilnehmenden erfordern.

Zugriff auf die für die Analyse erforderlichen Testdaten gewährleisten

In der Regel ist es wichtig, Testitems mit den Schwierigkeitswerten einer repräsentativen Population vorliegen zu haben. Bei prüfungsteilnehmerzentrierten Methoden, bei denen tatsächliche Leistungsbeispiele verwendet werden, kann eine Bewertung durch geschulte Bewertende erforderlich sein.

Kommunikation vorbereiten

Dies umfasst Einladungsschreiben für Expertinnen und Experten, in denen der Zweck des Projekts erläutert wird (siehe Kapitel 4); DSGVO-konforme Einwilligungformulare und Informationen rund um die Logistik für die Teilnehmenden. Es kann auch eine Geheimhaltungsvereinbarung (Non-Disclosure Agreement, NDA) erforderlich sein, wenn echte Testmaterialien und Leistungsbeispiele von Prüfungsteilnehmenden verwendet werden sollen.

Schulungsmaterialien für den GeR entwickeln

Für den GeR: Trainingsaspekte der Familiarisierung und Standardisierung, wie in den Kapiteln 2 und 4 beschrieben.

Für die Methode: Beschreibungen effektiver Trainings, angepasst an Beispiele der ausgewählten Methode.

Instrumente zur Datenerhebung vorbereiten

Formulare für das Expertenvotum: Zur Vermeidung von Eingabefehlern, müssen die Formulare für die Entscheidungen einzelner Expertinnen und Experten einfach zu verwenden sein. Es kann ratsam sein, digitale Versionen einzusetzen.

Feedbackformulare: Feedback von Expertinnen und Experten ist wichtig, um die Validität des Verfahrens nachzuweisen.

Tools zur Zusammenführung und Analyse von Daten vorbereiten

Bei vielen Methoden wird empfohlen, mehr als eine Runde zur Abgabe des Expertenvotums durchzuführen und die Zwischenergebnisse den Expertinnen und Experten zur Diskussion zu präsentieren. Dies erfordert, die Ergebnisdaten während der Aktivitäten zum Standard-Setting schnell zu erheben, zusammenzuführen und zu analysieren. Zu diesem Zweck sollten Sie formatierte Tabellen oder andere Möglichkeiten zum Zusammenführen einzelner Expertenergebnisse in Datensätze vorbereiten.

Räumlichkeiten vorbereiten

Der ausgewählte Raum sollte ausreichend Platz für das Gremium bieten. Sie müssen sicherstellen, dass alle erforderlichen Geräte und Ausstattungen bereitstehen. Es kann auch sein, dass Sie zusätzlichen Platz benötigen, um sensible Testmaterialien sicher zu verstauen. Übernachtungs- und Verpflegungsmöglichkeiten sowie Snacks können ebenfalls erforderlich sein.

In den Beispielen in Abschnitt 5.2 dauern die Veranstaltungen zum Standard-Setting zwei oder drei Tage. Anhand der Liste mit logischen Überlegungen in Tabelle 5.1 sollte allerdings klar sein, dass die Planung und Vorbereitung bereits einige Zeit vor der Veranstaltung beginnen muss. Wie zuvor erwähnt, dient der Dokumentationsbogen 5 dazu, Ihnen bei der Planung des Standard-Settings zu helfen. Es bietet eine Checkliste der wichtigsten Punkte, die zu berücksichtigen sind.

5.2 Beispiele zum Standard-Setting für Rezeption und Produktion

In diesem Abschnitt stellen wir Beispiele aus der Praxis für die Abfolge von Aktivitäten vor, die im Mittelpunkt der Datenerhebung für das Standard-Setting stehen, wenn diese auf Sprachtests zum Zweck der Anbindung an den GeR angewendet werden. Der Fokus liegt auf Zeitplänen und der Reihenfolge der Aktivitäten, nicht auf Details wie beispielsweise der Auswahl der Expertinnen und Experten oder dem Erheben und Analysieren von Daten des Expertenvotums. Die ausgewählten Methoden sind geeignet, vertretbar und praktikabel und wurden sowohl bei der Bewertung im Bildungsbereich allgemein als auch in Bezug auf den GeR schon häufig eingesetzt. Die zur Verwendung in den Beispielen in diesem Abschnitt angepassten Methoden werden in Abschnitt 5.6, Quellen 2, 3 und 4 beschrieben. Detaillierte technische Berichte zur Anwendung von bestimmten kombinierten Methoden sind in Abschnitt 5.6, Quellen 9 und 10 zu finden.

Diese Quellen befassen sich mit der Verlinkung von Studien für lokal entwickelte Standards, nicht für den GeR selbst. Sie umfassen jedoch ausreichend detaillierte Beschreibungen der Datenerhebung, der statistischen Analyseprozesse und Ergebnisse, der Verfahren zur Bestimmung von Bestehensgrenzen und der Validierungsnachweise.

Einer der Gründe für die Verwendung der unten aufgeführten Beispielmethoden ist die Tatsache, dass diese ausgewählten Methoden viele der wichtigsten Schritte und Merkmale enthalten, die alle häufig angewendeten Methoden aufweisen. Sie zeigen daher deutlich die Schritte, die im Rahmen einer Standard-Setting-Konferenz durchgeführt werden, unabhängig von der verwendeten Methode und der für eine erfolgreiche Durchführung erforderlichen Zeit. Dass sie als Beispiele verwendet werden, soll jedoch nicht implizieren, dass es sich dabei um die „besten“ Methoden für Ihren bestimmten Kontext handelt. Variationen und Alternativen werden in Abschnitt 5.3 behandelt.

5.2.1 Beispiel für das Standard-Setting für Rezeption

Nachfolgend wird ein Zeitplan für eine Standard-Setting-Konferenz zum Prüfungsteil Hören vorgestellt. Er kombiniert zwei häufig verwendete **testbasierte** Methoden: die **Basket-Methode** und die **modifizierte Angoff-Methode** (siehe Kasten 1 zur Methodik). In unserem Beispiel werden die Daten, die durch die modifizierte Angoff-Methode gewonnen wurden, von vornherein zum Festlegen der Mindestpunktzahlen verwendet. Die Methoden wurden kombiniert, um die Stärken zu ergänzen und die jeweiligen Schwächen auszugleichen. Die Basket-Methode wurde als Aufwärmübung eingesetzt, um die Expertinnen und Experten dazu anzuregen, sich über die Items bezüglich ihrer entsprechenden GeR-Niveaus Gedanken zu machen. Diese anfängliche Einschätzung dient als Ausgangspunkt zur Festlegung der Bandbreite für die Einschätzungen, die daraufhin mithilfe der modifizierten Angoff-Methode gewonnen werden. Ziel dieses Vorgehens ist es, die kognitive Komplexität der Aufgabe zu verringern.

Noch einmal als Hinweis: Durch die Präsentation dieser Beispiele soll nicht der Eindruck erweckt werden, dass es sich dabei um den einzigen oder „besten“ Ansatz beim Standard-Setting handelt. Vielmehr soll dadurch verdeutlicht werden, dass die Methoden je nach Kontext adaptiert werden können bzw. sollten, sofern notwendig. Dabei ist es jedoch entscheidend, diese Anpassungen klar zu begründen, sie explizit zu dokumentieren und durch eine Analyse der Ergebnisse zu validieren: Haben die Anpassungen funktioniert?

Tabelle 5.2: Beispielzeitplan für Aktivitäten zu einem Test der Fertigkeit Hören

Tag 1	
09:30	Einführung in den Umfang und die Ziele des Seminars: Selbstvorstellung
10:00	Überblick über den GeR: Prinzipien, Entwicklungsmethodik, Struktur
10:30	Pause
10:45	Auffrischung: Aktivitäten mit den wichtigsten Deskriptoren für Hören
11:30	Diskussion: Überlegungen zu den Aktivitäten im Booklet zum Selbststudium ⁶
12:30	Mittagessen
14:00	Überblick über die Prüfung in diesem Projekt
15:00	Standard-Setting mit Hör-Items: Erklärung und praktische Übungen
16:30	Pause
16:45	Abschließende Diskussion (Einigung auf eine Liste der Deskriptoren und entscheidenden GeR-Merkmale)
Tag 2	
09:00	Rückblick: Deskriptoren, Liste der entscheidenden GeR-Merkmale und der Test der Fertigkeit Hören
09:45	Test der Fertigkeit Hören (BASKET-METHODE)
11:00	Pause
11:30	Diskussion, Präsentation der Schwierigkeitswerte der Items (Mitglieder des Gremiums können Anpassungen vornehmen oder auf Wunsch bestimmte Items noch einmal anhören.)
12:30	Mittagessen
14:00	Test der Fertigkeit Hören (MODIFIZIERTE ANGOFF-METHODE), Runde 1
15:30	Pause
15:45	Organisation und Überprüfung des Plans für den letzten Tag
16:00	Diskussion, Präsentation der Schwierigkeitswerte der Items
16:30	Test der Fertigkeit Hören (MODIFIZIERTE ANGOFF-METHODE), Runde 2

⁶ Dieser Teil bezieht sich auf die Familiarisierung und besteht entweder aus einer papierbasierten oder digitalen Aufbereitung einer Reihe von Aufgaben zum Selbststudium, die von der Koordinatorin bzw. vom Koordinator auf Grundlage der exemplarischen Aufgaben des Europarats zu der jeweiligen GeR-Zielniveaustufe(n) angepasst werden. Das Booklet sollte von den Mitgliedern des Gremiums vor der Standard-Setting-Konferenz bearbeitet werden.

Es ist notwendig, einige der Kernpunkte dieses Zeitplans näher auszuführen.

- Das Expertengremium führt auch das Standard-Setting für den Prüfungsteil Lesen des gleichen Tests durch. Es ist leicht zu erkennen, dass sich der Zeitaufwand für die Aktivitäten schnell summiert. Das macht es für die Expertinnen und Experten schwierig, für alle Aktivitäten verfügbar zu sein, da diese sich über mehrere Tage erstrecken. Das ist eine der häufigsten logistischen Herausforderungen bei der Organisation eines Expertengremiums. Die Lösung, wie später in Abschnitt 5.3 beschrieben, besteht darin, viele der Familiarisierungs- und Standardisierungsaktivitäten als Material zum Selbststudium anzubieten, die die Expertinnen und Experten vor der zweitägigen Standard-Setting-Konferenz erarbeiten müssen. Dadurch ist es möglich, den größten Teil der verfügbaren Präsenzzeit für die Schulung der verwendeten Methode(n) und die tatsächliche Standardsetzung einzuplanen.
- Für die Hauptmethode des Standard-Settings, die für die Festlegung der Bestehensgrenze(n) verwendet wird, sind zwei Beurteilungsrunden mit einer Gesprächsrunde im Plenum dazwischen vorgesehen. Das ist ein gemeinsames Merkmal vieler Methoden. Da in diesem Beispiel die Basket-Methode für einen anderen Zweck verwendet wird, gibt es nur eine vollständige Beurteilungsrunde. Eine Diskussion ist aber trotzdem wichtig.
- Während der Gesprächsrunde werden üblicherweise die Beurteilungen der Expertinnen und Experten als Basis für die Identifikation von Hauptbereichen der Übereinstimmungen sowie Abweichungen vorgestellt. Eine 30-minütige Pause zwischen der ersten Beurteilungsrunde und der Diskussion erleichtert es den Moderatorinnen und Moderatoren, die Beurteilungen einzusammeln, sie im Excel-Format aufzubereiten, eine erste Analyse durchzuführen und die Ergebnisse zur Präsentation vor dem Expertengremium in ein grafisches Format zu übertragen.
- Bevor die Expertinnen und Experten gebeten werden, ihre ersten Beurteilungen anhand der Basket-Methode abzugeben, müssen sie den Prüfungsteil Hören unter denselben Bedingungen wie die Prüfungsteilnehmenden absolvieren, d. h. die im Test verwendeten Audioaufnahmen mit denselben Zeitvorgaben anhören und die Fragen dazu beantworten. Sie erhalten dann die Lösungen zu den Items, damit sie ihre Antworten überprüfen können, bevor sie mit ihrer Beurteilung fortfahren. Um sie in diesem Prozess zu unterstützen, erhalten die Expertinnen und Experten ein Transkript der Audios zu den entsprechenden Hör-Items. Das spart Zeit und hilft ihnen dabei, ihren Fokus von der Perspektive der Prüfungsteilnehmenden auf die einer Expertin bzw. eines Experten zu ändern. Ein weiterer hilfreicher Nachweis kann erbracht werden, indem die Expertinnen und Experten gebeten werden, sich auf die Genauigkeit ihrer Entscheidungen zu konzentrieren. Das kann entweder durch ein kurzes Interview oder mithilfe eines Berichtstools mit Likert-Skala geschehen (Skala von 1–5, auf der 1 für „geringe Genauigkeit“ und 5 für „hohe Genauigkeit“ steht).

5.2.2 Beispiel zum Standard-Setting für Produktion

Nachfolgend wird ein Beispiel für eine Standard-Setting-Konferenz zu den Prüfungsteilen Sprechen und Schreiben eines Tests dargestellt. Dabei kommt eine häufig verwendete prüfungsteilnehmerzentrierte Methode zum Einsatz, eine Adaptation der **Methode zur analytischen Beurteilung** (siehe Kasten 2 zur Methodik). Im Beispiel werden zwei Prüfungsteile betrachtet, um noch einmal hervorzuheben, wie man den gewählte Ansatz für unterschiedliche Situationen anpassen muss. Wie im Abschnitt 5.2.1 beschrieben, besteht das Ziel darin, einige gemeinsame Aspekte zu veranschaulichen, die auf alle Methoden angewendet werden können. Es soll nicht der Eindruck erweckt werden, dass die beschriebene Methode oder deren Anwendung der einzige oder beste Ansatz ist. Der Zeitplan unterscheidet sich deutlich von dem, der in Abschnitt 5.2.1 präsentiert wird. Auch das dient zur Demonstration, wie Anpassungen je nach Kontext durchgeführt werden können und sollten.

Tabelle 5.3: Zeitplan für Aktivitäten zum Prüfungsteil Sprechen und Schreiben

Tag 1	
09:00	Einführung in den Umfang und die Ziele des Workshops; Selbstvorstellung
09:30	Auffrischung der GeR-Skalen zu Sprechen/Schreiben: Aktivitäten mit wichtigen GeR-Deskriptoren für die mündliche oder schriftliche Produktion und/oder Interaktion (je nach Bedarf)
10:30	Pause
11:00	Finalisierung der Beschreibung der mündlichen Produktion und/oder Interaktion
12:00	Mittagessen
13:30	Auffrischung der GeR-Skala zu Schreiben: Aktivitäten mit wichtigen GeR-Deskriptoren für Schreiben
14:30	Finalisierung der Beschreibung der schriftlichen Produktion und/oder Interaktion
15:30	Pause
16:00	Überblick über die Prüfung in diesem Projekt
17:00	Standard-Setting mit Items zum Prüfungsteil Sprechen und Schreiben: Erklärung und praktische Übungen
17:30	Organisation und Plan für Tag 2
18:00	Ende
Tag 2	
08:30	Überprüfung der Unterscheidungsmerkmale der gesprochenen Produktion/Interaktion
09:00	Runde 1 der Beurteilungen der Leistungsbeispiele Sprechen (einschließlich Reflexion der Beurteilungsgenauigkeit)
12:30	Mittagessen
14:00	Präsentation der und Diskussion über die Beurteilungen der Leistungsbeispiele Sprechen durch das Expertengremium
14:30	Runde 2 der Beurteilungen der Leistungsbeispiele Sprechen (einschließlich Reflexion der Beurteilungsgenauigkeit)
17:45	Organisation und Überprüfung des Plans für den letzten Tag
18:00	Ende
Tag 3	
08:30	Überprüfung der Unterscheidungsmerkmale der schriftlichen Produktion/Interaktion
09:00	Runde 1 der Beurteilungen von Leistungsbeispielen Schreiben (einschließlich Reflexion der Beurteilungsgenauigkeit)
12:30	Mittagessen
14:00	Präsentation der und Diskussion über die Beurteilungen der Leistungsbeispiele Schreiben durch das Expertengremium
14:30	Runde 2 der Beurteilungen von Leistungsbeispielen Schreiben (einschließlich Reflexion der Beurteilungsgenauigkeit)
17:45	Sicherstellen, dass alle Materialien eingesammelt wurden
18:00	Ende

Auch hier ist es notwendig, einige Kernpunkte dieses Zeitplans näher auszuführen.

- Dieses Beispiel kombiniert zwei Fertigkeiten innerhalb einer einzelnen Standard-Setting-Konferenz. Wie oben bereits erwähnt, ist die Verfügbarkeit von Fachkundigen eine ständige Herausforderung. Umso wichtiger ist es, potenzielle Synergie-Effekte zu nutzen, um die Zeit, die für die Datenerhebung zur Verfügung steht, bestmöglich auszuschöpfen.
- In diesem Beispiel wird für Sprechen und Schreiben jeweils die gleiche Methode verwendet. Dadurch werden Teilnehmende im Laufe der dreitägigen Veranstaltung zunehmend mit der Methode vertraut.
- Dieses Beispiel setzt voraus, dass die Fachkundigen mit dem GeR vertraut sind. Es wird außerdem vorausgesetzt, dass das Expertengremium, das für das Standard-Setting zu den rezeptiven Fertigkeiten (das oben genannte Beispiel für den Prüfungsteil Hören sowie eine separate Standard-Setting-Konferenz für einen Prüfungsteil Lesen) ausgewählt wurde, ebenfalls bei den produktiven Fertigkeiten zum Einsatz kommt. So können im Zeitplan die allgemeine Übersicht des GeR sowie der Hintergrund zur Entwicklung usw. weggelassen werden.
- Wie beim Beispiel zur Fertigkeit Hören wird bei diesem Zeitplan angenommen, dass die Fachkundigen im Rahmen eines Schulungspakets zum Selbststudium vorab Informations- und Standardisierungsaktivitäten erhalten und bearbeitet haben. Es ist deshalb möglich, sich auf die GeR-Auffrischungsaktivitäten zu konzentrieren, gefolgt von einer Schulung zum Standard-Setting und zur Datenerhebung.
- Da für Sprechen und Schreiben derselbe Ansatz und dieselben Herangehensweisen verwendet werden, wurden Auffrischungsaktivitäten, die Beschreibung der Tests und die Schulungen zur Methodik für beide Fertigkeiten auf den ersten Tag gelegt.
- Beachten Sie, dass bei beiden unserer Beispiele zwei Bewertungsrounden eingeplant sind. Dazwischen finden jeweils die Präsentation und Besprechung der vorläufigen Ergebnisse statt. Das ist eine wichtige Vorgehensweise dieser häufig angewandten Methoden.
- Der Zeitplan für jede Bewertungsrunde variiert jedoch je nach angewandter Methode und Art der Testergebnisse, die der Bewertungsrunde als Basis zugrunde liegen.
 - In diesem Beispiel erhalten die Fachkundigen Zugang zu den Sprech- und Schreibleistungen auf sicheren Laptops, die speziell für die Veranstaltung bereitgestellt werden. Das erlaubt ihnen, im eigenen Tempo auf die Leistungen zuzugreifen und bei Bedarf noch einmal durchzulesen oder anzuhören. Dieser Zeitplan umfasst deshalb bei jeder Bewertungsrunde ein großes Zeitfenster ohne festgelegte Pause. Die Fachkundigen können selbst bestimmen, wann sie eine Pause einlegen, um etwas zu essen oder zu trinken, sich auszuruhen oder um frische Luft zu schnappen und den Kopf frei zu bekommen (Snacks und Getränke sollten also eher auf Nachfrage und nicht zu bestimmten Zeiten gereicht werden.)
 - Im Beispiel zur Fertigkeit Hören (5.2.1) bieten 30-minütige Kaffeepausen zwischen den Runden den Moderatorinnen und Moderatoren die Möglichkeit, die Bewertungsdaten zu sammeln und erste Analysen durchzuführen. Da die Fachkundigen mehr Zeit benötigen, um die mündlichen und schriftlichen Leistungsbeispiele zu bewerten, ist bei der Arbeit mit diesen Leistungsbeispielen eine längere Mittagspause vorgesehen, damit die Moderatorinnen und Moderatoren vor der Diskussionsrunde die Daten auswerten und die grafische Darstellung der Ergebnisse vorbereiten können.
 - Beachten Sie, dass der Fokus bei den Diskussionen zwischen den Bewertungsrounden auf „gemeinsamen Punkten“ liegt. Um eine ausreichende Anzahl an Bewertungen für eine ausreichend große Menge an Leistungsbeispiele zu erhalten, werden die Fachkundigen in kleinere Gruppen mit jeweils drei oder vier Personen aufgeteilt. Ihnen wird jeweils ein Set mit identischen Leistungsbeispielen sowie ein Set mit Leistungsbeispielen speziell für ihre Gruppe vorgelegt. Bei der Diskussion in der großen Runde geht es vor allem um die Ergebnisse der identischen Sets, da sich alle Fachkundigen diese Leistungen angesehen haben.

- Am Ende des dritten Tages ist etwas Zeit eingeplant, um sicherzustellen, dass alle Materialien eingesammelt wurden. In diesem Beispiel werden Leistungen aus Live-Prüfungen verwendet. Es ist äußerst wichtig, die Testinhalte und die Leistungsbeispiele zu schützen. Je nach verwendeten Materialien, Art des Tests und Veranstaltungsort ist dieser Schritt unter Umständen am Ende eines jeden Tages erforderlich.



Kasten 1 zur Methodik: Basket- und modifizierte Angoff-Methode

Bei der **Basket-Methode** legen die Expertinnen und Experten für jedes Item das GeR-Niveau fest, ab dem eine Kandidatin oder ein Kandidat dieses erfolgreich beantworten kann. Anschließend legen sie das Item in den entsprechenden „Korb“ („Basket“) des betreffenden GeR-Niveaus. Die Basket-Methode fand bei Verlinkungsprojekten schnell Anklang und war eine Zeit lang die am häufigsten für diesen Zweck verwendete Methode. Die Methode stand allerdings auch in der Kritik. Kaftandjewa (siehe Abschnitt 5.6, Referenz 9) ist der Meinung, dass die Aufgabenstellung zur Beurteilung zu einer inhärenten Verzerrung der Ergebnisse führt. Das ist vor allem dann problematisch, wenn ein Test für ein einzelnes GeR-Niveau entwickelt wurde und Items enthält, die eine sehr eingeschränkte Bandbreite bezüglich des Schwierigkeitsgrads innerhalb dieses Niveaus umfassen. Gleichzeitig ist die Methode leicht verständlich und deshalb für die Expertinnen und Experten gut zugänglich.

Die **modifizierte Angoff-Methode** stand zwar wegen ihrer kognitiv anspruchsvollen Art der Beurteilung auch in der Kritik, hat sich aber als statistisch robust erwiesen, wird oft verwendet und ist gut erforscht. Bei dieser Methode beurteilen die Expertinnen und Experten, wie wahrscheinlich es ist, dass eine Kandidatin oder ein Kandidat auf einem bestimmten Fähigkeitsniveau das Item erfolgreich beantwortet.

Hinweis: Da die Einführung der beiden Ansätze für das Expertengremium eine Herausforderung sein kann, ist es unter Umständen hilfreicher, die Basket-Methode als Aktivität vor dem Standard-Setting einzuführen und die modifizierte Angoff-Methode am Tag bzw. an den Tagen der Veranstaltung einzusetzen.



Kasten 2 zur Methodik: Die Methode zur analytischen Beurteilung

Bei der **Methode zur analytischen Beurteilung** (engl. „Analytic Judgement method“) ist die Aufgabe konzeptuell ähnlich der Basket-Methode, die Entscheidungsregeln für das Festlegen der Bestehensgrenze(n) unterscheiden sich jedoch dahingehend, dass die in Kasten 1 zur Methodik erwähnten Probleme mit der Verzerrung vermieden werden. Die Expertinnen und Experten prüfen reale Leistungsbeispiele der Prüfungsteile Sprechen oder Schreiben der Prüfung. Es ist entscheidend, dass die Leistungen vor der Standard-Setting-Konferenz von geschulten Bewerterinnen und Bewertern anhand der Bewertungskriterien für die jeweilige Prüfung beurteilt wurden. Die Expertinnen und Experten erhalten keine Informationen zur Bewertung, die die Leistungsbeispiele eingangs durch die geschulten Bewerterinnen und Bewerter erhalten haben. Sie ordnen jedes Leistungsbeispiel dem GeR-Niveau zu, das am besten die Fähigkeit der Kandidatin oder des Kandidaten, die/der diese Leistung erbringen kann, widerspiegelt.

Nachdem die Expertinnen und Experten alle Leistungsbeispiele den GeR-Niveaus zugeordnet haben, sind für das Festlegen der Bestehensgrenze(n) verschiedene Entscheidungsregeln verfügbar. Letztendlich geht es dabei immer um das Gleiche: Die Moderatorinnen und Moderatoren nutzen die bestehenden Ergebnisse, um die Bestehensgrenze(n) auf der Punkteskala zu definieren, die eine bestmögliche Unterscheidung zwischen den GeR-Niveaus ermöglicht. Wie in Kapitel 4, Abschnitt 4.3 erwähnt, ist es aus diesem Grund wichtig, eine große Anzahl an bewerteten Leistungsbeispielen zu sammeln, wobei jede Stufe auf der Punkteskala durch mehrere Leistungsbeispiele repräsentiert werden soll. So wird gewährleistet, dass die Expertinnen und Experten mehrere Leistungsbeispiele auf die gleiche Art und Weise lesen oder anhören/ansetzen müssen wie Prüfender, die mit Noten bewerten.



Kasten 3 zur Methodik: Wichtige Fragen bei der Auswahl der Mitglieder des Expertengremiums

Frage 1: Welches Spektrum an Fachwissen muss ich ermitteln?

1. Expertinnen und Experten für das entsprechende Material, z. B. Autorinnen und Autoren von Items und Prüfende und Bewertende, Autorinnen, Autoren, Redakteurinnen und Redakteure von Lehrwerken sowie Lehrkräfte aus der Unterrichtspraxis usw.
2. Personen, die mit den Lernenden der Zielgruppe des Materials vertraut sind, üblicherweise Lehrkräfte.
3. Expertinnen und Experten für den Standard (in diesem Fall GeR).

Frage 2: Wie kann das Expertengremium eine vielfältige Sichtweise repräsentieren?

1. Es ist wichtig zu verstehen, welche Untergruppen unterschiedliche wertvolle Perspektiven auf das Material liefern.
2. Menschen aus ähnlichen Gruppen (z. B. Erfahrung, Geschlecht usw.) müssen sich nicht zwangsläufig ähnlich verhalten. Das „Erzwingen“ einer geraden Zahl von Expertinnen und Experten mit gleichem Hintergrund ist also nicht notwendig.

5.3 Herausforderungen, Lösungen und Alternativen

In diesem Abschnitt wird kurz erläutert, wie sich der Umfang der Aktivitäten zum Standard-Setting an ressourcenbeschränkte Situationen wie an ein Programm in einer einzelnen Schule oder Institution anpassen lassen. Es werden auch Alternativen zu den in Abschnitt 5.2 beschriebenen Methoden erwähnt.

5.3.1 Anpassung des Umfangs

Während Cizeks Definition des Standard-Settings das „ordnungsgemäße Befolgen eines vorgegebenen, rationalen Systems an Regeln und Verfahren“ betont, sind wir eher der Meinung, dass es nicht nur akzeptabel, sondern tatsächlich sinnvoll ist, die Methodik je nach Kontext anzupassen. Bei der Testentwicklung werden Hinweise auf die technisch-handwerkliche Qualität, z. B. Schätzungen der Reliabilität, Übereinstimmungen zwischen den Bewertenden (inter-rater agreement) usw., oft im Hinblick auf den Akzeptanzbereich von Ergebnissen diskutiert. Dieser hängt davon ab, welche Bedeutung die auf der Grundlage der Prüfung getroffenen Entscheidungen haben und ob diese Entscheidungen von den Verwendenden nachträglich angepasst werden können. Ein niedriger Reliabilitätswert könnte beispielsweise für einen von Lehrkräften entwickelten Lesetest für Schülerinnen und Schüler akzeptabel erscheinen. Die Lehrkraft hat wahrscheinlich keinen Zugriff auf ein umfangreiches System für das Schreiben und Erproben von Items, sondern kann die Auswirkungen von Entscheidungen, die auf Basis der Testergebnisse getroffen werden, überwachen, wie beispielsweise die Zuweisung von Schülerinnen und Schülern zu verschiedenen Leistungsgruppen für ein Projekt. Zudem wird die Lehrkraft fähig sein, Probleme bestimmter Schülerinnen und Schüler zu erkennen, die einer falschen Gruppe zugeordnet wurden, und kann dies rückblickend korrigieren oder diese Lernenden entsprechend unterstützen. Eine ähnliche Logik können wir auf die Methodik des Standard-Settings anwenden.

Eine der größten Herausforderungen beim Standards-Setting ist das Bedürfnis, eine ausreichende Anzahl an Expertinnen und Experten zu finden. Beim Erstellen von Tests zum Klassenfortschritt mit einer Kollegin oder einem Kollegen in der Schule fehlen wahrscheinlich die finanziellen Mittel, um ein Gremium mit 15 Mitgliedern zu rekrutieren. Unter Umständen ist es aber möglich, ein kleineres Gremium bestehend aus Kolleginnen und Kollegen an der Schule zusammenzustellen. Während man bei den Ergebnissen des Projekts zum Standard-Setting hinsichtlich der Festlegung definitiver Bestehensgrenze(n) Vorsicht walten lassen sollte, stellt der Prozess der Familiarisierung und Standardisierung sowie die Beurteilung des Standard-Settings bezüglich tatsächlicher Test-Items eine enorm wertvolle Erfahrung zur beruflichen Weiterentwicklung für alle beteiligten Personen dar.

Im kleinen Rahmen oder in internen Situationen ist es unter Umständen nicht erforderlich, für einen vollständigen Test eine endgültige Bestehensgrenze festzulegen. Es könnte eher von Interesse sein sicherzustellen, dass ein angemessener Pool mit zielgerichteten Items für praktische Übungen, Hausaufgaben, laufende Leistungsmessungen usw. bereitsteht. In diesem Fall könnte die Durchführung einer Aktivität zur Basket-Methode mit einem kleinen Kreis an Personen aus dem Kollegium sinnvoll sein. Da keine Bestehensgrenze für den Test festgelegt wurde, muss die potenzielle Verzerrung bei der Methode keine Sorgen bereiten. Wenn jedoch ein Konsens erzielt wird, geben die Beurteilungen auf Item-Ebene Aufschluss darüber, ob die Aufgaben die jeweiligen Fähigkeiten bezüglich des entsprechenden GeR-Niveaus messen oder nicht.

Die Beurteilung der eigens entwickelten Aufgaben durch die Kolleginnen und Kollegen kann bestätigen, dass man auf dem richtigen Weg ist, oder Probleme aufzeigen, die bei der Beurteilung der eigenen Testaufgaben leicht übersehen werden. Die Zusammenarbeit ermöglicht eine Realitätsprüfung dessen, was in der Schule außerhalb der eigenen Klasse geschieht. *Wird den Schülerinnen und Schülern in verschiedenen Klassenräumen mit unterschiedlichen Lehrkräften der richtige Stoff für das richtige Niveau vermittelt? Ist die Behandlung und der Schwierigkeitsgrad der Items in den Tests einheitlich?* Selbst in einer kleinen Gruppe von fünf oder sechs Kolleginnen und Kollegen können mithilfe von Techniken zum Standard-Setting potenzielle Probleme erkannt und für sich selbst und die Kolleginnen und Kollegen geprüft werden, ob die Richtung stimmt – oder gegebenenfalls Anpassungen vorgenommen werden müssen. In solchen Situationen bietet sich auch mehr Flexibilität bei der Durchführung der Aktivitäten und der Erhebung der Daten als in den in Abschnitt 5.2 beschriebenen Beispielen.

Es ist zwar nach wie vor sinnvoll, Beurteilungen auf kontrollierte und konsistente Weise zu sammeln, durch eine interne Zusammenarbeit im Kollegium ist es jedoch möglich, viele der Aktivitäten einzeln durchzuführen, zwischendurch zu besprechen und die Endergebnisse zu evaluieren.

5.3.2 Anpassungen auf vertretbare Weise vornehmen

Vor der Überlegung, welche Anpassungen sinnvollerweise vorgenommen werden sollten, ist es ratsam, sich die folgenden Fragen zu stellen:

- Was möchte ich mit dem Nachweis der Verlinkung meiner Tests mit dem GeR erreichen?
- Wer sind die wichtigsten Stakeholder, denen ich diese Verlinkung nachweisen muss?
- Welche Arten von Entscheidungen müssen auf der Basis der Verlinkung mit dem GeR getroffen werden?
- Und vor allem: Auf welche Ressourcen und Fachkenntnisse kann ich zurückgreifen, um meine Ziele zu erreichen?

Es ist nicht empfehlenswert, leichtfertig Änderungen an den bewährten Methoden vorzunehmen. Für einen ressourcenbeschränkten Unterrichtskontext sind wir jedoch der Ansicht, dass eine Anpassung der Methoden zur Gewährleistung eines gewissen Maßes an empirischen Daten zur Verlinkung einem kompletten Verzicht auf eine Standardsetzung vorzuziehen ist. Bevor Änderungen vorgenommen werden, sollte man zuerst verstehen, was geändert wird und welche Auswirkung die Modifikationen haben könnten. Außerdem sollten die Gründe für die Änderungen sowie deren Umfang klar dokumentiert werden.

5.3.3 Online-Standard-Setting

Eine wichtige Entwicklung in den letzten Jahren ist die Umstellung auf ein Online-Standard-Setting ohne Präsenzveranstaltung. Das geschah generell bei der Leistungsbeurteilung im pädagogischen Bereich (siehe z. B. Abschnitt 5.6, Quelle 3) und beim Standard-Setting im Zusammenhang mit dem GeR (siehe Quellen 16a und 16b – da diese unveröffentlicht sind, empfehlen wir, sich für weitere Einzelheiten an die Autorinnen und Autoren zu wenden). Das Online-Standard-Setting ermöglicht die Durchführung der wichtigsten Datenerhebungsaktivitäten, ohne dass die Expertinnen und Experten an einem Ort zusammenkommen müssen. So lassen sich die Kosten von Großprojekten senken, da es nicht mehr notwendig ist, einen Raum für eine mehrtägige Veranstaltung anzumieten oder für die Reise- und Unterbringungskosten der Fachleute aufzukommen. Ein weiterer Vorteil ist die Möglichkeit, den Kreis der Expertinnen und Experten zu erweitern,

indem Personen einbezogen werden, die aus praktischen oder finanziellen Gründen nicht teilnehmen könnten, wenn sie zu der Veranstaltung reisen müssten. Natürlich bringt das Online-Standard-Setting auch eigene logistische und andere Herausforderungen mit sich – wie etwa die Notwendigkeit einer stabilen Internetverbindung für alle Teilnehmenden –, die sorgfältig bedacht werden müssen. Nach der umfassenden Umstellung auf Online-Unterricht und -Prüfungen im Jahr 2020 sind viele Menschen und Organisationen in Zukunft jedoch wahrscheinlich viel besser vorbereitet und offener für ein Online-Standard-Setting.

5.3.4 Weitere Methoden zum Standard-Setting

In Abschnitt 5.6 sind Quellen aufgelistet, die einen Überblick über die Methoden bieten. Es ist wichtig, sich mit dieser Literatur auseinanderzusetzen, um beurteilen zu können, was im jeweiligen Kontext am besten funktioniert. Hier möchten wir auf einige mögliche Alternativen zu den in Abschnitt 5.2 genannten Beispielen hinweisen. Dabei betonen wir noch einmal, wie wichtig es ist, selbst zu recherchieren, flexibel zu sein und zu prüfen, was in der jeweiligen Situation machbar ist.

Es gibt viele Alternativen zu den drei Methoden, die wir in Abschnitt 5.2 kurz beschrieben haben. Zwei alternative testbezogene Methoden sind die **Bookmark-Methode** und die **Item-Descriptor-Matching-Methode** (Abschnitt 5.6, Quellen 4, 5 und 6). Die Bookmark-Methode ist heute eine der beliebtesten Methoden in den Vereinigten Staaten und wurde dort auch für die Verlinkung von Tests und Prüfungen mit dem GeR verwendet. Diese Methode nutzt statistische, auf Grundlage der Item-Response-Theory (IRT) erzeugte Schwierigkeitswerte (IRT) und gibt dem Expertengremium eine relativ einfache Aufgabenstellung zur Beurteilung.

Die Item-Descriptor-Matching-Methode wird im Zusammenhang mit dem GeR seltener verwendet (Beispiel siehe Abschnitt 5.6, Quelle 12). Die Aufgabenstellung zur Beurteilung weist Ähnlichkeiten mit der Basket-Methode auf und hat den Vorteil, dass die GeR-Deskriptoren ausdrücklich einbezogen werden. Quelle 2 in Abschnitt 5.6 beschreibt eine interessante Anpassung der Basket-Methode, die das Potenzial hat, einige der in Abschnitt 5.2 erwähnten Verzerrungen zu vermeiden und gleichzeitig die konzeptionell zugängliche Aufgabenstellung zur Beurteilung, die die Methode attraktiv macht, beizubehalten.

Bei vielen prüfungsteilnehmerzentrierten Ansätzen besteht eine der größten Herausforderungen darin, Zugang zu einer ausreichenden Anzahl von Beispielen für Prüfungsteilnehmendenleistungen zu erhalten. Das gilt insbesondere dann, wenn diese Leistungen wie bei der in Abschnitt 5.2.2 beschriebenen Methode zur analytischen Beurteilung zuvor von geschultem Personal bewertet worden sein müssen. Bei produktiven Fertigkeiten ist die Verwendung realer Leistungen von Lernenden natürlich wünschenswert. Sollte das nicht möglich sein – wenn exemplarisch Leistungsbeispiele nicht im Voraus erfasst werden können, weil die Sicherheit des Testinhalts gefährdet sein könnte –, kann die Anwendung testzentrierter Methoden eine Option sein. Die Studie zur Item-Descriptor-Matching-Methode (Abschnitt 5.6, Quelle 12) wurde beispielsweise angewendet, um die Anforderungen an eine schriftliche Aufgabe zu beurteilen.

Quelle 6 in Abschnitt 5.6 beschreibt eine Anwendung des **Contrasting-Groups-Ansatzes**, ein weiterer Ansatz, bei dem die Prüfungsteilnehmenden im Fokus stehen. Hier werden die Lehrkräfte gebeten, das GeR-Niveau von ihnen bekannten Lernenden zu evaluieren, und dann Testdaten von ihnen zu sammeln. Die Testergebnisse von Lernenden, die z. B. in die Kategorie B1 eingestuft wurden, können dann mit den Ergebnissen von Lernenden verglichen werden, die von ihren Lehrkräften in die Kategorie B2 eingestuft wurden. Anschließend können die Bestehensgrenzen ermittelt werden. Das kann ein wirksames Mittel sein, um die Ergebnisse des Standard-Settings mit anderen Methoden oder dem Kern des Standard-Settings selbst zu vergleichen. Dies ist besonders nützlich, um Lehrkräfte und andere Stakeholder in das Projekt einzubeziehen.

5.3.5 Sprachenübergreifendes Standard-Setting

Während sich die meisten veröffentlichten Studien darauf konzentrieren, einen Test oder eine Reihe von Tests für eine einzelne Sprache mit dem GeR zu verlinken, beschreiben zwei der Quellen in Abschnitt 5.6 innovative Ansätze zur Angleichung von Testaufgaben und Leistungen in **verschiedenen** Sprachen. Insbesondere Quelle 1 wurde entwickelt, um zu zeigen, dass verschiedene sprachliche und pädagogische Kontexte eine gemeinsame Interpretation des GeR erlauben. Es gibt viele lokale und institutionelle Situationen, in denen ein solcher Ansatz nützlich sein könnte. In einem Fachbereich für moderne Sprachen an einer Universität, in dem Prüfungen für verschiedene Sprachen von verschiedenen Gruppen von Lehrkräften entwickelt werden, ist es zum Beispiel wichtig zu zeigen, dass die entscheidenden Merkmale der gleichen GeR-Niveaus auf kohärente und konsistente Weise berücksichtigt werden. Quelle 1 beschreibt einen prüfungsteilnehmerorientierten Ansatz für mündliche Leistungen in Englisch, Französisch, Deutsch, Italienisch und Spanisch. Diese Studie umfasst auch zwei verschiedene Methoden der Beurteilung: „Die Bewertung der Kandidatinnen und Kandidaten sowohl im Vergleich untereinander (Ranking: normbezogen) als auch in Bezug auf den GeR (Bewertung: kriterienbasiert)“ (S. 4). Quelle 8 beschreibt einen testzentrierten Ansatz, bei dem die beiden Methoden aus Abschnitt 5.2 zum Einsatz kommen, um Aufgaben zur Fertigkeit Lesen aus verschiedenen Tests für Englisch, Französisch, Deutsch und Spanisch auf einer gemeinsamen Skala zu platzieren (da dies unveröffentlicht ist, empfehlen wir, sich für weitere Einzelheiten an die Autorinnen und Autoren zu wenden). In beiden Studien wurde die Multifacetten-Rasch-Analyse verwendet, um Vergleiche auf Aufgabenebene zu tätigen, anstatt Bestehensgrenzen für vollständige Tests festzulegen. Dadurch sind sie für die in Kapitel 4 beschriebenen Aktivitäten zum Standard-Setting, insbesondere für das Benchmarking von Leistungsbeispielen, gleichermaßen relevant.

5.4 Anpassung der Methoden zum Standard-Setting bei anderen Materialien als Prüfungen

Wie bereits erwähnt, ist das Standard-Setting für Materialien wie Curricula oder Lehr- und Lernmaterialien möglicherweise nicht das am besten geeignete Instrument für die Verlinkung mit dem GeR. Das bedeutet jedoch nicht, dass das Standard-Setting nicht relevant ist: Sie bleibt durch die Anwendung bei Prüfungen, die im Curriculum und bei dessen Vermittlung mithilfe geeigneter Materialien integriert werden sollten, Teil des Gesamtsystems der Anbindung an den GeR. Beim Sammeln von Nachweisen für die Verlinkung, insbesondere bei Curricula, sind die in den Kapiteln 3 und 4 beschriebenen Phasen der Spezifikation und Standardisierung sehr nützlich. Es ist jedoch klar, dass die Beurteilung nicht von einer einzelnen Person vorgenommen werden sollte, auch wenn diese nicht direkt dem Team angehört, das das Curriculum entwickelt hat.

Die Entscheidungen sollten stattdessen von einem Team aus Personen getroffen werden, die ein Expertengremium bilden, ähnlich wie es oben für einen Test beschrieben wurde. Die Beurteilungen, die dieses Gremium vornimmt, beziehen sich auf das GeR-Niveau bzw. die GeR-Niveaus, die bestimmte Elemente des Curriculums widerspiegeln, und können auf die gleiche Weise gesammelt und analysiert werden, wie es oben vorgeschlagen wurde. Mit anderen Worten: Es ist eine Reihe von Bewertungs- und Diskussionsrunden erforderlich, um zu der treffendsten Entscheidung zu gelangen. Diesem Gremium sollten sowohl Personen aus dem Curriculumsteam (mit Fachwissen über dieses Dokument und die Gründe für die Entscheidungen, die zu seiner Erstellung geführt haben) als auch externe Expertinnen und Experten (die mit dem GeR sehr vertraut sind) angehören.

Im Falle von Lehr- und Lernmaterialien können Aufgaben, Aktivitäten usw. zwar einige Ähnlichkeiten mit Test-Items aufweisen – sie dienen jedoch in der Regel nicht dazu, eine Punktzahl zu erzielen. Das Durcharbeiten des Lehrwerks hängt nicht vom Erreichen einer numerischen Punktzahl ab, die sich aus der Leistung bezüglich einer bestimmten Anzahl von Items und Aufgaben ergibt. Darüber hinaus können die Lehr- und Lernmaterialien einen breiteren Fokus haben, um eine flexible und kreative Nutzung zu fördern.

Dennoch kann es sinnvoll sein, einige der in Abschnitt 5.3.1 beschriebenen Ansätze anzupassen. Die Festlegung einer Bestehensgrenze auf Testebene (oder auf Materialebene) ist vielleicht nicht das Ziel – und das ist vielleicht auch nicht möglich oder wünschenswert. Es lohnt sich jedoch, empirische Belege dafür zu sammeln, dass die als Lern- oder Fortschrittskontrollen verwendeten Aufgaben tatsächlich auf das GeR-Niveau, das damit

erreicht werden soll, ausgerichtet sind. Die Art von Beurteilungen auf Item-Ebene, die mithilfe von Methoden zum Standard-Setting gewonnen werden, können eindeutig wichtige Belege für die Behauptung liefern, dass ein Lehrwerk oder ein anderes Material an den GeR angebunden ist. Die Anpassung der in Abschnitt 5.3.1 erörterten Methoden des Standard-Setting an den Kontext von Tests im Klassenzimmer können auch für Lehr- und Lernmaterialien von Bedeutung sein.

5.5 Leitfaden für Berichte über die Standard-Settingphase

Wie in den vorherigen Kapiteln erwähnt, ist es unwahrscheinlich, dass ein und derselbe Bericht für alle Beteiligten gleichermaßen relevant oder nützlich ist. Folgende Fragestellungen sind in diesem Zusammenhang relevant:

- Für wen bereiten wir die Berichte vor?
- Welchen Zweck haben die Berichte?
- Welches Maß an Fach- und Hintergrundwissen können wir bei den anvisierten Lesenden der Berichte voraussetzen?

In den Berichten werden nicht nur die eigenen Ergebnisse klar und deutlich dargestellt, sondern es besteht auch die Möglichkeit, die Gültigkeit der Empfehlungen vorzulegen. Die **Validierung**, eine der fünf Phasen der Anbindung, die in Kapitel 1 zusammengefasst wurden, findet also auch im Mikrokosmos des Standard-Settings statt. Validitätsnachweise für das Standard-Setting werden manchmal in drei Kategorien unterteilt:

1. Verfahrenorientiert: klare Dokumentation der angewandten Methode und ihrer Umsetzung; Feedback von Expertinnen und Experten, um zu zeigen, dass sie die angebotene Schulung für angemessen hielten usw.
2. Intern: technische Nachweise zu den analysierten Daten, z. B. Grad der Übereinstimmung zwischen den Expertinnen und Experten usw.
3. Extern: Beleg, dass die Ergebnisse mit anderen Nachweisquellen übereinstimmen, z. B. Nachweise aus anderen Standard-Setting-Konferenzen, Anzahl der Lernenden, die anhand der Bestehensgrenzen wahrscheinlich in die GeR-Niveaus eingestuft werden, usw.

Es sollte klar sein, dass Validierung und Berichterstellung generell nicht bis nach der Veranstaltung warten können. Wir müssen im Voraus entscheiden, zu welcher Art von Nachweisen wir Berichte erstellen müssen, und sicherstellen, dass die Sammlung dieser Nachweise in unseren (Zeit-)Plänen berücksichtigt wurde. Um beispielsweise Nachweise für die Validität des Verfahrens zu liefern, müssen wir einen Fragebogen entwerfen, um Feedback von den Expertinnen und Experten einzuholen. Dabei gilt es zu beachten, dass sie genügend Zeit haben, um den Fragebogen auszufüllen. Im Nachhinein können wir nur selten Daten erheben, die wir übersehen haben, oder Daten neu erheben, die möglicherweise nicht richtig gespeichert wurden.

Wie in Kapitel 4 erwähnt, ist die vollständige Dokumentation aller Aspekte unseres Standard-Settings – von der Planung bis zu den Ergebnissen – ebenfalls ein wesentlicher Bestandteil des institutionellen Gedächtnisses. Während des Prozesses des Standard-Settings werden nicht nur die Grundlagen für die künftige Berichterstellung geschaffen, sondern es kommt auch immer wieder vor, dass wir aufgrund praktischer Probleme die ursprünglichen Pläne anpassen müssen. Wenn wir nicht dokumentieren, um welche Änderungen und welchen Grund dafür es sich handelt, ist es später nur schwer möglich, zu klären, was während der Veranstaltung geschehen ist, und dies gegenüber externen Beteiligten zu rechtfertigen.

Als hervorragendes Beispiel für eine umfassende Studie aus allen Perspektiven empfehlen wir Quelle 15. Die Quellen 11 und 17 bieten einen wertvollen Überblick über Verlinkungsstudien, die nach der Veröffentlichung der Pilotversion des *Manual for Relating Examinations to the CEFR* im Jahr 2003 durchgeführt wurden.

5.6 Weiterführende Literatur

Viele der in den vorangegangenen Kapiteln hervorgehobenen Quellen sind natürlich auch für das Standard-Setting von Bedeutung, z. B. die Links zur GeR-Deskriptorenbank (Excel) sowie zum GeR 2001 und zum Begleitband 2020. In diesem Abschnitt werden einige Quellen aufgelistet, die dabei helfen, sich näher mit der Methodik des Standard-Settings zu befassen. Es wird kein Anspruch auf Vollständigkeit erhoben, die Übersicht bietet aber einen guten Einstieg in das Thema und liefert weitere Informationen zu den spezifischen Beispielen, die in diesem Kapitel verwendet wurden. Wie bereits erwähnt, ist ein gewisses Maß an Vertrautheit mit der Literatur erforderlich, um fundierte Entscheidungen treffen zu können – unabhängig davon, ob das Standard-Setting selbst durchgeführt wird oder ein Expertengremium damit beauftragt wird.

- 1 Breton, G., Lepage, S. & North, B. (2008). [Cross-language benchmarking seminar to calibrate examples of spoken production in English, French, German, Italian and Spanish with regard to the six levels of the Common European Framework of Reference for Languages \(CEFR\)](#). CIEP, Sèvres, 23–25 Juni 2008. Straßburg: CIEP/Europarat.
- 2 Brunfaut, T. & Harding, L. (2014). [Linking the GEPT listening test to the Common European Framework of Reference](#). LTTC-GEPT Research Reports RG-05. Taipeh: Language Training and Testing Centre.
- 3 Bunch, M.B. (2020). [Digital Module 14: Planning and Conducting Standard Setting](#). ITEMS Module, NCME.
- 4 Cizek, G. (Hrsg.) (2012). *Setting performance standards*. New York: Routledge.
- 5 Cizek, G. & Bunch, M. (2007). *Standard Setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks: Sage Publications.
- 6 Europarat (2009). [Relating language examinations to the Common European Framework of References for Languages: Learning teaching, assessment](#). Straßburg: Abteilung für Sprachrichtlinien.
- 7 Dunlea, J., & Figueras, N. (2012). Replicating results from a CEFR test comparison project across continents. In D. Tsagari und I. Csepes (Hrsg.), *Collaboration in language testing and assessment* (S. 31–45). New York: Peter Lang.
- 8 Dunlea, J., Spiby, R., Malone, M. & Folny, V. (2021). *An innovative approach to evaluating claims of alignment to the CEFR for tests of different languages*. Im Rahmen der 17. EALTA-Onlinekonferenz präsentierte Publikation.
- 9 Dunlea, J., Motteram, J., Spiby, R., Jiang, R., Spoettl S., Zehentner, M., Kremmel B. & Holzknicht, F. (im Druck). Aligning Tests of Language Proficiency and Numeracy to the Workplace Literacy and Numeracy Framework. *British Council Validation Series, VS/2021/001*. London: British Council.
- 10 Dunlea, J., Spiby, R., Wu, S., Zhang, J. & Cheng, M. (2019). [China's Standards of English Language Ability \(CSE\): Linking UK Exams to the CSE](#). *British Council Validation Series, VS/2018/001*. London: British Council.
- 11 Figueras, N. & Noijons, J. (Hrsg.) (2009). [Linking to the CEFR levels: Research perspectives](#). Arnhem: Cito-EALTA.
- 12 Harsch, C. & Kanistra, V.P. (2020). Using an Innovative Standard-setting Approach to Align Integrated and Independent Writing Tasks to the CEFR. *Language Assessment Quarterly*, 17:3, 262–281, DOI: 10.1080/15434303.2020.1754828
- 13 Kaftandjieva, F. (2004). *Standard setting*. In Council of Europe, Reference supplement to the pilot version of the manual for relating language examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment (CEF). Straßburg: Abteilung für Sprachrichtlinien.

- 14 Kaftandjieva, F. (2010). [Methods for setting cut-off scores in criterion-referenced achievement tests: A comparative analysis of six recent methods with an application to tests of reading](#). Arnhem: CITO und EALTA.
- 15 Kantarcioğlu, E. (2012). [Relating an Institutional Proficiency Examination to the CEFR: a case study](#). Unveröffentlichte Dissertation, Universität Roehampton, London.
- 16a Kollias, C. (2017). Virtual standard setting in language testing: Exploring the feasibility and appropriateness of setting cut scores in synchronous audio and audio-visual environments. Unveröffentlichte Dissertation. Universität Lancaster, Lancaster.
- 16b Kollias, C. (2021). *Variations in setting cut scores: How comparable are cut scores across media, methods, time, and instrument length?* Im Rahmen der 17. EALTA-Onlinekonferenz präsentierte Publikation.
- 17 Martyniuk, W. (Hrsg.) (2010). Aligning tests with the CEFR: Reflections on using the Council of Europe's draft manual. *Studies in Language Testing* 33. Cambridge: Cambridge University Press.
- 18 Verhelst, N., Fugueras, N., Prokhorova, E., Takala, S. & Timofeeva, T. (2019). Standard setting for writing and speaking: The Saint Petersburg experience. In Ari Huhta, Gudrun Erickson & Neus Figueras (Hrsg.), *Developments in Language Education : A Memorial Volume in Honour of Sauli Takala* (S. 278–301). Universität Jyväskylä, Zentrum für angewandte Sprachstudien & EALTA. <https://jyx.jyu.fi/bitstream/handle/123456789/65608/978-951-39-7748-1.pdf?sequence=1&isAllowed=y>



Hinweise zur eigenen Implementierung ...

Das Standard-Setting erfordert technisches Know-how und zahlreiche Ressourcen. Dieses Kapitel bietet keine detaillierte Unterstützung bei der Planung und Durchführung eines Projekts zum Standard-Setting. Es ist jedoch eine hilfreiche Quelle, um herauszufinden, was Sie wissen müssen und wo Sie diese Informationen finden.

Die folgenden Punkte sind dabei sehr wichtig:

- Es ist erforderlich, sich mit technischer und wissenschaftlicher Literatur zu befassen.
- Wie in allen Kapiteln erwähnt, sollten Sie sich mit dieser Literatur kritisch auseinandersetzen und die Methoden an Ihr Material, die Projektanforderung und Ihre Ziele bezüglich der Anbindung an den GeR anpassen.
- Sie sollten alle Schritte im Prozess dokumentieren: der Grund für die Auswahl einer Methode, die entsprechend dem Kontext vorgenommenen Anpassungen, die Vorbereitung und Datenerhebung selbst sowie die Analyse, Ergebnisse und Validierungsnachweise.
- Nehmen Sie sich Zeit, die vorhandenen und benötigten Ressourcen sorgfältig zu prüfen, um das Standard-Setting so durchzuführen, wie es für Ihren Kontext angemessen ist.

Auch Folgendes sollte berücksichtigt werden:

- Die Planung der Logistik und praktischen Aspekte der Standard-Setting-Konferenzen ist genauso wichtig wie die Auswahl der richtigen Methode sowie das entsprechende technische Know-how bezüglich des Projekts.
- Während das Standard-Setting ein empirischer Prozess ist, der statistische Analysen der quantitativen Daten erfordert, geht es grundlegend auch um menschliche Interaktion(en):
 - Um die Expertinnen und Experten sowie die Gruppendynamik während der Standard-Setting-Konferenzen
 - Um Sie und Ihr Team bei der Durchführung der Schulungen und der Ermöglichung der Bewertungen
 - Um Prüfungsteilnehmenden und Stakeholder, die betroffen sind von den Entscheidungen, die basierend auf den von Ihnen gesammelten Nachweisen getroffen werden



KAPITEL 6: VALIDIERUNG DER ARGUMENTATION

Kapitel 6 befasst sich mit der Validierung der Anbindung an den GeR, die während des Prozesses herausgearbeitet und nachgewiesen wurde, außerdem mit der Präsentation des Validierungsberichts für die Stakeholder/-innen.

Dieses Kapitel ist folgendermaßen aufgeteilt:

- 6.1 Definition von Validierung
- 6.2 Wesentliche Einschränkungen der Validierungsargumente zur Anbindung
- 6.3 Aufbau einer Evidenzbasis für die Validierung
- 6.4 Hinweis zur Struktur des Berichts
- 6.5 Leitfaden für Berichte über die Validierungsphase
- 6.6 Hinweis zur Anbindung von Tests, die KI-unterstützte Bewertungssysteme verwenden

Am Ende des Kapitels sollten Sie eine klare Vorstellung davon haben,

- Was mit Validierung in diesem Kontext gemeint ist,
- Wie wichtig es ist, eine evidenzbasierte Basis für die Verlinkungsbehauptung zum GeR zu schaffen,
- Was die Inhalte des Validierungsberichts sind und wie dieser strukturiert ist.

6.1 Definition von Validierung

Das allgemein akzeptierte Verständnis von Validierung hinsichtlich der Leistungsmessung im Bildungsbereich besteht darin, dass die Validierung durch den Aufbau einer klar strukturierten Argumentation erfolgt, um spezifische Schlussfolgerungen innerhalb eines bestimmten Anwendungskontextes besser treffen zu können – daher der „Argument Based Approach“ nach Kane (1992, 2006, 2010)⁷.

Da dieses Handbuch nicht dafür gedacht ist, Lesenden einen detaillierten Überblick darüber zu geben, wie eine Prüfung validiert werden kann, betrachten wir die Validierung aus einer etwas anderen Perspektive. In diesem Handbuch erklären wir eine Reihe von Schritten, durch die Entwicklerinnen und Entwickler eines bestimmten Zielmaterials (Curriculum, Materialien, Ausbildungsprogramm für Lehrkräfte oder Test) einen Punkt erreichen können, an dem sie in der Lage sind, eine begründete Argumentation zu liefern, die nachweist, dass diese Aktivitäten auf eine bestimmte Weise mit dem GeR verbunden sind. Wenn wir von Validierung sprechen, beziehen wir uns daher auf das Sammeln und Präsentieren von Belegen, die unsere Argumentation einer Verbindung zwischen den anvisierten Aktivitäten und dem GeR unterstützen.

Bevor wir uns damit befassen, wie die oben genannten Nachweise aussehen und wie sie präsentiert werden können, sollten wir zunächst zwei maßgebliche Einschränkungen bei der Validierung unserer Argumentation berücksichtigen.

6.2 Wesentliche Einschränkungen der Validierungsargumente zur Anbindung

Wir können nicht sinnvoll argumentieren, dass ein Ziel per se gültig ist. Wie in Abschnitt 1.4 erklärt wurde, kann man nicht davon ausgehen, dass eine Reihe von Unterrichtsmaterialien, die für einen bestimmten Kontext (mit klarer Verbindung zu einem bestimmten Curriculum) entwickelt wurden, automatisch auch in einem anderen System funktionieren. Das bedeutet, dass die anvisierte Aktivität nur dann valide ist, wenn sie für einen bestimmten Zweck in einem bestimmten Kontext verwendet wird – normalerweise der Zweck und der Kontext, für den sie entwickelt wurde. Jede Änderung des Zwecks oder des Kontextes bedeutet, dass neue Nachweise vorgelegt werden müssen, um die Argumentation zur Validierung zu stützen. Die Validierung steht also in direktem Zusammenhang mit der Nutzung der anvisierten Aktivität in einem bestimmten Kontext.

Darüber hinaus ist die Vermutung nicht sinnvoll, dass eine Zielmaterial nur entweder valide oder nicht valide sein kann. Da es schwierig ist, sich eine Situation vorzustellen, in der die gesammelten Nachweise absolut eindeutig sind, ist es besser, sich die Validität als einen Prozess vorzustellen, bei dem es darum geht, eine evidenzbasierte Grundlage zu schaffen, mithilfe derer versucht wird, überzeugend darzulegen, dass die Nachweise, die die Argumentation einer Verlinkung mit dem GeR stützen, fundiert, präzise und zuverlässig sind.

⁷ Kane, M. T. (1992). An Argument-Based Approach to Validity. *Psychological Bulletin*, 112, 527.

Kane, M. T. (2006). Validation. In R. Brennan (Hrsg.), *Educational measurement*, 4. Ausgabe. (S. 17–64), Westport, CT: American Council on Education und Praeger.

Kane, M. T. (2010). *Validity, Fairness, and Testing*. Präsentation im Rahmen der Einladungskonferenz: Educational assessment, accountability, and equity: Conversations on validity around the World. New York: Columbia University.

https://www.ets.org/c/18486/pdf/19633_ets_tc_validityconference_Kane%202012_03_28.pdf

6.3 Aufbau einer Evidenzbasis für die Validierung

Wenn die verschiedenen Phasen des Prozesses, wie in diesem Handbuch beschrieben, sorgfältig dokumentiert sind, können die Zuständigen für Entwicklung und Validierung die oben erwähnten Belege kompilieren. Abbildung 6.1 macht deutlich, dass die Validierung des Prozesses im Wesentlichen darin besteht, die in den verschiedenen Phasen gesammelten Belege zu einem kohärenten und überzeugenden Argument zur Unterstützung der Verlinkung zusammenzustellen.

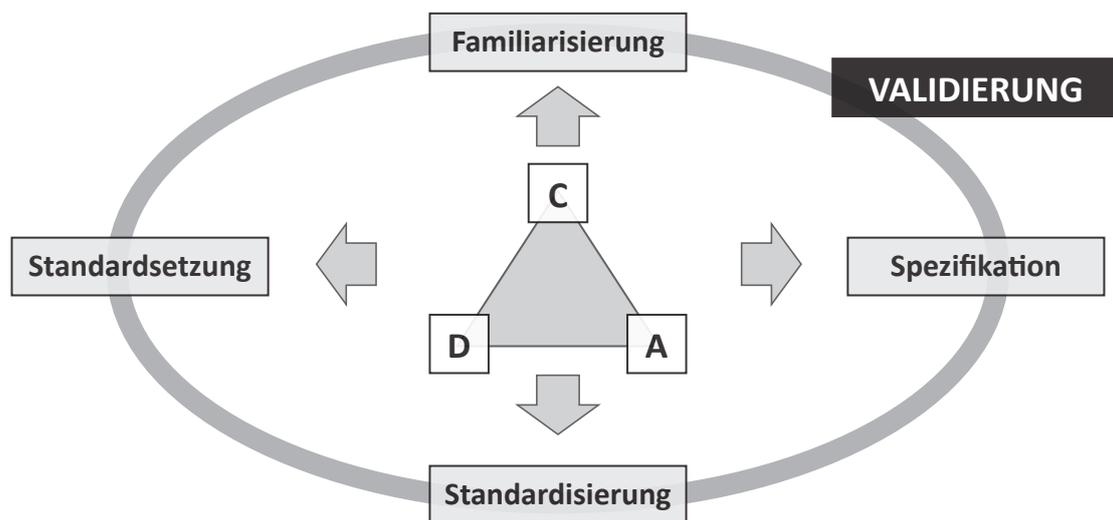


Abbildung 6.1: Übersicht zur Validierung der Verlinkungsbehauptung
[C – Curriculum, D – Durchführung, A – Prüfung (engl. Assessment)]

Dieser Prozess soll nicht als eine Alles-oder-nichts-Aktivität betrachtet werden. Wie in Abbildung 6.2 zu sehen, ist je nach Qualität der in den verschiedenen Phasen gesammelten Belege die Argumentation einer Verlinkung mit dem GeR unterschiedlich stark ausgeprägt. Auch hier ist keine Dichotomie im Spiel – selbst der stärkste Anspruch bedeutet nicht, dass die Argumentation korrekt ist. Und sie bedeutet auch nicht, dass die Verlinkung belegt ist. Es liegt auf der Hand, dass mehr und aussagekräftigere Belege zu einer stärkeren Argumentation führen, die schließlich überzeugender ist.

In dieser Abbildung stellen die dunkler schattierten Pfeile den idealen Verlauf des Anbindungsprozesses dar. Die helleren Pfeile zeigen ein ungünstiges Ergebnis, bei dem der Nachweis für eine Verlinkungsbehauptung, der gerechtfertigt werden soll, entweder schwach ist oder nicht existiert. So kann beispielsweise nach der Familiarisierung kein Anspruch formuliert werden; ebenso wenig kann nach der Spezifikation (den helleren Pfeilen folgend) kein Anspruch geltend gemacht werden, wenn die Belege schwach sind oder ganz fehlen. Wenn die Spezifikation erfolgreich ist und wir zur Standardisierung übergehen (gekennzeichnet durch den dunklen Pfeil), nur um festzustellen, dass die Belege hier schwach sind oder fehlen, wird erwartet, dass wir die schwächsten Verlinkungsbehauptungen aufstellen und den Prozess beenden. Sollte der Nachweis stark genug sein, fahren wir mit dem Standard-Setting fort.

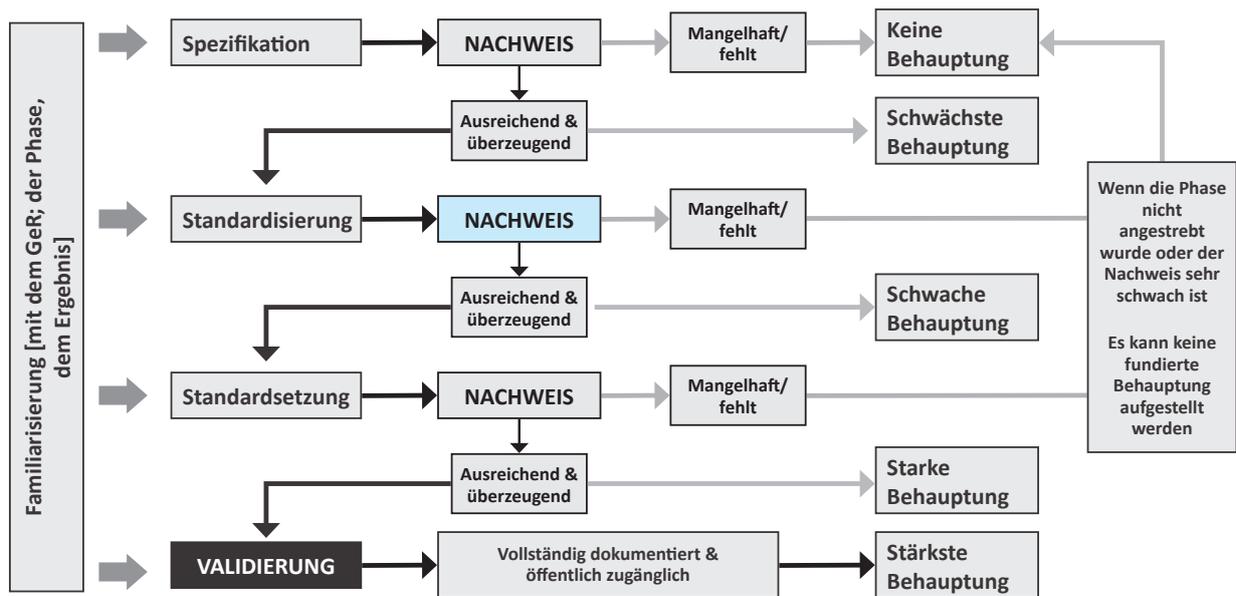


Abbildung 6.2: Hin zu einer starken Behauptung

6.4 Hinweis zur Struktur des Berichts

Wie bereits erläutert, sollte die Validierung des Anbindungsprozesses nicht im vollen Umfang die recht starr strukturierten Argumente widerspiegeln, die aktuell bei Leistungsmessungen im Bildungsbereich vorherrschen. Auch wenn ein gewisses Maß an Formalität erwartet wird, sollte stets bedacht werden, dass sich das Argument der Validierung bei Leistungsmessungen im Bildungsbereich in der Regel auf eine einzige, eng definierte Gruppe von Stakeholder/-innen konzentriert – in der Regel akademische Theoretiker/-innen. Unsere Vorstellung von einem Validierungsargument unterscheidet sich insofern stark, als dass wir wissen, dass es wahrscheinlich eine ganze Reihe von Stakeholderinnen und Stakeholdern gibt, die überzeugt werden müssen. Schließlich besteht der einzige Grund für die Präsentation eines Arguments darin, Lesende von der Richtigkeit einer bestimmten Behauptung zu überzeugen. Die verschiedenen Stakeholdergruppen werden sehr unterschiedliche Erwartungen an die Form dieses Arguments haben. Akademiker/-innen haben unter Umständen ein Interesse an einem bestimmten theoretischen Ansatz zur Argumentation und erwarten eine komplexe Struktur – wie beispielsweise die von Kane (2010) vorgeschlagene Anpassung der Toulmin-Struktur⁸. Lehrkräfte oder Autorinnen und Autoren von Materialien haben andererseits wahrscheinlich nicht die Zeit, sich mit einer langen und komplexen Argumentation zu befassen, und rechnen daher vielleicht mit einem praktischeren und gezielteren Überblick. Eltern oder Lernende wiederum finden es gut, wenn die Belege in einem leicht zugänglichen Format präsentiert werden, z. B. in Form eines Onlineblogs oder sogar in den sozialen Medien.

⁸ Das grundlegende Toulmin-Argument besteht aus den Argumenten (*grounds*), der Behauptung (*claim*) und der Schlussfolgerung (*warrant*). Es ist möglich, bei Bedarf drei weitere hinzuzufügen: Stützung (*backing*), Operator (*qualifier*) und Ausnahmebedingung (*rebuttal*). Mit diesen Elementen lässt sich eine strukturierte Argumentation bilden.

6.5 Leitfaden für Berichte über die Validierungsphase

Bezüglich der Erstellung des Abschlussberichts halten wir einen relativ unkomplizierten Ansatz für angemessen. Dies bedeutet im Wesentlichen, dass durch folgende Maßnahmen eine Reihe von Nachweisen erbracht werden: das Ausfüllen der in diesem Handbuch enthaltenen Bögen; eine Reflexion des Prozesses und der Qualität und Quantität der Nachweise in jeder Phase; das Ziehen von evidenzbasierten Schlussfolgerungen hinsichtlich der Stärke der Argumentation, die anhand der Nachweise Ihrer Meinung nach möglich sind.

Am Ende jeder Phase können auch eigene Argumente für den Übergang zur nächsten Phase präsentiert werden. Wir gehen davon aus, dass in den Fällen, in denen die Nachweise als angemessen für eine begründete Behauptung angesehen werden, automatisch die nächste Stufe erreicht wird. Um gegensätzliche Ansprüche (z. B. von externen Stellen oder der Konkurrenz) abzuwehren, empfiehlt es sich, die Dokumentation (Bögen, Überlegungen und Argumentation) von unabhängiger Seite überprüfen zu lassen.

Wenn wir alle in diesem Handbuch beschriebenen Phasen des Prozesses abgeschlossen haben, ist es nun an der Zeit, den Validierungsbericht zu erstellen. Wie im vorherigen Abschnitt erwähnt, ist das Spektrum von Stakeholderinnen und Stakeholdern breit gefächert, sodass Sie darüber nachdenken sollten, auf welche Weise, in welchem Umfang und mit welchen Nachweisen Sie ihnen die Erkenntnisse vermitteln. Entscheidend ist jedoch, dass die zugrunde liegende Botschaft sowohl konsistent als auch präzise ist. Aus diesem Grund wäre es hilfreich, sich bei der Erstellung des Validierungsberichts Unterstützung durch Kommunikationsexpertinnen oder -experten zu holen. Tabelle 6.1 zeigt eine grobe Gliederung zur Erstellung des Berichts.

Tabelle 6.1: Vorschlag für die Herangehensweise beim Verfassen des Validierungsberichts

Phase	Aktivität	Anmerkungen
1	Zweck	Skizzierung des wichtigsten Zwecks des Validierungsarguments. Bestimmung der Zielgruppe und der zu vermittelnden Punkte. Das ist vergleichbar mit dem Verfassen einer Theorie des Wandels (<i>Theory of Change</i>) für Ihr Anbindungsprojekt.
2	Struktur	Festlegung der endgültigen Struktur der verschiedenen Argumente, die zusammengestellt werden sollen. Dabei ist zu bedenken, dass für unterschiedliche Stakeholder/-innen unterschiedliche Ansätze erforderlich sind.
3	Nachweise	Beschreibung der angewandten Verfahren für das Sammeln der Nachweise. Hinweis: Der Schwerpunkt dieses Kapitels liegt auf dem Nachweis der <i>internen Validierung</i> . Das bedeutet also eine Ausrichtung auf das Endprodukt. Bei Prüfungen ist es auch ratsam, <i>externe</i> Nachweise zu sammeln. Das können zum Beispiel Einschätzungen von Lehrkräften sein, die mit den auf dem Verlinkeungsprojekt basierenden Einstufungen verglichen werden.
4	Verfassen	Aufbau der Argumentation auf der Grundlage aller verfügbaren Nachweise. Dies werden die Kapitel in diesem Handbuch widerspiegeln, wobei die hier vorgestellten Tabellen zusammen mit den Vorschlägen für die Kapitel, in denen keine Tabellen verfügbar sind, verwendet werden.
5	Behauptung	Am Ende jedes Abschnitts (also entsprechend den Kapiteln in diesem Handbuch) müssen die Nachweise zusammengefügt werden, die das Niveau und die Stärke der angeführten Argumentation untermauern. Anschließend ist zu begründen, warum zur nächsten Phase übergegangen wird. Im letzten Abschnitt können alle Argumente zusammengefügt werden, um eine finale Behauptung in Bezug auf Niveau und Stärke aufzustellen.

6	Kommunikation	Stellen Sie den Gruppen der Stakeholder/-innen das finale Argument / den finalen Bericht zur Verfügung. (Hinweis: Es kann verschiedene Berichte für unterschiedliche Gruppen geben.)
----------	----------------------	---

Tabelle 6.2 bietet eine kurze Übersicht über die spezifischen Details, die im Bericht berücksichtigt werden müssen. Diese Tabelle soll nur Vorschläge dafür liefern, welche Nachweise in die Argumentation einfließen, nicht aber, wie die Argumentation selbst aufgebaut ist.

Tabelle 6.2: Übersicht über die Phasen und möglichen Behauptungen

Phase	Erwartete Validierungsnachweise		
Familiarisierung	Ermittlung von Teilnehmenden und Dokumentation des Auswahlprozesses Ausfüllen der Dokumentationsbögen ausfüllen (Kapitel 2) Reflektierende Beurteilung und Untermauerung der Behauptung und/oder Übergang zur nächsten Phase		
Spezifikation	Ermittlung von Teilnehmenden und Dokumentation des Auswahlprozesses (sofern es Änderungen gab)		
	Curriculum: Ausfüllen der Bögen (Kapitel 3); entsprechende Anpassungen (mit Begründung)	Lehrwerk/Material: Ausfüllen der Bögen (Kapitel 3); entsprechende Anpassungen (mit Begründung)	Test/Prüfung. Ausfüllen der Bögen (Kapitel 3)
	Reflektierende Beurteilung und Untermauerung/Verwerfen der Behauptung und/oder Übergang zur nächsten Phase		
Standardisierung	Ermittlung von Teilnehmenden und Dokumentation des Auswahlprozesses Dokumentation des verwendeten Materials und Begründung der Auswahl (Ausfüllen der Bögen in Kapitel 4) Erstellung des Berichts zu allen durchgeführten Aktivitäten Reflektierende Beurteilung und Untermauerung/Verwerfen der Behauptung und/oder Übergang zur nächsten Phase		
Standard-Setting	Auswahl der zu verwendenden Methode Ermittlung der erforderlichen Ressourcen (physisch, finanziell und personell) Ermittlung der relevanten Expertinnen/Experten für das Gremium (Rekrutierungsprozess dokumentieren) Auswahl und Dokumentation der zu verwendenden Testmaterialien Ermittlung und Dokumentation der zu verwendenden Testdaten für die Veranstaltung Vorbereitung der gesamten Kommunikation (insbesondere für das Expertengremium, aber auch Festlegung einer Strategie für die Verbreitung) Ermittlung und Vorbereitung aller Schulungsmaterialien (Verweis auf Familiarisierung) Vorbereitung aller Bögen/Materialien zur Datenerhebung Buchung und Vorbereitung eines Raums Erstellung von Berichten zu allen durchgeführten Aktivitäten– einschließlich der Ergebnisse aller statistischer Analysen Reflektierende Beurteilung und Untermauerung/Verwerfen der Behauptung und/oder Übergang zur nächsten Phase		

Validierung	Definition der Zielgruppe der Stakeholderinnen und Stakeholder Skizzierung des wichtigsten Zwecks des Validierungsarguments Beschreibung der Struktur der Argumentation Beschreibung des Prozesses des Sammelns von Nachweisen (Hinweis: In diesem Kapitel lag der Fokus auf internen, also materialbezogenen Nachweisen. Bei Prüfungen ist es auch ratsam, externe Belege zu sammeln. Das können zum Beispiel Einschätzungen von Lehrkräften sein, die die mit den auf dem Verlinkungsprojekt basierenden Einstufungen verglichen werden.) Aufbau der auf allen verfügbaren Nachweisen basierenden Argumentation Benennen Sie das Level und den Grad der finalen Behauptung (dies kann in jeder Phase geschehen, wodurch am Ende des Prozesses der finale Behauptung entsteht) Bereitstellung der finalen Argumentation / des finalen Berichts für die Zielgruppe der Stakeholder/-innen (Hinweis: Es kann verschiedene Berichte für unterschiedliche Gruppen geben)
--------------------	---

Viele der in diesem Handbuch beschriebenen Prozesse basieren entweder auf dem früheren *Manual for Relating Examinations to the CEFR* (2009) des Europarats oder spiegeln Erfahrungen im Umgang mit Leistungsmessung im Bildungsbereich wider. Wir denken, dass die in den Kapiteln 2 bis 4 enthaltenen Dokumentationsbogen in den drei von uns beschriebenen Bereichen (Curriculum, Durchführung und Testen/Prüfen) wahrscheinlich übereinstimmen. Der Dokumentationsbogen für Kapitel 5 unterscheidet sich von den anderen, da es nicht als Teil des Validierungsberichts gedacht ist, sondern vielmehr als Checkliste für die Planung die Prozessphase des Standard-Settings. Natürlich können die dort hervorgehobenen Überschriften verwendet werden, um über die Inhalte des Berichts über die Phase des Standard-Settings zu informieren. Es ist zu erwarten, dass sich Berichte über diese Phase je nach Material, das verlinkt wird, deutlich unterscheiden. Während sich das Kapitel hauptsächlich auf die Anbindung von Tests an den GeR konzentriert, wiederholen sich zahlreiche Schritte bei anderen Materialien, auch wenn sich die Details ändern (Übersicht siehe Tabelle 6.3.).

Die Auswahl der zu verwendenden Methode unterscheidet sich darin, dass das Fehlen von Daten zu Curricula oder Lehr-/Lernmaterialien dazu führt, dass einzig die tatsächlichen Inhalte des Materials direkt mit den relevanten GeR-Deskriptorenskalen verglichen werden können. Das bedeutet, dass sich diese Veranstaltungen zum Standard-Setting von denen für Prüfungen insofern unterscheiden, als dass es keine Mess- oder Ergebnisdaten gibt, mit denen man die Beurteilungen der Expertinnen und Experten vergleichen könnte. Der Prozess der Urteilsfindung und die damit verbundene Analyse sowie die Diskussionen im Anschluss an die Beurteilungsrunden werden jedoch ähnlich ablaufen wie bei Prüfungen.

Tabelle 6.3: Übersicht über die Schritte, die für die verschiedenen Arten von Materialien erforderlich sind

	Leistungsmessung	Curriculum	Material/Lehrwerk
Auswahl der Methode	Auswahl (bezogen auf Prüfung oder Prüflingen)	Nur auf Material bezogen	Nur auf Material bezogen
Ermittlung der benötigten Ressourcen	Physisch, finanziell und personell	Physisch, finanziell und personell	Physisch, finanziell und personell
Ermittlung von relevante Expertinnen und Experten	Erforderliche Anzahl Rekrutierung	Erforderliche Anzahl Rekrutierung	Erforderliche Anzahl Rekrutierung
Bereitstellung Zugang zu Testmaterialien	Materialien ermitteln und vorbereiten	Materialien ermitteln und vorbereiten	Materialien ermitteln und vorbereiten

Bereitstellung Zugriff auf Daten	Testdaten ermitteln und vorbereiten	–	–
Vorbereitung Kommunikation	Einladung, Einverständnis, Anleitung für die Teilnahme	Einladung, Einverständnis, Anleitung für die Teilnahme	Einladung, Einverständnis, Anleitung für die Teilnahme
GeR-Schulungsmaterial	Ermittlung oder Erstellung und Vorbereitung	Ermittlung oder Erstellung und Vorbereitung	Ermittlung oder Erstellung und Vorbereitung
Materialien zur Datenerhebung	Beurteilungs- und Feedbackformulare für Expertengremium Plan zur Datenanalyse	Beurteilungs- und Feedbackformulare für Expertengremium Plan zur Datenanalyse	Beurteilungs- und Feedbackformulare für Expertengremium Plan zur Datenanalyse
Räumlichkeiten vorbereiten	Veranstaltungsort finden, buchen und vorbereiten	Veranstaltungsort finden, buchen und vorbereiten	Veranstaltungsort finden, buchen und vorbereiten

Schließlich enthält die Originalversion des Manual for Relating Examinations to the CEFR (Europarat, 2009) eine ausführliche Erläuterung interner und externer Validierungsnachweise. Wie in Tabelle 6.1 beschrieben, liegt der Schwerpunkt der in diesem Handbuch beschriebenen Aktivitäten auf dem Sammeln interner Nachweise. Mit anderen Worten: Wir suchen nach Nachweisen für die Behauptungen, die auf der Dokumentation des Prozesses und der damit verbundenen Reflexion basieren. Die unabhängige Überprüfung in jeder Phase bietet ein gewisses Maß an externer Evidenz und stützt die in diesen Phasen getätigten Aussagen. Während des Standard-Settings variieren die Arten der externen Nachweise je nach dem betreffenden Element:

- Curriculum – Überprüfung von Prozessen, Nachweisen und Behauptungen durch eine externe Gruppe, die sich aus mindestens zwei unabhängigen Expertinnen/Experten zusammensetzt.
- Lehr-/Lernmaterial – ähnlicher Ansatz.
- Test – ähnlicher Ansatz, wir würden aber ergänzend die Erhebung zusätzlicher datengestützter Nachweise empfehlen. Das kann durch Daten aus einer ähnlichen Leistungsmessung geschehen, um die Ergebnisse des Expertengremiums zu untermauern, oder durch die Verwendung von Einschätzungen der Lehrkräfte bezüglich des GeR-Niveaus ihrer eigenen Lernenden im Vergleich zu den GeR-Niveaus basierend auf den Leistungen bei der untersuchten Leistungsmessung.

Die in diesem Handbuch dargestellten Aktivitäten beziehen sich alle auf die Einschätzung, inwiefern das Zielmaterial in Bezug auf Gestaltung und Entwicklung mit dem GeR übereinstimmt. Bei einer Podiumsdiskussion auf der Konferenz der Chinese Association for Language Testing and Assessment (CALTA) im Oktober 2021 erinnerte uns Fan Jia von der Southwest Jiaotong University daran, dass dies nicht unbedingt ein überzeugender Beweis dafür ist, dass das Material tatsächlich auf dem erwarteten Niveau funktioniert. Dies gilt insbesondere für die Einschätzung der Übereinstimmung eines Curriculums oder Lehrwerks mit dem GeR, bei der es vor allem darauf ankommt, wie das Material in der Praxis interpretiert und verwendet wird (die standardisierte Natur von Tests bedeutet, dass sie höchstwahrscheinlich einheitlich administriert werden, wodurch derartige Interpretationsprobleme vermieden werden). Natürlich kann diese Untersuchung nicht während der Entwicklungsphase eines Curriculums oder Lehrwerks durchgeführt werden. Sie sollte aber die Grundlage für ein systematisches Monitoring nach der Einführung bilden.

Darüber hinaus sollte beachtet werden, dass die meisten schulgebundenen Lehrwerke Personen nur für einen relativ begrenzten Zeitraum beim Lernen unterstützen sollen. In Anbetracht der Tatsache, dass es lange dauert, eine Sensibilität für das Kasussystem zu entwickeln und die Verwendung von Kasus und Kasusendungen dann in gewissem Maß zu beherrschen, sollte ein Lehrbuchverlag, der beispielsweise ein A1-Lehrbuch für Deutsch herausgibt, deutlich machen, dass dieses spezielle Lehrbuch nur eine Stufe in der längerfristigen Entwicklung dieses spezifischen Aspekts der Sprachbeherrschung abdeckt.

6.6 Hinweis zur Anbindung von Tests, die KI-unterstützte Bewertungssysteme verwenden

Der zunehmende Einsatz von automatischen KI-gestützten Bewertungssystemen bei Tests zum Sprechen und Schreiben muss in Bezug auf die Validierungstheorie und -praxis noch angemessen behandelt werden (O’Sullivan, Breakspear & Bayliss, im Druck).

Da die mutmaßliche Anbindung an den GeR (oder einen anderen Referenzwerk oder Standard für Sprachen) im Wesentlichen eine Frage der Validierung ist, sollte unserer Meinung nach in diesem Handbuch darauf eingegangen werden. Der Vorschlag von O’Sullivan, Breakspear und Bayliss lautet, dass Testentwicklerinnen und -entwickler, die sich bei der Bewertung auf KI-Modelle stützen, Belege sowohl über die ursprüngliche Modellentwicklung als auch über die spezifische Validierung des Modells für die Verwendung in ihrem Test vorlegen sollten – und zwar unabhängig davon, ob es sich um ein vollständig maschinelles oder ein maschinengestütztes System handelt. Wenn wir es mit einem „Black Box“-System zu tun haben, bei dem der zur Bewertung implementierte Algorithmus nicht transparent ist und im Wesentlichen weder den Entwickler/-innen noch den Nutzerinnen und Nutzern des Tests bekannt ist, stellt sich das für jede Verlinkungsbehauptung als Problem dar. Das kann durch die Veröffentlichung eines Datenblatts (Gebru et al., 2021)⁹, einer Modellkarte (Mitchell et al., 2019)¹⁰ und einer Studie mit gemischten Methoden (quantitativ und qualitativ), die speziell auf die Verlinkungsbehauptung abzielt, angegangen werden.

Das Datenblatt stellt eine Spezifikation für das grundlegende oder zugrunde liegende KI-Bewertungsmodell dar und beschreibt, wie es entwickelt wurde, während die Modellkarte ein kurzer, leicht zugänglicher (d. h. wenig technischer) Validierungsbericht ist, in dem beschrieben wird, wie das Bewertungsmodell für die Verwendung in einem bestimmten Kontext getestet wurde (z. B. wo potenzielle Quellen für Bias in diesem Kontext identifiziert wurden und das Modell mit einem geeigneten Datensatz trainiert wird, um potenziellen Bias zu beseitigen oder zumindest zu begrenzen).

Es ist nicht mehr akzeptabel, einfach ein KI-gesteuertes oder -gestütztes Bewertungssystem zu verwenden, ohne zu belegen, dass es nachweislich für den Einsatz bezüglich bestimmter Entscheidungen in bestimmten Kontexten geeignet ist. Es ist auch nicht akzeptabel zu behaupten, dass eine Anbindung an einen Referenzrahmen oder einen Standard für Sprachen vorgenommen wurde, ohne einen solchen Nachweis mit einer geeigneten Anbindungsstudie zu verknüpfen, wie sie in diesem Handbuch vorgeschlagen wird.

9 Gebru, T., Morgenstern, J., Vecchione, B., Wortman Vaughan, J., Wallach, H., Daum III, H. & Crawford, K. (2021). Datasheets for Datasets. *Communications of the ACM (CACM)*, Band 64, Ausgabe 12: 86–92. <https://doi.org/10.1145/3458723>

10 Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D. & Gebru, T. (2019). *Model Cards for Model Reporting*. FAT* ’19, 29.–31. Januar 2019, Atlanta, GA, USA. Auszug von <https://arxiv.org/pdf/1810.03993.pdf> (3. August 2021)



Hinweise zur eigenen Implementierung ...

Der Zweck eines Verlinkungsprojekts besteht darin, Nachweise zu sammeln, die Ihre Behauptungen bezüglich einer Anbindung Ihres Lehrplans, Ihrer Lehr-/Lernmaterialien, Ihrer Aktivitäten zur Leistungsmessung oder Tests an den GeR untermauern. Die Nachweise setzen sich zusammen aus verschiedenen Strängen, die explizit mit den Konzepten verknüpft sind, die den Kapiteln dieses Handbuchs zugrunde liegen. Es wird nicht erwartet, dass Sie allen hier vorgeschlagenen Prozessen folgen. Sie können sich auch für einen anderen Ansatz entscheiden, müssen dies aber in Ihrer Dokumentation überzeugend begründen. Die verschiedenen Belege bilden zusammen die Grundlage für Ihren Validitätsbericht.

Hier sind einige Aspekte, die Sie beachten sollten:

- Es ist wichtig, sich von Anfang an darüber im Klaren zu sein, dass das Verlinkungsprojekt keinen definitiven Beweis dafür liefern kann, dass Ihre Arbeit exakt an den GeR angebunden ist. Stattdessen sollte man es als einen Prozess betrachten, in dessen Rahmen ausreichend Nachweise gesammelt werden, um ein überzeugendes Argument für die Richtigkeit der eigenen Behauptung zu finden.
- Überlegen Sie sich, wer Ihr Zielpublikum ist, bevor Sie mit dem Schreiben des Berichts beginnen. Wahrscheinlich haben Sie dabei mehrere Zielgruppen im Auge. Erstellen Sie daher einen Kommunikationsplan, bevor Sie den endgültigen Bericht verfassen, damit verschiedene (wenn auch einheitliche) Versionen von Anfang an berücksichtigt werden können.
- Nach Möglichkeit sollten Sie unabhängige externe Belege zur Untermauerung Ihrer Behauptungen hinzufügen. Versuchen Sie beispielsweise, auch externe Personen einzubeziehen, anstatt nur Beurteilungen von Kolleginnen und Kollegen in Ihrer Organisation einzuholen. Das gilt sowohl für das Standard-Setting bei Tests als auch für Beurteilungen von Materialien und Teilen von Curricula.
- Damit die Belege zur Untermauerung einer Behauptung als „ausreichend und überzeugend“ gelten (siehe Abbildung 6.2), wird ein gewisses Maß an unabhängigem externen Input erwartet.
- Die Anbindung bezieht sich auf ein bestimmtes Material (Test, Material oder Curriculum) und auf die Verwendung in einem bestimmten Kontext (z. B. weiterführende Schulen in Malaysia). Wenn wesentliche Änderungen an dem Material oder dem Kontext vorgenommen werden (z. B. weiterführende Schulen in Spanien), kann es sein, dass die Verlinkungsbehauptung nicht mehr unterstützt wird. Es kann daher eine vollständige oder teilweise Neuanknüpfung erforderlich sein.
- Fügen Sie ein reflektierendes Element mit internen und externen Perspektiven ein. Dies kann Ihnen helfen, eine vernünftige und dennoch detaillierte Argumentation zu entwickeln, die dann an verschiedene Kontexte angepasst werden kann, beispielsweise für die Ausrichtung auf verschiedene Stakeholdergruppen, wie im vorherigen Punkt beschrieben.

**ANHANG:
DER PROZESS DER
ANBINDUNG UND
ARGUMENTATION –
FORMULARE**

Dokumentationsbogen zu Kapitel 2: Familiarisierung

Dokumentationsbogen 2.1: Allgemeine Familiarisierung

<p>Dieser Übersichtsbogen zeigt, dass Sie die vorgeschlagenen Aktivitäten abgeschlossen haben. Ihre Antworten können später als Referenz, zur Validierung und zur Berichterstellung verwendet werden.</p>		
Vor- und Nachname		
Datum		
Name der Institution		
Bezeichnung des Materials		
Art des Materials	<input type="checkbox"/> Curriculum <input type="checkbox"/> Kursmaterial <input type="checkbox"/> Lehrwerk <input type="checkbox"/> Prüfung <input type="checkbox"/> Sonstiges: (bitte angeben)	
AKTIVITÄTEN		Antworten
1	Kapitel 1 & 2 – Was bedeutet die Anbindung an den GeR? – Was bedeutet das für mich? Und für meinen Kontext?	
2	Abschnitt 2.4 – Welche Kompetenzen, Aktivitäten und Strategien sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant? Wenn nicht, wie kann ich sie priorisieren?	
3	GeR-Begleitband, Anhänge 1 & 2 – Mit welchen Niveaus werde ich hauptsächlich arbeiten? – Was sind die wichtigsten Aspekte für dieses Niveau und für die angrenzenden? – Welche Aspekte sind in meinem Kontext für die jeweiligen Modi der Kommunikation am relevantesten?	
4	GeR-Begleitband, Abbildung 11 Aktivitäten und Strategien für die Rezeption – Welche Skalen sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant?	
	GeR-Begleitband, Abbildung 12 Aktivitäten und Strategien für die Produktion – Welche Skalen sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant?	

<p>GeR-Begleitband, Abbildung 13 Aktivitäten und Strategien für die Interaktion</p> <ul style="list-style-type: none"> – Welche Skalen sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant? 	
<p>GeR-Begleitband, Abbildung 14 Aktivitäten und Strategien für die Mediation</p> <ul style="list-style-type: none"> – Welche Skalen sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant? 	
<p>GeR-Begleitband, Abbildung 15 Plurilinguale und plurikulturelle Kompetenz</p> <ul style="list-style-type: none"> – Welche Skalen sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant? 	
<p>GeR-Begleitband, Abbildung 16 Kommunikative Sprachkompetenzen</p> <ul style="list-style-type: none"> – Welche Skalen sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant? 	
<p>GeR-Begleitband, Abbildung 17 Gebärdenkompetenzen</p> <ul style="list-style-type: none"> – Welche Skalen sind in meinem Kontext am relevantesten? Warum? – Sind sie alle gleich relevant? 	
<p>Wie begründe ich meine Entscheidungen oben? (wichtig, da dies als Grundlage für die spätere Argumentation gilt)</p>	
<p>Auf einer Skala von 1 bis 10, wie zufrieden bin ich mit meiner allgemeinen Familiarisierung mit dem GeR?</p> <p style="text-align: center;">1 2 3 4 5 6 7 8 9 10</p> <p>Welche weitere Familiarisierung benötige ich?</p> <p>.....</p> <p>.....</p>	

Dokumentationsbogen 2.2: Spezifische Familiarisierung – Sitzung 1

<p>Dieser Übersichtsbogen zeigt, dass Sie die vorgeschlagenen Aktivitäten abgeschlossen haben. Ihre Antworten können später als Referenz, zur Validierung und zur Berichterstellung verwendet werden.</p>	
Vor- und Nachname	
Datum	
Name der Institution	
Bezeichnung des Materials	
Art des Materials	<input type="checkbox"/> Curriculum <input type="checkbox"/> Kursmaterial <input type="checkbox"/> Lehrwerk <input type="checkbox"/> Prüfung <input type="checkbox"/> Sonstiges: (bitte angeben)
AKTIVITÄTEN	Antworten [inklusive Begründung]
TEIL 1	
<p>1. Einführung und Überprüfung</p> <p>Welche Eigenschaften der GeR-Niveaus sind in meinem Kontext am relevantesten?</p>	
<p>2. Schritt 1: GeR-Begleitband, Anhang 2. Raster zur Selbstevaluation: Rezeption/ Produktion/Interaktion</p> <p>Welche Stichworte (Aktion[en], Arten von Informationen, Bedingungen und Einschränkungen) beschreiben die folgenden Aspekte am besten?</p> <ul style="list-style-type: none"> • Hörverständnis • Leseverständnis • Mündliche Produktion • Schriftliche Produktion • Mündliche Interaktion • Schriftliche und Online-Interaktion 	
<p>3. Schritt 2: GeR-Begleitband, Anhang 2. Raster zur Selbstevaluation: Mediation</p> <p>Welche Stichworte (Aktion[en], Arten von Informationen, Bedingungen und Einschränkungen) beschreiben die folgenden Aspekte am besten?</p> <ul style="list-style-type: none"> • Mediation eines Textes • Mediation von Konzepten • Mediation von Kommunikation 	

TEIL 2

4. Aktivitäten: XXXXXXXXXXXX Skalen

Name der Skala X

Warum ist diese Skala in meinem Kontext relevant?

Welche Stichwörter beschreiben meine(n) Zweck(e) am besten?

Name der Skala X

Warum ist diese Skala in meinem Kontext relevant?

Welche Stichwörter beschreiben meine(n) Zweck(e) am besten?

Name der Skala X

Warum ist diese Skala in meinem Kontext relevant?

Welche Stichwörter beschreiben meine(n) Zweck(e) am besten?

Dokumentationsbogen 2.3: Allgemeine Zusammenfassung des Familiarisierungsprozesses

Dieser Übersichtsbogen zeigt, dass Sie die vorgeschlagenen Aktivitäten abgeschlossen haben. Ihre Antworten können später als Referenz, zur Validierung und zur Berichterstellung verwendet werden.	
Vor- und Nachname	
Datum	
Name der Institution	
Bezeichnung des Materials	
Art des Materials	<input type="checkbox"/> Curriculum <input type="checkbox"/> Kursmaterial <input type="checkbox"/> Lehrwerk <input type="checkbox"/> Prüfung <input type="checkbox"/> Sonstiges: (bitte angeben)
AKTIVITÄTEN	Antworten [inklusive Begründung]
1. Einführung und Überprüfung – Welche der im Abschnitt „Spezifische Familiarisierung 1“ besprochenen Skalen und Deskriptoren sind in meinem Kontext am relevantesten? – Warum? – Welche Aspekte in den in dieser Sitzung besprochenen Skalen sind in meinem Kontext am hilfreichsten?	
2. Aktivitäten: XXXXXXXXXXXX Skalen Name der Skala X – Warum ist diese Skala in meinem Kontext relevant? – Welche Stichwörter beschreiben meine(n) Zweck(e) am besten? Name der Skala X – Warum ist diese Skala in meinem Kontext relevant? – Welche Stichwörter beschreiben meine(n) Zweck(e) am besten? Name der Skala X – Warum ist diese Skala in meinem Kontext relevant? – Welche Stichwörter beschreiben meine(n) Zweck(e) am besten? ../..	

3. Aktivitäten: XXXXXXXXXXXX Skalen

Name der Skala X

- Warum ist diese Skala in meinem Kontext relevant?
- Welche Stichwörter beschreiben meine(n) Zweck(e) am besten?

Name der Skala X

- Warum ist diese Skala in meinem Kontext relevant?
- Welche Stichwörter beschreiben meine(n) Zweck(e) am besten?

Name der Skala X

- Warum ist diese Skala in meinem Kontext relevant?
- Welche Stichwörter beschreiben meine(n) Zweck(e) am besten?

../..

Auf einer Skala von 1 bis 10, wie zufrieden sind Sie mit der Sitzung zur spezifischen Familiarisierung mit dem GeR?

1 2 3 4 5 6 7 8 9 10

Dokumentationsbogen zu Kapitel 3: Spezifikation

Dokumentationsbogen 3.1: Beschreibung der Verlinkung und des Materials

ALLGEMEINE INFORMATIONEN	
Vor- und Nachname	
Datum	
Name der Institution	
1. Bezeichnung des Materials	
2. Art des Materials	<input type="checkbox"/> Curriculum <input type="checkbox"/> Kursmaterial <input type="checkbox"/> Lehrwerk <input type="checkbox"/> Prüfung <input type="checkbox"/> Sonstiges: (bitte angeben)
3. Art der Anbindung	<input type="checkbox"/> Neuerstellung <input type="checkbox"/> Nachjustierung
4. Zielsprache	
5. Name des Materials	
6. Kurze Beschreibung des Materials, einschließlich des Zwecks und der Ziele	
7. Reichweite des Materials	<input type="checkbox"/> International <input type="checkbox"/> Regional <input type="checkbox"/> National <input type="checkbox"/> Institutionell
8. Zielgruppe	<input type="checkbox"/> Junge Lernende <input type="checkbox"/> Studierende <input type="checkbox"/> Unterstufe <input type="checkbox"/> Oberstufe <input type="checkbox"/> Erwachsene <input type="checkbox"/> Sonstiges: (bitte angeben)
9. Hauptdomäne(e)	<input type="checkbox"/> Öffentlich <input type="checkbox"/> Privat <input type="checkbox"/> Beruflich <input type="checkbox"/> Schulisch
10. Angestrebte(s) GeR-Niveau(s)	<input type="checkbox"/> A1 <input type="checkbox"/> A2 <input type="checkbox"/> B1 <input type="checkbox"/> B2 <input type="checkbox"/> C1 <input type="checkbox"/> C2
11. Zentrale Merkmale Notieren Sie die zentralen Merkmale des Materials in Bezug auf die angestrebten allgemeinen GeR-Niveaus.	

Dokumentationsbogen 3.2: Sprachaktivitäten, Strategien und Kompetenzen

Vor- und Nachname	
Datum	
Name der Institution	
Bezeichnung des Materials	
Angestrebte kommunikative Sprachaktivitäten und Strategien	<input type="checkbox"/> Hörverstehen
	<input type="checkbox"/> Audio-visuelles Verstehen
	<input type="checkbox"/> Leseverstehen
	<input type="checkbox"/> Strategien für die Rezeption
	<input type="checkbox"/> Mündliche Produktion
	<input type="checkbox"/> Schriftliche Produktion
	<input type="checkbox"/> Produktionsstrategien
	<input type="checkbox"/> Mündliche Interaktion
	<input type="checkbox"/> Schriftliche Interaktion
	<input type="checkbox"/> Online-Interaktion
	<input type="checkbox"/> Interaktionsstrategien
	<input type="checkbox"/> Mediation von Texten
	<input type="checkbox"/> Mediation von Konzepten
	<input type="checkbox"/> Mediation von Kommunikation
<input type="checkbox"/> Mediationsstrategien	
Angestrebte plurilinguale und plurikulturelle Kompetenz	<input type="checkbox"/> Plurilinguales Verstehen
	<input type="checkbox"/> Auf einem plurilingualen Repertoire aufbauen
	<input type="checkbox"/> Auf einem plurikulturellen Repertoire aufbauen
Angestrebte kommunikative Sprachkompetenzen	<input type="checkbox"/> Linguistische Kompetenz
	<input type="checkbox"/> Soziolinguistische Kompetenz
	<input type="checkbox"/> Pragmatische Kompetenz
Angestrebte Gebärdenkompetenzen	<input type="checkbox"/> Linguistische Kompetenz
	<input type="checkbox"/> Soziolinguistische Kompetenz
	<input type="checkbox"/> Pragmatische Kompetenz

Dokumentationsbogen 3.3: Grafische Profilierung der Verlinkung

Dokumentationsbogen 3.3: Grafische Profilierung der Verlinkung

Vor- und Nachname																			
Datum																			
Name der Institution																			
Bezeichnung des Materials																			
C2																			
C1																			
B2																			
B1																			
A2																			
A1																			
	Mündlich	Lesen	Mündlich	Schriftlich	Mündlich	Schriftlich	Mündlich	Schriftlich	Mündlich	Schriftlich	Online	Text	Konzept	Kommunikation					
	Rezeption		Produktion		Interaktion		Mediation												

Dokumentationsbogen 3.4: Illustrative Deskriptoren von Aktivitäten und Strategien beim Spracherwerb

Das folgende Formular sollte für alle Sprachaktivitäten, Strategien und Kompetenzen ausgefüllt werden, auf die das Material abzielt.

Vor- und Nachname	
Datum	
Name der Institution	
Bezeichnung des Materials	

	Kurze Beschreibung und/oder Referenz
Name der Sprachaktivität / Strategie / Kompetenz	
Mit welchen Kommunikationsthemen sollten die Lernenden zurechtkommen? <i>Die Listen im GeR 4.2 können als Referenz dienen (2001: S. 58).</i>	
Mit welchen Kommunikationsaktivitäten und -strategien sollten die Lernenden zurechtkommen? <i>Kapitel 3 des GeR-Begleitbands kann als Referenz dienen (2020: S. 57–143).</i>	
Mit welchen Textsorten sollten die Lernenden zurechtkommen? <i>Die Listen im GeR 4.6.2 und 4.6.3 können als Referenz dienen (2001: S. 96–957).</i>	
Nachdem Sie die Skala für die zugehörige Sprachaktivität durchgelesen haben, geben Sie an und begründen Sie, auf welchem Niveau bzw. welchen Niveaus der Skala die Lernenden eine Leistung erbringen sollten.	

Dokumentationsbogen 3.5: Detaillierte Zuordnung

Vor- und Nachname	
Datum	
Name der Institution	
Bezeichnung des Materials	

Prüfungsteil (Aktivität, Item, Ziel)	Kommunikations- aktivität/ -strategie/ -kompetenz	Domäne(n)	Thema/ Themen	Textsorte(n)	GeR- Skala	GeR- Niveau	GeR- Deskriptor

Dokumentationsbogen zu Kapitel 4: Standardisierung

Dokumentationsbogen 4.1:

Zusammenfassung der Standardisierung – Teilnehmende

ZUSAMMENFASSUNG DER STANDARDISIERUNG: Überblick über den Standardisierungsprozess

<p>Dieser Übersichtsbogen beschreibt die durchgeführten Standardisierungsaktivitäten und kann als Grundlage für die Berichterstellung dienen. Für jeden Bereich muss ein unterschiedlicher Bogen ausgefüllt werden.</p>		
Vor- und Nachname		
Datum		
Name der Institution		
Bezeichnung des Materials		
Art des Materials <input type="checkbox"/> Rezeption (Hören) <input type="checkbox"/> Rezeption (Lesen) <input type="checkbox"/> Schriftliche Produktion <input type="checkbox"/> Mündliche Produktion <input type="checkbox"/> Sonstiges: (bitte angeben)		
AKTIVITÄTEN		ANTWORTEN/DETAILS
1	Familiarisierung	Überprüfung vorheriger Aktivitäten (Kapitel 2) <input type="checkbox"/> Zusätzliche Aktivitäten (kurz auflisten)
2	Spezifikation	Überprüfung vorheriger Aktivitäten (Kapitel 3) <input type="checkbox"/> Angaben zum Kontext (Beschreibung)
3	Standardisierte Beschreibung der erwarteten Lernziele	Beschreibung:
4a	Analyse der illustrativen Beispiele	
	1. Sind sie für das/die Leistungs-niveau(s) relevant, um das/die es mir geht?	Ja <input type="checkbox"/> Nein <input type="checkbox"/> Kommentare:

	2. Sind sie relevant für den Kontext und Zweck meines Materials?	Ja <input type="checkbox"/> Nein <input type="checkbox"/> Kommentare:
	3. Sind sie relevant für die Kommunikationsmethode(n), Sprachaktivitäten und Sprachstrategien in meinem Material?	Ja <input type="checkbox"/> Nein <input type="checkbox"/> Kommentare:
4b	Benchmarking	
	a) Vorhandenes Material Spiegelt mein Material das/die GeR-Niveau(s) präzise wider, auf das/die es ausgerichtet ist?	Ja <input type="checkbox"/> Nein <input type="checkbox"/> Kommentare:
	b). Neues Material Spiegelt der Prüfungsteil das/die GeR-Niveau(s) präzise wider, auf das/die er ausgerichtet ist?	Ja <input type="checkbox"/> Nein <input type="checkbox"/> Kommentare:
5	Fertigstellung des Materials	Beschreiben Sie kurz den Prozess des Schreibens/Bearbeitens:
	a) Vorhandenes Material	
	b) Neues Material	
<p>Auf einer Skala von 1 bis 10, wie zufrieden sind Sie mit dieser Sitzung zur Standardisierung?</p> <p style="text-align: center;">1 2 3 4 5 6 7 8 9 10</p> <p>Wie könnten die Standardisierungssitzungen noch effektiver gestaltet werden?</p> <p>.....</p> <p>.....</p> <p>.....</p>		

Dokumentationsbogen 4.2: Zusammenfassung der Standardisierung – Koordinatorinnen/Koordinatoren

ZUSAMMENFASSUNG DER STANDARDISIERUNG: Überblick über den Standardisierungsprozess

Diese Zusammenfassung beschreibt die durchgeführten Standardisierungsaktivitäten und kann als Grundlage für die Berichterstellung dienen.
Für jeden Bereich muss ein einzelner Dokumentationsbogen (4.1.1, 4.1.2, 4.1.3, ...) ausgefüllt werden.

Vor- und Nachname		
Datum		
Name der Institution		
Bezeichnung des Materials		
Ansatz	<input type="checkbox"/> Top-down (neues Material)	<input type="checkbox"/> Bottom-up (vorhandenes Material)
Bereich	<input type="checkbox"/> Rezeption (Hören)	<input type="checkbox"/> Rezeption (Lesen)
	<input type="checkbox"/> Schriftliche Produktion	<input type="checkbox"/> Mündliche Produktion
	<input type="checkbox"/> Sonstiges: (bitte angeben)	
AKTIVITÄTEN		ANTWORTEN/DETAILS
1	Familiarisierung	Überprüfung vorheriger Aktivitäten (Kapitel 2) <input type="checkbox"/> Zusätzliche Aktivitäten (kurz auflisten)
2	Spezifikation	Überprüfung vorheriger Aktivitäten (Kapitel 3) <input type="checkbox"/> Angaben zum Kontext (Beschreibung)
3	Standardisierte Beschreibung der erwarteten Lernziele	Beschreibung:
	Verwendete illustrative Beispiele (Namen und Quellen) a b c	Rechtfertigung/Ergebnisse der Analyse a b c
4a	Analyse der illustrativen Beispiele	Ja oder nein? Fügen Sie Kommentare hinzu.

	<p>1. Sind sie für das/die Leistungs-niveau(s) relevant, um das es mir geht?</p> <p>2. Sind sie relevant für den Kontext und Zweck meines Materials?</p> <p>3. Sind sie relevant für die Kommunikationsmethode(n), Sprachaktivitäten und Sprachstrategien in meinem Material?</p>	<p>1.</p> <p>2.</p> <p>3.</p>
4b	<p>Benchmarking</p> <p>a). Vorhandenes Material</p> <p>Spiegelt mein Material das/die GeR-Niveau(s) präzise wider, auf das/die es ausgerichtet ist?</p>	<p>Ja oder nein? Fügen Sie Kommentare hinzu.</p> <p>a)</p>
	<p>b). Neues Material</p> <p>Spiegelt der Prüfungsteil das/die GeR-Niveau(s) präzise wider, auf das/die er ausgerichtet ist?</p>	<p>b)</p>
5	<p>Fertigstellung des Materials</p> <p>a) Vorhandenes Material</p> <p>b) Neues Material</p>	<p>Beschreiben Sie kurz den Prozess des Schreibens/Bearbeitens:</p> <p>a)</p> <p>b)</p>

Vorerprobung und Konsultation

Gab es bei Ihrem Material eine Vorerprobung mit den Lernenden, um zu prüfen, ob sie wie gewünscht funktioniert?

Wo?

Mit wem?

Wie?

Wer wurde bezgl. des finalen Entwurfs des Materials konsultiert?

Auf einer Skala von 1 bis 10, wie zufrieden sind die Teilnehmenden mit dieser Standardisierungssitzung?

1 2 3 4 5 6 7 8 9 10

Wie könnten die Standardisierungssitzungen noch effektiver gestaltet werden?

.....

.....

.....

Dokumentationsbogen 4.3: Analyse illustrativer Beispiele – Holistisches Formular

Aufgrund der vielfältigen Eigenschaften von Materialien ist dies lediglich ein Beispielformular, das individuell angepasst werden muss.

Fertigkeit	Dem GeR-Niveau zugeordnete	Kommentare
Beispiel/Aufgabe 1		
Beispiel/Aufgabe 2		
Beispiel/Aufgabe 3		
Beispiel/Aufgabe ...		
Beispiel/Aufgabe N		

Dies ist ein Beispiel für ein Blatt mit einzelner Bewertung, bei dem die Teilnehmenden eine allgemeine Beurteilung über das Niveau jedes Beispiels bzw. jeder Aufgabe abgeben müssen. Das Bewertungsblatt kann für die Beurteilungen von Leistungen (schriftlich oder mündlich) oder Test-Items verwendet werden.

Dokumentationsbogen zu Kapitel 5: Standard-Setting

<p>Im Folgenden finden Sie eine Checkliste mit wichtigen Fragen, die bei der Planung des Standard-Settings in Betracht gezogen werden sollten.</p>	
Vor- und Nachname	
Datum	
Name der Institution	
Bezeichnung des Materials	
Art des Materials	<input type="checkbox"/> Curriculum <input type="checkbox"/> Kursmaterial <input type="checkbox"/> Lehrwerk <input type="checkbox"/> Prüfung <input type="checkbox"/> Sonstiges: (bitte angeben)
Methode auswählen	<input type="checkbox"/> Testbasiert Angaben
	<input type="checkbox"/> Prüfungsteilnehmer-basiert Angaben
Was ist der Grund für die ausgewählte(n) Methode(n)?	
Benötigte Ressourcen ermitteln (Geben Sie jeweils an, ob Sie bereits Zugriff auf diese Ressourcen haben.)	Physisch
	Finanziell
	Personell
Relevante Expertinnen/Experten ermitteln Wie viele Expertinnen/Experten werde ich einberufen? Warum diese Zahl?	
Details dazu, wie die Expertinnen/Experten ermittelt und kontaktiert werden	
Was sind die Mindestvoraussetzungen für die Expertinnen/Experten? (Erfahrung, Qualifikation usw.)	

Zugang zu Testmaterialien gewährleisten Welche Materialien sind erforderlich?	<input type="checkbox"/> Mündliche Interaktion	Details:
	<input type="checkbox"/> Mündliche Produktion	Details:
	<input type="checkbox"/> Schriftliche Interaktion	Details:
	<input type="checkbox"/> Schriftliche Produktion	Details:
	<input type="checkbox"/> Rezeption – Hören	Details:
	<input type="checkbox"/> Rezeption – Lesen	Details:
Zugang zu Testdaten gewährleisten Werde ich Zugriff auf die Leistungen der Prüfungsteilnehmenden und/oder empirische Daten haben?	<input type="checkbox"/> Mündliche Interaktion	Details:
	<input type="checkbox"/> Mündliche Produktion	Details:
	<input type="checkbox"/> Schriftliche Interaktion	Details:
	<input type="checkbox"/> Schriftliche Produktion	Details:
	<input type="checkbox"/> Rezeption – Hören	Details:
	<input type="checkbox"/> Rezeption – Lesen	Details:
Kommunikation vorbereiten	<input type="checkbox"/> Einladungsschreiben vorbereitet	<input type="checkbox"/> Verschickt
	<input type="checkbox"/> Einverständniserklärung vorbereitet	<input type="checkbox"/> Verschickt
	<input type="checkbox"/> Teilnahmeanleitung vorbereitet	<input type="checkbox"/> Verschickt
GeR-Schulungsmaterial (Falls die Materialien noch nicht verfügbar sind, geben Sie an, wie Sie sie produzieren werden.)	<input type="checkbox"/> Mündliche Interaktion	Details:
	<input type="checkbox"/> Mündliche Produktion	Details:
	<input type="checkbox"/> Schriftliche Interaktion	Details:
	<input type="checkbox"/> Schriftliche Produktion	Details:
	<input type="checkbox"/> Rezeption – Hören	Details:
	<input type="checkbox"/> Rezeption – Lesen	Details:

Materialien zur Datenerhebung In welchem Format werde ich die Formulare zur Datenerhebung (Einschätzung der Expertinnen/ Experten) produzieren?	<input type="checkbox"/> Mündliche Interaktion	Details:
	<input type="checkbox"/> Mündliche Produktion	Details:
	<input type="checkbox"/> Schriftliche Interaktion	Details:
	<input type="checkbox"/> Schriftliche Produktion	Details:
	<input type="checkbox"/> Rezeption – Hören	Details:
	<input type="checkbox"/> Rezeption – Lesen	Details:
	<input type="checkbox"/> Feedbackformulare	Details:
Habe ich Feedbackformulare für die Expertinnen/Experten?	<input type="checkbox"/> Mündliche Interaktion	Details:
	<input type="checkbox"/> Mündliche Produktion	Details:
	<input type="checkbox"/> Schriftliche Interaktion	Details:
	<input type="checkbox"/> Schriftliche Produktion	Details:
	<input type="checkbox"/> Rezeption – Hören	Details:
	<input type="checkbox"/> Rezeption – Lesen	Details:
Welche Art von Datenanalyse ist erforderlich, einschließlich der Softwareanforderungen?	<input type="checkbox"/> Mündliche Interaktion	Details:
	<input type="checkbox"/> Mündliche Produktion	Details:
	<input type="checkbox"/> Schriftliche Interaktion	Details:
	<input type="checkbox"/> Schriftliche Produktion	Details:
	<input type="checkbox"/> Rezeption – Hören	Details:
	<input type="checkbox"/> Rezeption – Lesen	Details:
Räumlichkeiten vorbereiten Wo werden die Sitzungen stattfinden?	<input type="checkbox"/> Mündliche Interaktion	Details:
	<input type="checkbox"/> Mündliche Produktion	Details:
	<input type="checkbox"/> Schriftliche Interaktion	Details:
	<input type="checkbox"/> Schriftliche Produktion	Details:
	<input type="checkbox"/> Rezeption – Hören	Details:
	<input type="checkbox"/> Rezeption – Lesen	Details:

Das Handbuch wurde von der European Association for Language Testing and Assessment (EALTA), der UK Association for Language Testing and Assessment (UKALTA), dem British Council und der Association for Language Testers in Europe (ALTE) herausgegeben.

Übersetzung ins Deutsche: Goethe-Institut e. V., Abteilung Sprache
www.goethe.de

© British Council, UKALTA, EALTA und ALTE. 2022.