

Standardüberprüfung E8 – 2013

Glossar



Bundesinstitut für Bildungsforschung, Innovation & Entwicklung
des österreichischen Schulwesens

Alpenstraße 121/5020 Salzburg

Direktoren: MMag. Christian Wiesner & Mag. Martin Netzer, MBA

Für den Inhalt verantwortlich: BIFIE – Department Bildungsstandards & Internationale Assessments

(Mag. Dr. Claudia Schreiner, Mag. Simone Breit)

Kontakt: 0662/620088-3000; office.bista@bifie.at

www.bifie.at

Bei Fragen zur Rückmeldung kontaktieren Sie bitte unsere Hotline:

Telefon: 0662/620088-3700; E-Mail: bist8-rm@bifie.at

Die Überprüfung und Rückmeldung der Bildungsstandards ist rechtlich verpflichtend verankert und zählt zu den gesetzlichen Kernaufgaben des Bundesinstituts BIFIE (BIFIE-Gesetz 2008).



Bundesministerium für Unterricht, Kunst und Kultur
Minoritenplatz 5/1014 Wien

Inhalt

3	A
3	An Gesprächen teilnehmen
3	Antwortformate
3	Assessment Scale (Bewertungsskala für Sprechen)
4	Assessor
4	B
4	Baseline-Testung
4	BIFIE
4	Bildungsstandards (BIST)
5	E
5	Englischwert
5	Erwartungsbereich und fairer Vergleich
6	F
6	Fairer Vergleich
6	Fragebögen
6	G
6	GERS-Referenzniveaus
9	H
9	Hören
10	I
10	Informelle Kompetenzmessung (IKM)
10	Interlocutor
11	Item (Testitem)
11	K
11	Kompetenzen
11	Kompetenzmodell/Kompetenzbereiche
12	Kontextfragebögen
12	Kriteriale Rückmeldung
12	L
12	Leistungsabstand der mittleren 75 %
13	Lesen
13	M
13	Migrationshintergrund
14	Mittelwert
14	Mittlere Differenz
15	N
15	n
15	R
15	Rater
15	Referenzwerte
15	Rückmeldemoderatorinnen/-moderatoren (RMM)

- 15 S**
15 Schreiben
16 Schülerfragebogen
16 Sozialstatus
17 Sprechen
18 Standardüberprüfung bzw. Bildungsstandardüberprüfung
18 Standard Setting
18 Streuung
- 19 T**
19 Teilkompetenzen/Sektionen
19 Testhefte/Testformen
19 Testitem
- 19 U**
19 Überprüfung
19 Unterrichtsgruppe
- 20 V**
20 Verordnung zu den Bildungsstandards
20 Vertrauensintervall
- 21 W**
21 Writing Rating Scale (Beurteilungsskala für Schreiben)
- 21 Z**
21 Zusammenhängend sprechen

A

An Gesprächen teilnehmen

Siehe unter *Sprechen*

Antwortformate

Im Rahmen der *Überprüfungen* werden *Testitems* mit unterschiedlichen Antwortformaten eingesetzt. Das Antwortformat bestimmt, in welcher Art und Weise die Schüler/innen die *Testitems* beantworten.

Offenes Antwortformat

Beim offenen Antwortformat generieren die Schüler/innen ihre Antwort selbst. Diese *Items* überlassen es vollständig den Schülerinnen und Schülern, wie sie die Aufgabenstellung lösen. Die Aufgabenlösung kann verbal frei gestaltet werden. Die Schüler/innen schreiben z. B. einen Text in der Fremdsprache.

Halboffenes Antwortformat

Halboffene *Items* überlassen die Antwortformulierung dem Schüler/der Schülerin. Die Aufgabe ist jedoch im Unterschied zum offenen *Item* so präzise gestellt, dass die Antwort mit geringem Aufwand als richtig oder falsch bewertet werden kann. Vervollständigungsitems bzw. Prediction-Items in *Hören* und Lückentexte bzw. Gap-Fill-Items in Lesen zählen zu den halboffenen Antwortformaten. Prediction-Items sind jene Aufgaben, bei denen die Schüler/innen einen Satz(teil) hören, in dem das letzte Wort fehlt. Die Schüler/innen müssen den gehörten Satz mit genau einem Wort sinngemäß vervollständigen. Gap-Fill-Items erfordern in einem Lückentext das Ergänzen jeweils eines passenden Worts pro Lücke. Bei manchen Aufgaben werden die Schüler/innen in der Angabe explizit angewiesen, hierfür nur ein Wort zu verwenden, das bereits anderswo im Lückentext vorkommt.

Geschlossenes Antwortformat

Beim geschlossenen Antwortformat stehen den Schülerinnen und Schülern zu einer Frage mehrere vorformulierte Antworten zur Auswahl. Die Antworten werden eingescannt und elektronisch weiterverarbeitet.

- Richtig-falsch-Items (True/False zum Ankreuzen) bestehen aus einer Aussage und zwei Antwortalternativen („richtig“ oder „falsch“), von denen eine zutrifft. Um die Ratewahrscheinlichkeit zu minimieren, werden solche *Items* nur blockweise vorgelegt.
- Multiple-Choice-Items (zum Ankreuzen): Die Schülerin/der Schüler muss hierbei aus mehreren zur Wahl gestellten Antwortmöglichkeiten diejenige auswählen, die sie/er für richtig hält.
- Zuordnungsitems (Multiple-Matching-Items) erfordern, dass der Schüler/die Schülerin vorgegebenen Begriffen eine entsprechende Beschreibung oder ein Bild korrekt zuweist.

Illustrierende Beispielitems finden Sie auf der BIFIE-Homepage unter www.bifie.at/node/1822.

Assessment Scale (Bewertungsskala für Sprechen)

Die Beurteilung der Sprechleistungen der Schüler/innen erfolgt durch sogenannte *Assessors* anhand eines Kriterienrasters (Assessment Scale). Für jede der vier *Teilkompetenzen* (Erfüllung der Aufgabenstellung, klare & natürliche Sprache, Grammatik und Wortschatz) des *Kompetenzbereichs Sprechen* werden dabei zwischen 0 und 7 Punkte vergeben. Die Assessment Scale mit einer inhaltlichen Beschreibung der einzelnen Punkte finden Sie unter www.bifie.at/node/1821.

Assessor

Die *Überprüfung* in Englisch *Sprechen* wird von einem Testleitungspaar, bestehend aus einem *Interlocutor* und einem Assessor, durchgeführt. Assessors beurteilen dabei die Sprechleistungen der Schüler/innen unter Zuhilfenahme eines Kriterienrasters (= *Assessment Scale*).

Assessors sind erfahrene Lehrer/innen, die in einer mehrphasigen Fortbildung sowohl zur Bewertung von Sprechperformanzen als auch als *Interlocutors* zur Durchführung von Sprechtests qualifiziert wurden. Im Rahmen der *Überprüfungen* konnten so auch die Rollen getauscht werden.

B

Baseline-Testung

Bei der sogenannten Baseline-Testung, die im Frühjahr 2009 auf der 8. Schulstufe und im Frühjahr 2010 auf der 4. Schulstufe stattgefunden hat, handelt es sich um die Ausgangsmessung für die *Standardüberprüfung*. Das Ziel der Testung war die objektive Feststellung des Ist-Stands vorhandener *Kompetenzen* bei den Schülerinnen und Schülern der 8. und 4. Schulstufe. Sie erfasste die *Kompetenzen* von Schülerinnen und Schülern vor Einführung der *Bildungsstandards* und dient als Vergleichsmaßstab für die *Standardüberprüfungen*. Die für die Baseline-Testung ausgewählten Schulen wurden durch eine Zufallsziehung eruiert, die getesteten Schüler/innen entsprachen einer repräsentativen Stichprobe.

BIFIE

Das Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (BIFIE) wurde vom Nationalrat mit 1. Jänner 2008 gegründet. Das BIFIE hat folgende Kernaufgaben zu erfüllen:

- Bildungsmonitoring (z. B. PISA, PIRLS, TIMSS)
- Angewandte Bildungsforschung (Begleitung und Evaluation bildungspolitischer Reformen)
- Qualitätsentwicklung (z. B. Entwicklung, Implementierung und begleitende Evaluierung der standardisierten kompetenzorientierten Reife- und Diplomprüfung)
- Erstellung des Nationalen Bildungsberichts

Die Durchführung, Auswertung und Rückmeldung der *Standardüberprüfungen* liefert sowohl zum Bildungsmonitoring als auch zur Qualitätsentwicklung einen Beitrag. Bildungsmonitoring bedeutet, dass das Bildungssystem hinsichtlich seiner Praxis und der erzielten Ergebnisse kontinuierlich beobachtet wird. Die Ergebnismeldungen der *Bildungsstandards* bieten eine wichtige Informationsgrundlage, um Qualitätsentwicklungs- und Qualitätssicherungsprozesse auf allen Ebenen des Schulsystems anzuregen.

Detaillierte Informationen zum BIFIE finden Sie auf der BIFIE-Homepage unter www.bifie.at.

Bildungsstandards (BIST)

Bildungsstandards legen fest, was Schüler/innen in der Regel am Ende der 4. und der 8. Schulstufe in den Gegenständen Deutsch und Mathematik sowie auf der 8. Schulstufe auch in Englisch (sofern erste lebende Fremdsprache) können sollen. Sie sind konkret formulierte Lernergebnisse, die sich aus dem Lehrplan ableiten.

Informationen zu den Bildungsstandards finden Sie auf der BIFIE-Homepage unter www.bifie.at/bildungsstandards.

E

Englischwert

Die bewerteten Schülerantworten der *Kompetenzbereiche Hören, Lesen und Schreiben* bilden zusammen den Englischwert. Diese drei *Kompetenzbereiche* werden dabei gleichwertig berücksichtigt. Der Englischwert wird mittels Skalierung (statistische Überführung) auf einer Punktskala von 200 bis 800 dargestellt. Durch die Verwendung dieser einheitlichen Skala sind die Leistungen aller Schüler/innen trotz unterschiedlicher *Testformen* miteinander vergleichbar.

Die Skala wurde im Rahmen der *Baseline-Testung* im Jahr 2009 so festgelegt, dass sich ein *Mittelwert* (MW) von 500 Skaleneinheiten mit einer Standardabweichung von 100 ergab. Um die Leistungen aller österreichischen Schüler/innen aus der *Standardüberprüfung 2013* mit denen der *Baseline-Testung* vergleichen und somit eine Veränderung sichtbar machen zu können, wurden die Ergebnisse der *Standardüberprüfung 2013* auf die Punktskala der Ausgangsmessung von 2009 übertragen.

Ein Englischwert wird auch für Schüler/innen berechnet, die die Testung vorzeitig abgebrochen haben. Keinen Schülerwert gibt es nur für jene *Kompetenzbereiche*, in denen die Schülerin/der Schüler kein *Testheft* bearbeitet hat. Diese Vorgehensweise kann dazu führen, dass der *Mittelwert* aller *Kompetenzbereiche* nicht dem Englischwert entspricht.

Erwartungsbereich und fairer Vergleich

Die Leistungen der Schüler/innen werden wesentlich von Rahmenbedingungen bestimmt, die von der Schule bzw. der unterrichtenden Lehrkraft nicht beeinflusst werden können. Daher beinhalten die Rückmeldungen auf Schul- und Unterrichtsgruppenebene zusätzlich zu kriterialen und sozialen Vergleichen faire Vergleiche in Bezug auf die Ergebnisse der *Standardüberprüfung*, indem ein Erwartungsbereich für die Schule/*Unterrichtsgruppe* berechnet wird. Der Erwartungsbereich einer Schule/*Unterrichtsgruppe* ist jener Bereich um einen berechneten Wert, der aufgrund der gegebenen strukturellen Rahmenbedingungen (statistisch) zu erwarten wäre. Für alle anderen Schulen/*Unterrichtsgruppen* mit vergleichbaren strukturellen Rahmenbedingungen würde sich somit auch der gleiche Erwartungsbereich ergeben. Man spricht demzufolge vom fairen Vergleich. Für die Berechnung des Erwartungsbereichs im Rahmen des fairen Vergleichs werden standortbezogene Merkmale, schulbezogene Merkmale sowie Merkmale der Zusammensetzung der Schülerpopulation (hinsichtlich demografischer und sozioökonomischer Aspekte) herangezogen. In der aktuellen *Standardüberprüfung* auf der 8. Schulstufe wurden folgende Merkmale berücksichtigt:

Standortbezogene Merkmale:

- Gemeindegröße
- Urbanisierungsgrad
- Entfernung zur nächstgelegenen AHS

Schul- bzw. unterrichtsgruppenbezogene Merkmale:

- Schulart (AHS/APS)
- Schulerhalter (öffentlich, privat)
- Schulgröße/Größe der *Unterrichtsgruppe*

Schülerbezogene Merkmale:

- Anteil der Mädchen/Burschen
- Anteil der Schüler/innen mit/ohne *Migrationshintergrund*
- Anteil der Schüler/innen, deren Erstsprache Deutsch ist

- *Sozialstatus* der Schüler/innen (Ausbildung und beruflicher Status der Eltern, Anzahl der Bücher zu Hause)
- Anzahl bzw. Anteil der von der Testung ausgenommenen Schüler/innen

Anteile und Anzahlen beziehen sich dabei (mit Ausnahme der Anzahl der von der Testung ausgenommenen Schüler/innen) ausschließlich auf die getesteten Schüler/innen der betrachteten Schule (im Schulbericht) bzw. der betrachteten *Unterrichtsgruppe* (in den Rückmeldungen für Lehrer/innen). Die Angaben dazu stammen entweder von der Statistik Austria (Gemeindegröße, Urbanisierungsgrad) oder werden den jeweiligen *Kontextfragebögen* entnommen.

Das Hauptaugenmerk liegt dabei auf jenen Merkmalen, die nicht direkt von Schule und Unterricht beeinflusst werden können. Nicht enthalten sind deshalb z. B. Leistungsgruppen-Einteilungen, Schulschwerpunkte, Stundentafeln etc.

F

Fairer Vergleich

Siehe unter *Erwartungsbereich und fairer Vergleich*

Fragebögen

Siehe unter *Kontextfragebögen*

G

GERS-Referenzniveaus

Die *Überprüfung der Kompetenzen* in Englisch basiert auf dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (GERS), der eine gemeinsame Basis für die Entwicklung von zielsprachlichen Lehrplänen, curricularen Richtlinien, Prüfungen, Lehrwerken usw. in ganz Europa darstellt.

Der GERS beschreibt umfassend, was zu lernen ist, um eine Sprache für kommunikative Zwecke benutzen zu können, und welche Kenntnisse und Fertigkeiten dementsprechend entwickelt werden müssen. Er bietet Sprachkompetenzbeschreibungen für sechs verschiedene Niveaus an (A1, A2, B1, B2, C1, C2). Dabei wird definiert, was auf den einzelnen Niveaus von Sprachlernenden erwartet werden kann. Die drei großen Niveaus A, B und C entsprechen einer elementaren, einer selbstständigen und einer kompetenten Sprachverwendung (vgl. Tabelle 1).

Kompetente Sprachverwendung (hohes Sprachniveau bis hin zu fast muttersprachlicher Kompetenz)	C	C2
		C1
Selbstständige Sprachverwendung (mittleres allgemeines und berufsbezogenes Niveau)	B	B2
		B1
Elementare Sprachverwendung (grundlegendes oder fortgeschrittenes Anfängerniveau)	A	A2
		A1

Tabelle 1: GERS-Niveaus

In Österreich ist der GERS sowohl Grundlage des Lehrplans für die lebenden Fremdsprachen als auch der *Verordnung der Bildungsstandards*. Nach dem 4. Lernjahr (für Englisch in der Regel auf der 8. Schulstufe) ist im Lehrplan für die lebenden Fremdsprachen als Grundanforderung definiert, dass Schüler/innen in allen *Kompetenzbereichen* das Niveau A2 erreichen sollen, unter förderlichen Begleitumständen jedoch auch die Erreichung von ausgewiesenen Fertigkeiten auf B1 möglich ist. Die *Bildungsstandards* wiederum beschreiben fachliche Grundkompetenzen auf den Niveaus A2 und B1, die die Schüler/innen in der Regel erreichen sollten (vgl. Abbildung 1). Um den Grad der Kompetenzerreichung darzustellen, werden im Fach Englisch, basierend auf dem GERS, folgende *Referenzniveaus* rückgemeldet:

- B1 oder darüber
- A2
- A1 oder darunter

Folgende Grafik verdeutlicht die Zusammenhänge:

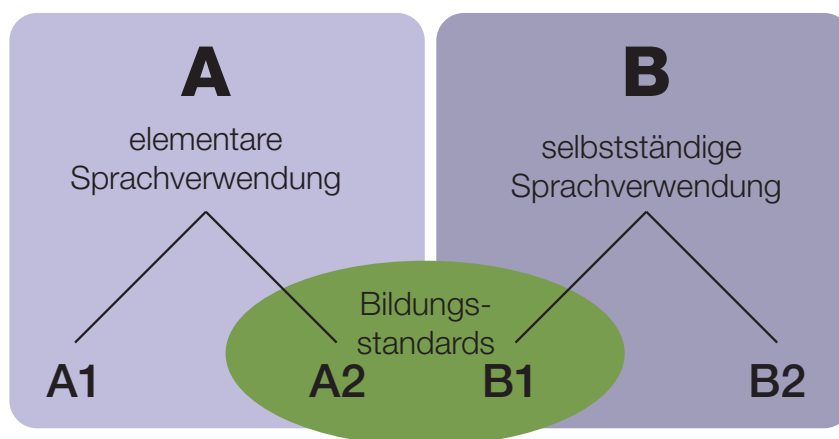


Abbildung 1: GERS-Levels und Bildungsstandards

Die in der Verordnung zu den *Bildungsstandards* angeführten Deskriptoren (Beschreibungen) stellen die Fähigkeiten und Fertigkeiten dar, die im jeweiligen *Kompetenzbereich* in der Regel erreicht werden sollten. Wie aus Tabelle 2 ersichtlich, sind die Deskriptoren der *Bildungsstandards* für die erste lebende Fremdsprache Englisch überwiegend den Niveaus A2 und B1 zuzuordnen. Bei der Betrachtung der verschiedenen *Kompetenzbereiche* fällt auf, dass in den einzelnen *Kompetenzbereichen* die Anzahl und Anteile der Deskriptoren auf unterschiedlichem Niveau variieren.

Kompetenzbereiche	A1		A2		B1		Anzahl der BIST-Deskriptoren
	Anzahl	%	Anzahl	%	Anzahl	%	
Hören	1	16,7	1	16,7	4	66,7	6
Lesen	0	0	3	50,0	3	50,0	6
Schreiben	0	0	4	50,0	4	50,0	8
Sprechen	0	0	9	60,0	6	40,0	15

Tabelle 2: Anzahl der Deskriptoren

Für die GERS-Referenzniveaus wurden in einem ersten Schritt inhaltliche Beschreibungen für die Niveaus A2 und B1 oder darüber entwickelt. Um die Schüler/innen auf Basis ihrer erbrachten Testleistung den GERS-Referenzniveaus zuordnen zu können, wurde in weiterer Folge eine repräsentative Auswahl an verwendeten

Testitems (bei *Hören* und *Lesen*) und Schülerperformanzen (produzierte Texte bei *Schreiben* bzw. Videos bei *Sprechen*) durch Expertinnen und Experten aus Schulpraxis (Sekundarstufe I und II), Fachdidaktik, Pädagogik und Psychologie den verschiedenen *Referenzniveaus* zugeordnet. Die Schwierigkeiten der *Items* und Schülerperformanzen wurden dabei auf derselben Punktskala gemessen wie die Leistungen. Daher ließen sich Schwellenwerte auf der Punktskala bestimmen, die die *Referenzniveaus* voneinander abgrenzen.

Nachstehend finden Sie für jeden *Kompetenzbereich* die **inhaltlichen Beschreibungen der Referenzniveaus für A2 und B1 oder darüber**. Auf eine **inhaltliche Beschreibung des Referenzniveaus A1** wurde verzichtet, da man im Rahmen der *Standardüberprüfung* keine Aussagen darüber tätigen kann, inwieweit Schüler/innen das Referenzniveau tatsächlich erreichen. In der Regel können Schüler/innen auf diesem Niveau in den *Kompetenzbereichen Lesen* und *Hören* vertraute, alltägliche Ausdrücke und ganz einfache Sätze verstehen, die auf die Befriedigung konkreter Bedürfnisse zielen. Es geht dabei v. a. um persönliche Themen (Bekanntmachung, Wohnort, Bekanntschaften, Besitz). In den *Kompetenzbereichen Schreiben* und *Sprechen* können Schüler/innen in der Regel sehr kurze und sehr einfach strukturierte Sätze oder Aussagen produzieren. Dabei werden sehr einfache grammatikalische Strukturen und ein extrem eingeschränktes Vokabular verwendet.

Beschreibung der GERS-Referenzniveaus – Kompetenzbereich Hören	
B1 oder darüber	Zusätzlich zu den Kompetenzen des Niveaus A2 können Schüler/innen auf dem Niveau B1 Gesprächen und Monologen die Hauptpunkte entnehmen. Sie können Geschichten und Erzählungen aus dem Alltag verstehen sowie einfachen Interviews oder Berichten folgen, wenn es sich um vertraute Themenbereiche handelt, Standardsprache verwendet und deutlich gesprochen wird. Schüler/innen können Äußerungen kontextbezogen zuordnen, Schlussfolgerungen aufgrund von Textinformationen ziehen und die Haltung der sprechenden Person erkennen. Sie können außerdem ein fehlendes Wort aus dem Kontext und den lexikalischen Zusammenhängen heraus erschließen. ab 539 Punkten
A2	Schüler/innen auf dem Niveau A2 können einfache Wendungen und Wörter aus einer begrenzten Auswahl von gesprochenen Texten – wie etwa Anweisungen, Mitteilungen oder Fragen – verstehen, wenn die Gespräche oder Monologe deutlich und langsam gesprochen werden sowie von unmittelbarer Bedeutung sind. Schüler/innen können die Kompetenzen <i>globales Verstehen</i> , <i>selektives Verstehen</i> und <i>detailliertes Verstehen</i> nachweisen. Sie können außerdem ein fehlendes Wort aus dem Kontext und den lexikalischen Zusammenhängen heraus meistens/weitestgehend erschließen. 358–538 Punkte
A1 oder darunter *)	bis 357 Punkte

*) Erläuterungen zum **GERS-Referenzniveau A1 oder darunter** finden Sie im Text zu Beginn dieses Abschnitts.

Abbildung 2: Inhaltliche Beschreibung der GERS-Referenzniveaus für den Kompetenzbereich Hören

Beschreibung der GERS-Referenzniveaus – Kompetenzbereich Lesen	
B1 oder darüber	Zusätzlich zu den Kompetenzen des Niveaus A2 können Schüler/innen auf dem Niveau B1 einfachen Texten zu vertrauten Themenbereichen in Zeitungen und Zeitschriften die wesentlichen Informationen entnehmen, wenn sie klar gegliedert sind. Sie können unkomplizierte Sachtexte über Themen, die mit den eigenen Interessen und den Themenbereichen des Lehrplans in Zusammenhang stehen, und einfache literarische Texte (z. B. Gedichte) verstehen. Schüler/innen können stilistische Merkmale eines Textes erfassen sowie diesen einer bestimmten Situation zuordnen. Die Schüler/innen können Folgerungen und Schlüsse aus einem Text ziehen und so Haltung, Gesinnung und Absicht der schreibenden Person ableiten. Sie können außerdem die logische Organisation eines schriftlichen Textes und lexikalische Zusammenhänge verstehen sowie unbekannte sprachliche Elemente aus dem Kontext heraus erschließen. ab 575 Punkten
A2	Schüler/innen auf dem Niveau A2 können kurze, einfache persönliche Briefe, Karten oder E-Mails verstehen. Sie können vertrauten Alltagstexten (z. B. Schildern und Broschüren – auch aus dem Internet) die wichtigsten Informationen entnehmen und können einfache, alltägliche Anleitungen und Vorschriften verstehen. Schüler/innen können die Kompetenzen <i>globales Verstehen</i> , <i>selektives Verstehen</i> und <i>detailliertes Verstehen</i> nachweisen. Die Schüler/innen können außerdem die logische Organisation eines schriftlichen Textes und lexikalische Zusammenhänge einigermaßen verstehen. 406–574 Punkte
A1 oder darunter *)	bis 405 Punkte

*) Erläuterungen zum **GERS-Referenzniveau A1 oder darunter** finden Sie im Text zu Beginn dieses Abschnitts.

Abbildung 3: Inhaltliche Beschreibung der GERS-Referenzniveaus für den Kompetenzbereich Lesen

Beschreibung der GERS-Referenzniveaus – Kompetenzbereich Schreiben	
B1 oder darüber	Schüler/innen auf dem Niveau B1 können in einem unkomplizierten, zusammenhängenden, aber bereits etwas detaillierten Text Informationen zu vertrauten Themen erfolgreich schriftlich kommunizieren und damit Situationen in ihrer Lebenswelt sprachlich differenzierter bewältigen. Sie können Elemente eines Textes in einer schlüssigen Abfolge verbinden und eine größere Bandbreite von grammatikalischen Strukturen und lexikalischen Wendungen flexibel einsetzen. ab 560 Punkten
A2	Schüler/innen auf dem Niveau A2 können in einfachen Sätzen Informationen zu einer Reihe von vertrauten Themen erfolgreich schriftlich kommunizieren und damit einfache Alltagssituationen bewältigen. Sie können Sätze und Satzteile mit einfachen Mitteln verbinden und ein einfaches Repertoire von grammatikalischen Strukturen und lexikalischen Phrasen ausreichend korrekt verwenden. 464–559 Punkte
A1 oder darunter *)	bis 463 Punkte

*) Erläuterungen zum **GERS-Referenzniveau A1 oder darunter** finden Sie im Text zu Beginn dieses Abschnitts.

Abbildung 4: Inhaltliche Beschreibung der GERS-Referenzniveaus für den Kompetenzbereich Schreiben

Beschreibung der GERS-Referenzniveaus – Kompetenzbereich Sprechen	
B1 oder darüber	Schüler/innen auf dem Niveau B1 können über vertraute Themen alleine und mit anderen Personen sprechen bzw. Alltagssituationen bewältigen, dabei aber bereits unkomplizierte Informationen verständlich kommunizieren. Sie verfügen über ausreichend strukturelle und lexikalische Mittel, um sich auszudrücken, Wiederholungen oder Schwierigkeiten mit der Formulierung kommen jedoch vor. Sie verwenden ihr Repertoire mit relativer Korrektheit, Fehler können aber noch passieren. Gelegentliche Pausen für lexikalische oder grammatikalische Planung sind notwendig. Ihre Aussprache ist trotz eines möglichen Akzents eindeutig verständlich. ab 551 Punkten
A2	Schüler/innen auf dem Niveau A2 können über vertraute Themen alleine und mit anderen Personen sprechen bzw. Alltagssituationen bewältigen, dabei aber nur beschränkt Informationen kommunizieren. Sie verwenden einfache Satzstrukturen, Phrasen oder Wortgruppen. Sie müssen gelegentlich aufgrund von fehlendem Vokabular Kompromisse eingehen und machen noch häufig Fehler. Sie kommunizieren noch sehr zögernd und der Redefluss ist häufig unterbrochen. Ihre Aussprache ist trotz eines möglichen Akzents deutlich genug, um verstanden zu werden. 469–550 Punkte
A1 oder darunter *)	bis 468 Punkte

*) Erläuterungen zum **GERS-Referenzniveau A1 oder darunter** finden Sie im Text zu Beginn dieses Abschnitts.

Abbildung 5: Inhaltliche Beschreibung der GERS-Referenzniveaus für den Kompetenzbereich Sprechen

H

Hören

Beim Kompetenzbereich Hören wurde überprüft, wie gut Schüler/innen einen gesprochenen Text, der ihnen per CD dargeboten wurde, verstehen und verarbeiten konnten. Dafür wurden mehrere kurze Texte (*Items*) vorgespielt (in der Regel 33 *Items*) und die Schüler/innen mussten dazugehörige Aufgabenstellungen im *Testheft* bearbeiten. Die Testdauer dauerte ca. 32 Minuten. Zu den Höraktivitäten gehörten unter anderem:

- Anweisungen, Mitteilungen oder Fragen hören und das Wesentliche verstehen,
- Geschichten und Erzählungen aus dem Alltag hören und bestimmte Details verstehen,
- einfache Gespräche oder Interviews mithören und Grundsätzliches aus dem Kontext heraus verstehen (z. B. den Beruf oder die Absicht der Sprechenden Person),
- einen gehörten Satz am Ende mit einem dazupassenden fehlenden Wort vervollständigen.

Sämtliche Hörbeispiele kamen dabei aus einem Themenbereich, der für das Alter der Schüler/innen angemessen war. Die Hörbeispiele sind deutlich und in englischer Standardsprache gesprochen. Im Rahmen der *Standardüberprüfung* wurde der *Kompetenzbereich Hören* in die folgenden *Teilkompetenzen (Sektionen)* untergliedert:

Beschreibungen der einzelnen Sektionen im Kompetenzbereich Hören	
Ergänzendes Verständnis (Prediction)	Bei diesen halboffenen Aufgaben hören die Schüler/innen jeweils einen Satz, in dem das letzte Wort fehlt. Der gehörte Satz ist mit genau einem Wort sinngemäß zu vervollständigen.
Schlussfolgerndes Hören (Inferring meaning from text and context)	Aufgaben aus diesem Bereich umfassen das detailgetreue Hören mit dem Augenmerk auf Schlussfolgerungen. Man stellt damit fest, ob die Schüler/innen allein aus den sachlichen Informationen in einem Hörtext Rückschlüsse ziehen oder Ableitungen herstellen können.
Kontextbezogenes Hören (Relating utterances to social/situational contexts)	Dieser Bereich stellt fest, ob Schüler/innen das Gehörte einer bestimmten Situation aus dem Alltag zuordnen können.
Identifizieren der Haltung einer sprechenden Person (Identifying speaker attitude)	Dies umfasst die Kompetenz, die Einstellung einer Person mithilfe des Gehörten ermitteln zu können. Dies kann sowohl durch inhaltliche als auch durch sprachliche Faktoren wie Intonation oder Betonung(en) bestimmt werden.
Globales Hören (Listening for gist/main idea)	Globales Hören überprüft das überblicksmäßige Erfassen der Hauptaussage oder Funktion eines Hörtextes.
Detailliertes Hören (Listening for specific information/details)	Dieser Bereich umfasst jene Aufgaben, bei denen überprüft wird, ob konkret gemachte Aussagen im Detail verstanden werden können.

Abbildung 6: Inhaltliche Beschreibung der Teilkompetenzen (Sektionen) – Kompetenzbereich Hören

I

Informelle Kompetenzmessung (IKM)

Dieses Diagnoseinstrument wird vom *BIFIE* in Zusammenarbeit mit Projektpartnern erstellt. Es ist als Selbst-evaluierungstool konzipiert und besteht aus Aufgaben zu verschiedenen *Kompetenzen*, die analog zu den *Items* der *Standardüberprüfungen* erstellt worden sind. Damit wird Lehrerinnen und Lehrern ermöglicht, den bis dato erreichten Leistungsstand ihrer Schüler/innen in Bezug auf die in den Standards formulierten Zielkompetenzen differenziert nach *Kompetenzbereichen* festzustellen. Die Ergebnisse daraus bieten den Lehrerinnen und Lehrern eine Grundlage, mögliche Stärken und Schwächen der Schüler/innen zu identifizieren und etwaige notwendige Fördermaßnahmen zu setzen. Weitere Informationen zur IKM finden Sie unter www.bifie.at/ikm.

Interlocutor

Die *Überprüfung* in Englisch *Sprechen* wird von einem Testleitungspaar, bestehend aus einem Interlocutor und einem *Assessor*, durchgeführt. Der Interlocutor übernimmt dabei die Rolle der/des Gesprächsleiterin/des Gesprächsleiters. Zu ihren/seinen Aufgaben zählt die Verteilung der Aufgabenstellungen (prompt cards), die Überwachung der Vorbereitungs- und Sprechzeiten und, wenn nötig, die Unterstützung in Form vorgegebener Hilfestellungen.

Interlocutors sind erfahrene Lehrer/innen, die in einer mehrphasigen Fortbildung sowohl zur Durchführung von Sprechtests als auch als Assessors zur Bewertung von Sprechperformanzen qualifiziert wurden. Im Rahmen der *Überprüfungen* konnten so auch die Rollen getauscht werden.

Item (Testitem)

Als Item bzw. Testitem werden einzelne Aufgaben und/oder Fragen bezeichnet, die bei der Konstruktion von Tests verwendet werden. Sie dienen zur Messung eines nicht direkt beobachtbaren Sachverhalts, wie beispielsweise einer bestimmten *Kompetenz* aus Englisch.

K

Kompetenzen

Im Sinne der österreichischen *Verordnung über Bildungsstandards* im Schulwesen sind Kompetenzen definiert als:

„längerfristig verfügbare kognitive Fähigkeiten und Fertigkeiten, die von Lernenden entwickelt werden und die sie befähigen, Aufgaben in variablen Situationen erfolgreich und verantwortungsbewusst zu lösen und die damit verbundene motivationale und soziale Bereitschaft zu zeigen.“ (BGBl. II Nr. 1/2009, §2 (2))

Kompetent zu sein bedeutet somit, das vorhandene Wissen sowie Fähigkeiten und Fertigkeiten in unterschiedlichen Situationen anwenden zu können. So sollten Schüler/innen z. B. lernen, kurze, einfache Notizen und Mitteilungen zu schreiben, die sich auf unmittelbare Bedürfnisse beziehen.

Kompetenzmodell/Kompetenzbereiche

Den in den *Bildungsstandards* festgelegten *Kompetenzen* liegt ein aus dem jeweiligen Lehrplan abgeleitetes fachspezifisches Kompetenzmodell zugrunde. Das Kompetenzmodell strukturiert die wesentlichen Kernbereiche eines Unterrichtsgegenstands.

Für die lebende Fremdsprache Englisch basiert der Lehrplan – wie die *Bildungsstandards* – auf dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (*GERS*). Im Zentrum des *GERS* stehen Sprachkompetenzbeschreibungen für die in Abbildung 7 angeführten Kompetenzbereiche.

Als Kompetenzbereiche werden fertigkeitenbezogene Teilbereiche innerhalb eines Kompetenzmodells bezeichnet. Für Englisch auf der 8. Schulstufe sind das *Hören*, *Lesen*, *Schreiben* und *Sprechen* (bestehend aus *An Gesprächen teilnehmen* und *Zusammenhängend sprechen*).

Für die *Standardüberprüfung* und Rückmeldung werden die Fertigkeiten *An Gesprächen teilnehmen* und *Zusammenhängend sprechen* zum Kompetenzbereich *Sprechen* zusammengefasst. Beide, Lehrplan wie *Bildungsstandards*, definieren als wesentliche Kriterien des Fremdsprachenunterrichts die grundsätzliche Gleichwertigkeit der Kompetenzbereiche sowie seine Handlungsorientierung (die Entwicklung kommunikativer Kompetenz). Die Lernenden sollen demnach durch den Unterricht befähigt werden, Alltagssituationen kommunikativ erfolgreich und sozial angemessen zu bewältigen.

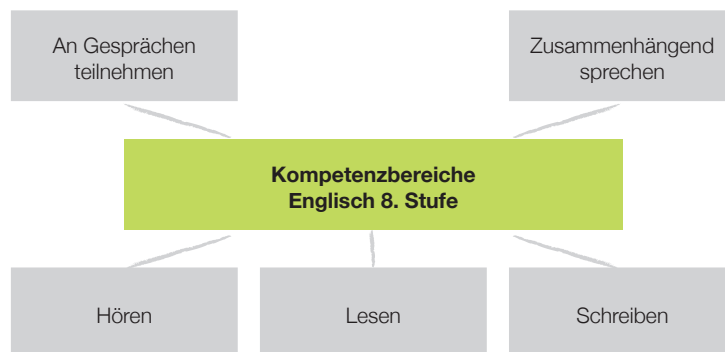


Abbildung 7: Kompetenzmodell Englisch 8. Schulstufe

Detaillierte Informationen zum Kompetenzmodell für Englisch, zu den *Bildungsstandards* und Aufgabenbeispielen finden Sie unter www.bifie.at/node/1822.

Kontextfragebögen

In Schülerleistungsstudien werden zumeist neben Leistungstests, welche die kognitiven Fähigkeiten der Schüler/innen erfassen, Kontextfragebögen eingesetzt. Aufgabe der Kontextfragebögen ist es, die Rahmenbedingungen, unter denen Lehren und Lernen stattfindet, zu erheben. Wesentliche Hintergrundinformationen umfassen individuelle, demografische und sozioökonomische Aspekte auf Schülerebene sowie Merkmale des Unterrichts und Standortbedingungen auf Schulebene. In Verbindung mit den Leistungstests der Schüler/innen ist es möglich, die Ergebnisse so zu kommunizieren, dass Qualitätssicherungs- und -entwicklungsprozesse unterstützt werden. Des Weiteren können dadurch auf Systemebene Rahmenbedingungen analysiert und mögliche Einflussfaktoren dargestellt werden. Im Rahmen der *Standardüberprüfungen* in Englisch auf der 8. Schulstufe wurden ein Schul- und ein *Schülerfragebogen* eingesetzt, welche auf der BIFIE-Homepage im Downloadbereich zur Verfügung stehen. Die Themenfelder der *Fragebögen* finden Sie unter www.bifie.at/node/2119.

Kriteriale Rückmeldung

Die *Verordnung zu den Bildungsstandards* im Schulwesen sieht vor, dass der Grad der Kompetenzerreichung durch die Schüler/innen gemessen und rückgemeldet wird. Eine Schülerleistung wird nach einem bestimmten Kriterium bewertet, nämlich danach, welches *GERS-Referenzniveau* erreicht wurde. Diese Rückmeldung anhand vorher definierter *Referenzniveaus* nennt sich „kriteriale Rückmeldung“.

L

Leistungsabstand der mittleren 75 %

Als Leistungsabstand der mittleren 75 % wird der Abstand zwischen den besten 12,5 % und den schwächsten 12,5 % der Schüler/innen einer Schule/Gruppe bezeichnet. In ihm liegen also die mittleren 75 % der Leistungen einer Schule/Gruppe. Je größer der Abstand, desto heterogener sind die Leistungen bzw. je kleiner der Abstand, desto homogener sind die Leistungen. Dieses *Streuungsmaß* gewährleistet eine bessere Vergleichbarkeit als beispielsweise die gesamte Spannweite, in der auch extrem gute oder extrem schwache Schüler/innen vorkommen (können). Die Spannweite ergäbe sich aus der Leistungsdifferenz zwischen dem/der besten und schlechtesten Schüler/in.

Als *Referenzwert* dient der Median (mittlere Wert) der Leistungsabstände der mittleren 75 % aller österreichischen Schulen, d. h. in der Hälfte aller Schulen ist der Leistungsabstand kleiner als der angegebene Wert, in der

anderen Hälfte ist er größer. Ein Beispiel zur Ermittlung des Medians ist im Glossareintrag zur *mittleren Differenz* zu finden.

Lesen

Das Leseverständnis wurde mittels kürzerer (max. 35 Wörter) und längerer (max. 220 Wörter) Texte überprüft. Dabei wurden in der Regel 16 Items pro *Testheft* in einem Zeitraum von 35 Minuten bearbeitet. Zu den Lesetätigkeiten gehörten unter anderem:

- einfache persönliche Briefe, Karten oder E-Mails verstehen,
- einfache, alltägliche Anleitungen und Vorschriften verstehen,
- vertrauten Alltagstexten (z. B. Schildern und Broschüren – auch aus dem Internet) die wichtigsten Informationen entnehmen,
- einfachen Texten in Zeitungen und Zeitschriften das Wesentliche entnehmen und Schlussfolgerungen ziehen (z. B. Absicht der schreibenden Person),
- unkomplizierte Sachtexte und einfache literarische Texte (z. B. Gedichte) verstehen.

Sämtliche Lesebeispiele kamen dabei aus einem Themenbereich, der für das Alter der Schüler/innen angemessen war. Im Rahmen der *Standardüberprüfung* wurde der *Kompetenzbereich Lesen* in die folgenden *Teilkompetenzen (Sektionen)* untergliedert:

Beschreibungen der einzelnen Sektionen im Kompetenzbereich Lesen	
Allgemeines Leseverständnis (Identifying gist/formality/function/setting)	Dieser Bereich umfasst jene Aufgaben, in denen nach wesentlichen Informationen – wie etwa dem Thema, einer Person oder der Textsorte – gefragt wird.
Überprüfen von vorab definierten Details (Reading for pre-specified information)	Schüler/innen zeigen bei diesen Aufgaben die Kompetenz, Informationen aus einem vorgegebenen Satz mit den relevanten Informationen eines Lesetextes abgleichen zu können.
Detailliertes Leseverständnis (Reading for detailed comprehension, including attitude, opinion & writer purpose)	Aufgaben aus diesem Bereich überprüfen das Leseverständnis im Detail. Hierfür können auch Schlussfolgerungen notwendig sein, etwa wenn die Einstellung bzw. Meinung einer Person oder der Zweck des Schreibens ermittelt werden soll.
Verständnis der Textstruktur (Making meaning: coherence and cohesion; Understanding text structure)	Aufgaben aus diesem Bereich überprüfen, ob Schüler/innen die logische Organisation eines Textes verstehen, indem sie jeweils entweder eine Phrase oder genau ein Wort passend in die vorhandenen Lücken integrieren können.

Abbildung 8: Inhaltliche Beschreibung der Teilkompetenzen (Sektionen) – Kompetenzbereich Lesen

M

Migrationshintergrund

Die Definition des Begriffs erfolgt in Anlehnung an jene der OECD, welche als Kriterium das Geburtsland der Eltern und nicht die Sprachgewohnheiten heranzieht. Ein Kind gilt demnach als Schüler/in mit Migrationshintergrund, wenn beide Elternteile im Ausland geboren wurden. Als Schüler/in ohne Migrationshintergrund wird ein Kind bezeichnet, wenn mindestens ein Elternteil in Österreich geboren wurde.

Einzige Ausnahme für diese Regel bilden Schüler/innen, deren Eltern (ein Elternteil oder beide) in Deutschland geboren wurden – sie werden aufgrund der gleichen Sprache zur Gruppe der Schüler/innen ohne Migrationshintergrund gezählt.

Mittelwert

Der Mittelwert wird häufig auch als Durchschnitt oder in der Statistik als arithmetisches Mittel bezeichnet. Zur Berechnung wird die Summe aller Werte durch deren Anzahl dividiert.

Mittlere Differenz

Die mittlere (Leistungs-)Differenz gibt an, wie stark sich zwei Gruppen (z. B. Burschen und Mädchen) im Mittel voneinander unterscheiden. Für den Schulbericht wird diese mittlere Differenz über alle österreichischen Schulen ermittelt, für die Lehrerrückmeldung über alle österreichischen *Unterrichtsruppen*. Unter „im Mittel“ wird hier der mittlere Wert (Median) verstanden, d. h. es werden die Differenzen für alle Schulen/*Unterrichtsruppen* berechnet, der Größe nach gereiht und dann der Wert jener Schule/*Unterrichtsruppe* bestimmt, der in dieser Rangfolge genau in der Mitte liegt. Die Berechnung dieser mittleren Differenz soll anhand des nachfolgenden Beispiels näher erläutert werden.

In einem ersten Schritt wurden die Geschlechterdifferenzen für jede der fünf aufgelisteten Schulen bestimmt. Dabei ist es wichtig, die Differenz immer gleich zu berechnen (z. B.: immer Burschenmittelwert abzgl. Mädchenmittelwert). Anschließend wurden die Geschlechterdifferenzen (Mittelwertdifferenz) dieser (fiktiven) Schulen der Größe nach gereiht. Durch die Reihung der Mittelwertdifferenzen kann in weiterer Folge jene Differenz bestimmt werden, die in dieser Rangfolge genau in der Mitte liegt. Im angeführten Beispiel entspricht dies der Geschlechterdifferenz von Schule NU mit 13 Punkten zugunsten der Mädchen (Burschen-MW abzgl. Mädchen-MW ergibt –13; vgl. Tabelle 3). Somit liegen in der Hälfte dieser Schulen die Mädchen mehr als 13 Punkte vor den Burschen, in der anderen Hälfte ist der Leistungsvorsprung der Mädchen kleiner als 13 Punkte oder die Burschen sind besser als die Mädchen.

Schule	Mittelwert Burschen	Mittelwert Mädchen	Mittelwertdifferenz (der Größe nach gereiht)
Schule XY	477	507	–30 Pkt. (Mädchen besser)
Schule AB	405	425	–20 Pkt. (Mädchen besser)
Schule NU	607	620	–13 Pkt. (Mädchen besser)
Schule DE	480	480	0 Pkt.
Schule JH	582	597	13 Pkt. (Burschen besser)

Schulen mit noch größerem Leistungsvorsprung der Mädchen

mittlere Differenz

Schulen mit kleinerem Leistungsvorsprung der Mädchen oder Vorsprung der Burschen

Tabelle 3: Berechnung der mittleren Differenz (fiktives Beispiel)

Diese Berechnung wird zur Ermittlung der mittleren Differenz aller österreichischen Schulen/*Unterrichtsruppen* angewendet. Damit kann verglichen werden, wie sich der Unterschied zwischen zwei Gruppen in der Schule/*Unterrichtsruppe* von der mittleren Differenz in österreichischen Schulen/*Unterrichtsruppen* unterscheidet. Generell gilt: Je geringer die Gruppenunterschiede, desto höher die Chancengleichheit für die einzelnen Gruppen.

N

n

n ist die Anzahl der getesteten Schüler/innen in der Schule/*Unterrichtsgruppe*. Diese Anzahl kann von der tatsächlichen Anzahl der Schüler/innen in der Schule/*Unterrichtsgruppe* aus verschiedenen Gründen abweichen:

- Einige Schüler/innen der Schule/*Unterrichtsgruppe* waren von der *Überprüfung* ausgenommen (z. B. Schüler/innen mit sonderpädagogischem Förderbedarf). Detaillierte Informationen zu diesen Kriterien finden Sie in der Novelle der Verordnung (BGBl. II 282/2011) unter www.bifie.at/node/48.
- Zudem kann es vorkommen, dass nicht alle Schüler/innen der Schule/*Unterrichtsgruppe*, die zur *Überprüfung* zugelassen waren, bei der Testung anwesend waren.

R

Rater

Raters sind Personen, die die von den Schülerinnen und Schülern verfassten Texte im *Kompetenzbereich Schreiben* nach einer vorgegebenen Bewertungsskala (*Writing Rating Scale*) beurteilen. Die Texte wurden hierfür direkt nach der *Standardüberprüfung* an das BIFIE retourniert, gescannt und den jeweiligen Raters in personalisierten Booklets zur Verfügung gestellt. Jeder Rater bewertete bis zu 600 Schülertexte (ca. 300 kurze und 300 lange Texte). Raters sind erfahrene Lehrer/innen, welche in einer mehrphasigen Fortbildung für ihre Tätigkeiten vom BIFIE geschult wurden.

Referenzwerte

Bei der Rückmeldung der Ergebnisse der *Standardüberprüfung* werden zusätzlich zu den erzielten Leistungen (z. B. der Schülerin/des Schülers, der *Unterrichtsgruppe* oder der Schule) Werte angegeben, mit denen die Ergebnisse verglichen werden können. Welche Werte das sind, richtet sich nach der Zielgruppe. Es wird meist sowohl der *Mittelwert* aller getesteten Schüler/innen Österreichs als auch ein *Erwartungsbereich* für die Schule/*Unterrichtsgruppe* als Referenz (Vergleich) angegeben, zusätzlich können aber auch *mittlere Differenzen* zwischen Subgruppen (z. B. zwischen Mädchen und Burschen) ausgewiesen werden.

Rückmeldemoderatorinnen/-moderatoren (RMM)

RMM unterstützen die Schulleiter/innen sowie die Lehrer/innen bei der sachlichen Analyse und objektiven Interpretation der Ergebnisse aus den *Standardüberprüfungen* und bei einer faktenbasierten Ergebnisaufarbeitung (z. B. Erstellen eines Stärken-Schwächen-Profiles, Identifizierung von Handlungsfeldern). Sie helfen den Schulen dabei, sich der Chancen der Ergebnismeldung bewusst zu werden und Qualitätsentwicklungspotenzial, welches sich aus der Ergebnismeldung ergibt, zu erkennen. Die konkrete nachfolgende Schul- und Unterrichtsentwicklung liegt in der Verantwortung der Schulleiter/innen und Lehrer/innen. Schulleiter/innen haben die Möglichkeit, ausgebildete RMM über die Pädagogischen Hochschulen anzufordern.

S

Schreiben

Im *Kompetenzbereich Schreiben* verfassten die Schüler/innen zwei Texte unterschiedlicher Länge. Der kürzere Text sollte 40 bis 70 Wörter umfassen und war in zehn Minuten fertigzustellen. Der längere Text sollte 120 bis

180 Wörter umfassen und war in 20 Minuten zu erstellen. Die inhaltlichen Punkte, die zu bearbeiten waren, wurden in der Aufgabenstellung vorgegeben. Nach den 30 Minuten Arbeitszeit standen den Schülerinnen und Schülern fünf Minuten für die Überarbeitung der beiden Texte zur Verfügung. Die von den Schülerinnen und Schülern verfassten Texte wurden von sogenannten „Raters“ mithilfe einer vorgegebenen Bewertungsskala (*Writing Rating Scale*) beurteilt. Beim *Kompetenzbereich Schreiben* werden folgende *Teilkompetenzen* (Dimensionen) eines Textes bewertet:

Beschreibungen der einzelnen Dimensionen im Kompetenzbereich Schreiben	
Erfüllung der Aufgabenstellung	Das Kriterium <i>Erfüllung der Aufgabenstellung (Task Achievement)</i> bewertet die inhaltliche Leistung bei der Bearbeitung einer Aufgabenstellung. Dazu gehören vor allem die Erwähnung und Bearbeitung aller inhaltlichen Punkte (einige davon im Detail) sowie die Wahl der richtigen Textsorte und das Erreichen der geforderten Wortanzahl.
Struktur & Aufbau (Kohärenz & Kohäsion)	Das Kriterium <i>Struktur & Aufbau (Coherence & Cohesion)</i> bewertet den logischen und verständlichen Aufbau eines Textes. Für eine gute Beurteilung in diesem Bereich müssen einzelne Sätze z. B. durch die Verwendung von Verknüpfungsmitteln gut miteinander verbunden sein und längere Abschnitte (Absätze) eine nachvollziehbare und klare Struktur aufweisen. Längere Texte sind durch Absätze zu gliedern.
Grammatik	Im Bereich <i>Grammatik (Grammar)</i> liegt der Schwerpunkt auf der Verwendung vielfältiger grammatikalischer Strukturen. Dabei wird vor allem ein größeres grammatikalisches Repertoire positiv beurteilt. Inkorrekt verwendete Strukturen fließen vor allem dann negativ in die Bewertung ein, wenn sie zu Missverständnissen führen oder die Kommunikation nachhaltig stören oder verhindern.
Wortschatz	Im Bereich <i>Wortschatz (Vocabulary)</i> liegt der Schwerpunkt auf der Anwendung eines für das Thema angemessenen, „möglichst breiten“ Wortschatzes. Der verwendete Wortschatz soll eine erfolgreiche Kommunikation ermöglichen, sodass – abhängig von der Komplexität der Botschaft – Missverständnisse oder gar der Abbruch der Kommunikation vermieden werden.

Abbildung 9: Inhaltliche Beschreibung der Teilkompetenzen (Dimensionen) – Kompetenzbereich Schreiben

Schülerfragebogen

Siehe unter *Kontextfragebögen*

Sozialstatus

Aus den Angaben zur Bildung und beruflichen Stellung von Vater und Mutter sowie der Bücheranzahl im Haushalt (erhoben im *Schülerfragebogen*) wird ein Sozialstatus-Index gebildet. Dieser Index dient als Grundlage für weitere Berechnungen und fließt in den *fairen Vergleich* ein. Zur anschaulichen Darstellung dieser Verteilung in der Rückmeldung werden aus dem Index drei Sozialstatusgruppen gebildet. Dabei werden alle getesteten Schüler/innen in Österreich nach ihrem Sozialstatus gereiht und anschließend wie folgt unterteilt:

- Die 25 % mit dem niedrigsten Sozialstatus (unteres Viertel der österr. Verteilung)
- Die mittleren 50 % (= Interquartilabstand; mittlere 50 % der österr. Verteilung)
- Die 25 % mit dem höchsten Sozialstatus (oberes Viertel der österr. Verteilung)

Diese Kategoriegrenzen werden also ausschließlich nach statistischen Kriterien festgelegt. Zur Beschreibung der Sozialstatusverteilung einer *Schule/Unterrichtsguppe* werden die österreichweiten Kategoriegrenzen herangezogen. So können die Prozentanteile in den einzelnen Sozialstatusgruppen mit der österreichweiten Verteilung verglichen werden. Wenn z. B. 30 % der Schüler/innen einer *Schule/Unterrichtsguppe* dem „oberen

Viertel der österr. Verteilung“ zugeordnet sind, würde dies bedeuten, dass in dieser Schule prozentuell etwas mehr Schüler/innen zur obersten Sozialstatusgruppe zählen als in Österreich insgesamt.

Sprechen

Die *Überprüfung* des Kompetenzbereichs Sprechen fand aus organisatorischen Gründen nur an einer Teilstichprobe von 120 Schulen statt. Die Schüler/innen traten paarweise zur *Überprüfung* an. Das Gespräch wurde von einem sogenannten „*Interlocutor*“ geleitet, ein „*Assessor*“ beurteilte die Schülerleistungen anhand einer vorgegebenen Bewertungsskala (*Assessment Scale*). Das Gespräch begann mit einem kurzen Interview, das als Aufwärmphase gedacht war und nicht in die Beurteilung einbezogen wurde.

Im ersten Testteil, dem monologischen Teil erhielt jede Schülerin/jeder Schüler drei Themen zur Wahl und sprach nach einer Vorbereitungszeit von einer Minute, in der keine Notizen gemacht werden durften, ca. zwei bis drei Minuten lang über das gewählte Thema. Der zweite, dialogische Teil der *Überprüfung* war zweiteilig. Hier wurde ein schriftlicher und/oder bildlicher Impuls gegeben. Die beiden Schüler/innen sprachen miteinander über ein vorgegebenes Thema. Im ersten Dialog sollten die Schüler/innen ein einfaches Alltagsgespräch führen (z. B. Vereinbarungen treffen). Der zweite Dialog war ein ausführlicheres Gespräch, das auch den Austausch von Meinungen verlangte. Die Vorbereitungszeit für den zweiten Teil betrug eine Minute, die gemeinsame Sprechzeit vier bis fünf Minuten. Beim *Kompetenzbereich* Sprechen werden folgende *Teilkompetenzen* (Dimensionen) bewertet:

Beschreibungen der einzelnen Dimensionen im Kompetenzbereich Sprechen	
Erfüllung der Aufgabenstellung und Aufrechterhaltung der Kommunikation	Das Kriterium <i>Erfüllung der Aufgabenstellung und Aufrechterhaltung der Kommunikation (Task Achievement & Communication Skills)</i> bewertet die erfolgreiche Erfüllung der Aufgabenstellung. Hohe Kompetenz zeigt sich z. B. beim monologischen Sprechen in der Fähigkeit, die erforderlichen Informationen zu geben und wichtige Punkte durch zusätzliche Details ergänzen zu können. Beim dialogischen Sprechen steht neben der Erfüllung der Aufgabe auf funktionaler Ebene vor allem die erfolgreiche Interaktion im Fokus der Bewertung. Dazu gehören neben der Erfüllung des Auftrags das Aufgreifen von Inhalten des Gesprächspartners bzw. der Gesprächspartnerin, aktiver Sprecherwechsel oder das Beenden von Gedankengängen.
(klare & natürliche) Sprache	Für eine gute Beurteilung im Bereich <i>(klare & natürliche) Sprache (Clarity & Naturalness of Speech)</i> müssen die Schüler/innen zeigen, dass sie zusammenhängend und verständlich sprechen können. Bewertet wird hier vor allem die Fähigkeit, längere und zusammenhängende Sprechakte flüssig und mit klarer und natürlicher Aussprache und Intonation produzieren zu können.
Grammatik	Im Bereich <i>Grammatik (Grammar)</i> liegt der Fokus auf der Verwendung vielfältiger und situationsadäquater grammatikalischer Strukturen. Dabei wird ein breites strukturelles Repertoire positiv beurteilt. Die Sprachrichtigkeit wird unter besonderer Berücksichtigung der Charakteristika gesprochener Sprache (z. B. kurze Einheiten, unvollständige Sätze, Fehler durch syntaktische Überlagerungen während des Sprechens) bewertet. Inkorrekt verwendete Strukturen fließen nur dann in die Bewertung ein, wenn sie zu Missverständnissen führen oder die Kommunikation stören oder verhindern.
Wortschatz	Der Fokus der Beurteilung vom <i>Wortschatz (Vocabulary)</i> liegt auf dem Einsatz eines „reichhaltigen“ Wortschatzes. Bewertet wird die Fähigkeit, ein der gesprochenen Sprache entsprechendes lexikalisches Repertoire situationsadäquat zu verwenden. Inkorrekt verwendetes Vokabular fließt nur dann in die Bewertung ein, wenn es zu Missverständnissen führt, die Kommunikation stört oder sogar verhindert.

Abbildung 10: Inhaltliche Beschreibung der Teilkompetenzen (Dimensionen) – Kompetenzbereich Sprechen

Standardüberprüfung bzw. Bildungsstandardüberprüfung

Seit dem Schuljahr 2011/12 finden auf der 8. Schulstufe in den Gegenständen Deutsch, Englisch und Mathematik periodische *Standardüberprüfungen* statt. Auf der 4. Schulstufe werden diese Tests ab dem Schuljahr 2012/13 in Mathematik und Deutsch/Lesen/Schreiben durchgeführt. Dabei werden Lernergebnisse objektiv festgestellt und mit den angestrebten Standards verglichen.

Im Vorfeld wurden auf der 8. Schulstufe im Frühjahr 2009 *Baseline-Testungen* für Deutsch, Englisch und Mathematik durchgeführt. Die Ergebnisse dieser Testung dienen als Ausgangspunkt für die Beobachtung der Entwicklung der Schülerleistungen. Für die 4. Schulstufe fand die *Baseline-Testung* im Frühjahr 2010 statt.

Weitere Informationen zur Standardüberprüfung finden Sie auf der BIFIE-Homepage unter www.bifie.at/standardueberpruefung.

Standard Setting

Beim Standard Setting handelt es sich um ein komplexes Verfahren zur Bestimmung von theoretisch wie auch empirisch festgelegten *Referenzniveaus*, die in Englisch auf Basis der *GERS-Referenzniveaus* beschrieben werden. Die kontinuierliche Punktskala in den einzelnen *Kompetenzbereichen* der *Standardüberprüfung* in Englisch wird jeweils durch zwei Schwellenwerte in drei *GERS-Referenzniveaus* unterteilt:

- B1 oder darüber
- A2
- A1 oder darunter

Die Festlegung der Schwellenwerte, die den Übergang zwischen den *GERS-Referenzniveaus* markieren, wird als Standard Setting bezeichnet. Für die *GERS-Referenzniveaus* wurden in einem ersten Schritt inhaltliche Beschreibungen für die Niveaus A2 und B1 oder darüber entwickelt. Um die Schüler/innen auf Basis ihrer erbrachten Testleistung den *GERS-Referenzniveaus* zuordnen zu können, wurde in weiterer Folge eine repräsentative Auswahl an verwendeten *Testitems* (bei *Hören* und *Lesen*) und Schülerperformanzen (produzierte Texte bei *Schreiben* bzw. Videos bei *Sprechen*) durch Expertinnen und Experten aus Schulpraxis (Sekundarstufe I und II), Fachdidaktik, Pädagogik und Psychologie den verschiedenen *Referenzniveaus* zugeordnet. Die Schwierigkeiten der *Items* und Schülerperformanzen wurden dabei auf derselben Punktskala gemessen wie die Leistungen. Daher ließen sich Schwellenwerte auf der Punktskala bestimmen, die die Referenzniveaus voneinander abgrenzen.

Streuung

Als Streuung wird die Verteilung von einzelnen Werten um den *Mittelwert* bezeichnet. Die Werte 450, 520 und 530 haben beispielsweise denselben *Mittelwert* (500) wie die Werte 350, 420 und 730, sie verteilen sich aber unterschiedlich weit um ihn herum. Liegen die einzelnen Werte sehr dicht am *Mittelwert*, spricht man von einer kleinen Streuung bzw. homogenen Verteilung, liegen sie weit entfernt von ihm, dann handelt es sich um eine große Streuung bzw. heterogene Verteilung. Als Maß für diese Streuung wird in der vorliegenden Rückmeldung der *Leistungsabstand der mittleren 75 %* angeführt.

Im Rahmen der *Baseline-Testungen* wurden die Ergebnisse so auf die Skala transformiert, dass österreichweit 2/3 aller Schüler/innen einen Wert zwischen 400 und 600 bzw. je 1/6 aller getesteten Schüler/innen ein Ergebnis über 600 bzw. weniger als 400 erzielten.

Bei der *Standardüberprüfung* in Englisch auf der 8. Schulstufe erreichten 2/3 aller getesteten Schüler/innen ein Ergebnis zwischen 410 und 600 Punkten.

T

Teilkompetenzen

Als Teilkompetenzen (Sektionen bzw. Dimensionen) werden die in der Rückmeldung zur *Standardüberprüfung* ausgewiesenen Unterkategorien eines *Kompetenzbereichs* verstanden. Welche Teilkompetenzen/Sektionen eines *Kompetenzbereichs* in der Rückmeldung ausgewiesen werden, kann den Glossareinträgen zu den einzelnen *Kompetenzbereichen Hören, Lesen, Schreiben* und *Sprechen* entnommen werden.

Testhefte/Testformen

Die Testhefte sind das Kernelement der *Standardüberprüfung*, da auf Basis der Leistung im Test für jede Schülerin/jeden Schüler der individuelle Grad der Kompetenzerreichung ermittelt wird. Damit jeder *Kompetenzbereich* auf Systemebene möglichst breit erfasst und auch die Problematik des Abschreibens möglichst gering gehalten wird, werden verschiedene Testformen mit unterschiedlichen *Items* erstellt. Die Verwendung vieler *Items* ist notwendig, um alle *Kompetenzen* entsprechend breit abdecken zu können. Die individuelle Testzeit ist jedoch so kurz wie möglich gehalten, weshalb nicht alle *Items* von jeder Schülerin und jedem Schüler bearbeitet werden. Alle Testhefte, die dieselben *Items* in derselben Reihenfolge beinhalten, entsprechen also einer Testform. Bei der Zusammenstellung der Testformen wird berücksichtigt, dass alle Testformen möglichst gleich schwer sind und eine ähnliche Anzahl an offenen, halboffenen und geschlossenen *Items* beinhalten. Bei der Zuordnung der Testhefte in den *Kompetenzbereichen Lesen* und *Schreiben* wird darauf geachtet, dass möglichst viele verschiedene Testformen innerhalb einer Klasse verwendet werden.

Testitem

Siehe unter *Item*

U

Überprüfung

Siehe unter *Standardüberprüfung*

Unterrichtsgruppe

Die Unterrichtsgruppe ist die Einheit, in der die Schüler/innen im jeweiligen Testfach gemeinsam unterrichtet werden. Wird eine Klasse beispielsweise im Fach Englisch in zwei Gruppen A und B unterrichtet, so sind diese beiden Gruppen die jeweiligen Unterrichtsgruppen der Schüler/innen.

Werden an einer Hauptschule Schüler/innen mehrerer Leistungsgruppen gemeinsam unterrichtet, so wird diese Gruppe als gemischte Unterrichtsgruppe bezeichnet. In den Rückmeldungen an die Lehrer/innen dieser gemischten *Unterrichtgruppen* werden zusätzliche Auswertungen getrennt nach Leistungsgruppen dargestellt, sofern:

- jede Schülerin/jeder Schüler dieser gemischten Unterrichtsgruppe einer bestimmten Leistungsgruppe zugeordnet wurde (vollständige Angaben sind erforderlich) und
- sich in keiner Leistungsgruppe dieser gemischten Unterrichtsgruppe nur ein/e Schüler/in befindet (Datenschutz).

V

Verordnung zu den Bildungsstandards im Schulwesen

Eine Novellierung des Schulunterrichtsgesetzes vom August 2008 legte die rechtliche Grundlage für die Einführung von *Bildungsstandards* im österreichischen Schulwesen. Diese erfolgte durch eine entsprechende Verordnung über Bildungsstandards im Schulwesen im Jänner 2009 (BGBl. II Nr. 1/2009) sowie deren Novellen im August 2011 (BGBl. II Nr. 282/2011) und Mai 2012 (BGBl. II Nr. 185/2012).

Zur Verordnung über Bildungsstandards und der entsprechenden Novelle gelangen Sie über folgenden Link: www.bifie.at/node/48.

In der Anlage zur Verordnung finden Sie die *Bildungsstandards* für die 4. Schulstufe der Volksschulen in Deutsch/Lesen/Schreiben und Mathematik bzw. für die 8. Schulstufe der Volksschuloberstufen, der Hauptschulen, der Neuen Mittelschulen und der allgemeinbildenden höheren Schulen in Deutsch, erster lebender Fremdsprache (*Englisch*) und Mathematik.

Vertrauensintervall

Das Vertrauensintervall ist ein Wertebereich, der um das Testergebnis der Schule/*Unterrichtsgruppe* angegeben ist und in dem die wahre Leistung der Schule/*Unterrichtsgruppe* mit 90 % Sicherheit liegt. Dieser Wertebereich wird angegeben, da das Testergebnis aufgrund des Messfehlers einer Messung möglicherweise nicht exakt der wahren Leistung entspricht. Ein messfehlerfreies Testen wäre nur möglich, wenn unendlich viele verschiedene *Items* eingesetzt würden. Da das nicht möglich ist, werden die *Testhefte* nach bestimmten Kriterien zusammengestellt. Zum einen müssen sie in einer zumutbaren Zeit zu bearbeiten sein. Zum anderen muss sichergestellt werden, dass ausreichend Informationen vorliegen, um zuverlässige und informative Ergebnismeldungen an alle Zielgruppen geben zu können. Bei der Zusammenstellung der *Testhefte* wird auf eine ausgewogene Verteilung der *Items* auf die *Testformen* nach *Antwortformaten* geachtet. Trotzdem könnte es sein, dass bei einer anderen Auswahl an Aufgaben die Testergebnisse unter Umständen etwas besser oder schlechter ausfallen.

Die Unsicherheit, die durch den Messfehler bedingt ist, spiegelt sich in der Breite des Vertrauensintervalls wider. Je größer das Vertrauensintervall, desto größer die Unsicherheit. Ein Ergebnis von 525 (± 7) Punkten würde demnach bedeuten, dass die wahre Leistung der Schülergruppe mit 90 % Sicherheit zwischen 518 und 532 Punkten liegt. Die Unsicherheit ist dabei abhängig von der Anzahl (n) der getesteten Schüler/innen sowie der eingesetzten *Items* (je mehr Schüler/innen und/oder *Items*, desto zuverlässiger ist die Messung der Kompetenz) und der *Streuung* der Testwerte (je kleiner die *Streuung*, d. h. je homogener die Leistungsverteilung, desto zuverlässiger die Messung).

Wann unterscheiden sich Gruppen (Schülergruppe, Klasse etc.) voneinander? Überlappen die Vertrauensintervalle zweier Gruppen nicht, dann unterscheiden sich die tatsächlichen Ergebnisse dieser Gruppen substantiell voneinander und Unterschiede sind nur mit sehr geringer Wahrscheinlichkeit durch den Messfehler bedingt. Bei überlappenden Vertrauensintervallen kann nicht mit Sicherheit gesagt werden, dass sich diese Gruppen voneinander unterscheiden, da eine Abweichung zwischen den Testergebnissen eventuell vollständig auf den Messfehler zurückzuführen sein könnte (vgl. Abbildung 11).

Bei der Interpretation des fairen Vergleichs ist das Ergebnis einer Schule bzw. *Unterrichtsgruppe* dem *Erwartungsbereich* ohne Berücksichtigung des Vertrauensintervalls gegenüberzustellen.

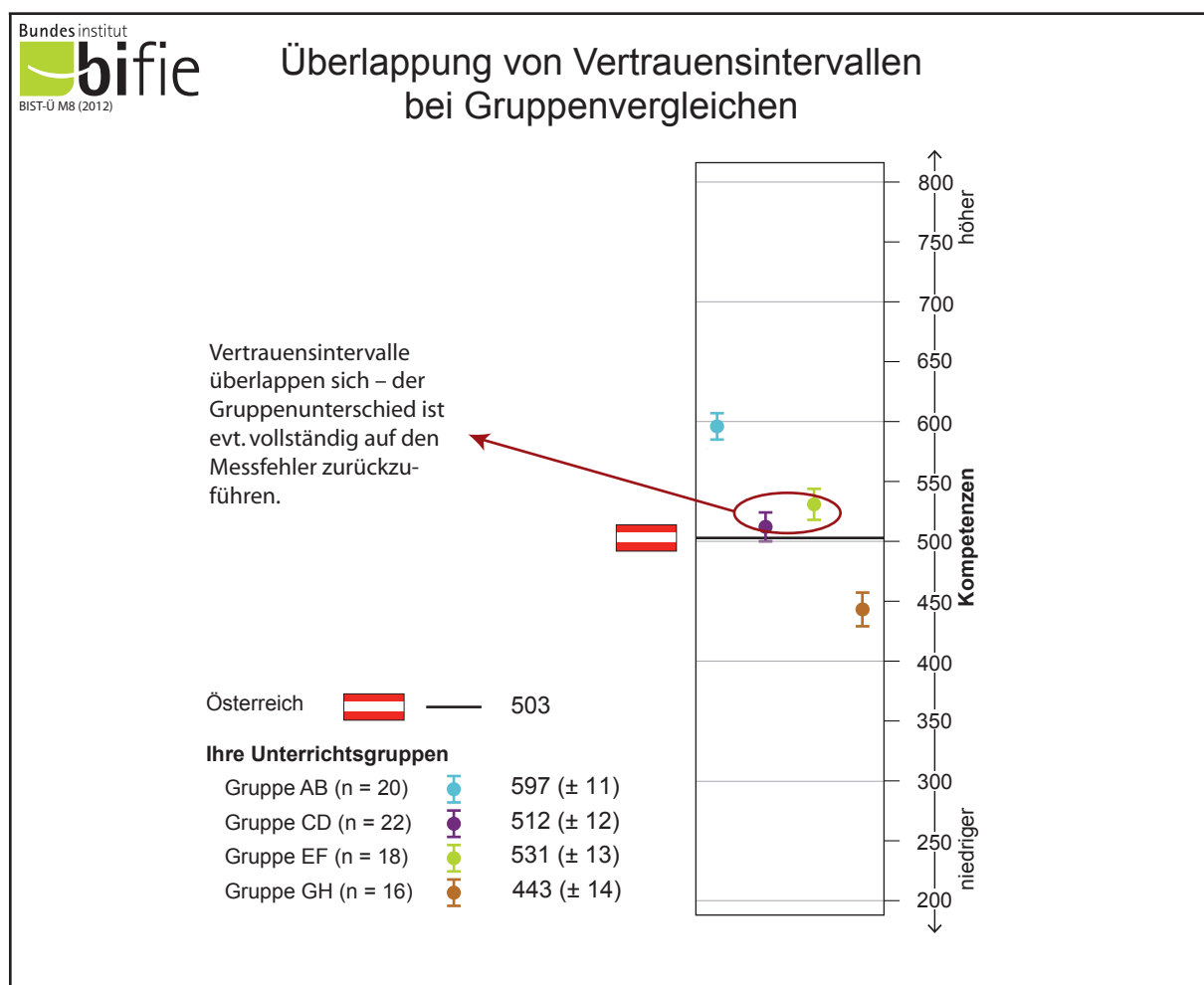


Abbildung 11: Überlappung von Vertrauensintervallen

W**Writing Rating Scale (Beurteilungsskala für Schreiben)**

Zur *Überprüfung* der Schreibkompetenz in Englisch verfassten die Schüler/innen zwei Texte unterschiedlicher Länge. Die Beurteilung dieser Texte erfolgt durch speziell ausgebildete Personen, sogenannte *Raters*. Sie beurteilen die Texte in den vier *Teilkompetenzen* (Erfüllung der Aufgabenstellung, Struktur und Aufbau, Grammatik und Wortschatz) des *Kompetenzbereichs Schreiben*. Für jede der vier *Teilkompetenzen* werden dabei zwischen 0 und 7 Punkte vergeben. Die Writing Rating Scale mit einer inhaltlichen Beschreibung der einzelnen Punkte finden Sie unter www.bifie.at/node/1826.

Z**Zusammenhängend sprechen**

Siehe unter *Sprechen*

