



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG



**MEZ**  
Mehrsprachigkeitsentwicklung  
im Zeitverlauf

# MEZ Arbeitspapiere

Thorsten Klinger, Anouk Ticheloven, Irina Usanova

**Das Konstrukt und die Messung produktiver Sprachfähigkeiten in den Projekten „Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“ und „Mehrsprachigkeit an der Schwelle zum Beruf (MEZ-2)“**

MEZ-Arbeitspapier Nr. 5  
Hamburg, März 2024

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

Arbeitspapiere der Forschungsprojekte Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ) und Mehrsprachigkeit an der Schwelle zum Beruf (MEZ-2) an der Universität Hamburg

In der Reihe erscheinen Beiträge zu Themen, die den Arbeitsprozess des MEZ-Projekts betreffen. Die Beiträge stellen vorläufige Ergebnisse zu den theoretischen und empirischen Fragestellungen des Projekts sowie Lösungen zum technischen Umgang mit den Projektdaten vor.

Die in den *MEZ-Arbeitspapieren* vertretenen Meinungen sind die der Autor(inn)en und entsprechen nicht notwendigerweise den Auffassungen der Projektverantwortlichen.

**Principal Investigator MEZ-2-Projekt:**

Ingrid Gogolin, Universität Hamburg (Projektleitung)

**Bezug:**

[www.mez.uni-hamburg.de](http://www.mez.uni-hamburg.de)

**Kontakt:**

Mehrsprachigkeit an der Schwelle zum Beruf – MEZ-2

Universität Hamburg

Institut für Interkulturelle und International Vergleichende Erziehungswissenschaft

Postanschrift: Von-Melle-Park 8, 20146 Hamburg

Besucheranschrift: Alsterterrasse 1, 5. Stock, 20354 Hamburg

E-Mail: [mez@uni-hamburg.de](mailto:mez@uni-hamburg.de)

Tel.: +49 40 42838-2708

**Bitte zitieren Sie dieses Arbeitspapier wie folgt:**

Klinger, Thorsten; Ticheloven, Anouk; Usanova, Irina (2024): Das Konstrukt und die Messung produktiver Sprachfähigkeiten in den Projekten „Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“ und „Mehrsprachigkeit an der Schwelle zum Beruf (MEZ-2)“. MEZ Arbeitspapier Nr. 5. Hamburg (Universität Hamburg): [www.mez.uni-hamburg.de](http://www.mez.uni-hamburg.de)

Thorsten Klinger, Anouk Ticheloven und Irina Usanova

## Das Konstrukt und die Messung produktiver Sprachfähigkeiten in den Projekten „Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“ und „Mehrsprachigkeit an der Schwelle zum Beruf (MEZ-2)“

MEZ-Arbeitspapier Nr. 5

### Zusammenfassung:

Der Beitrag befasst sich mit dem im Projekt „Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“ entwickelten Konstrukt der mehrsprachigen Schreibfähigkeiten von Schüler(inne)n in der Sekundarstufe. Entsprechend dem Studiendesign spielt dabei der Aspekt der Messung der Schreibfähigkeiten in verschiedenen Sprachen eine wesentliche Rolle. Untersucht wurden Jugendliche mit deutsch-russischem, deutsch-türkischem und monolingual deutschem familialen Sprachhintergrund. In vier Erhebungswellen wurde ihre Schreibfähigkeit in den Sprachen Deutsch (Majoritätssprache), Russisch oder Türkisch (Herkunftssprachen der Familien) sowie Englisch, Französisch und/oder Russisch (Schulfremdsprachen) getestet. Vorgestellt wird ein Konstrukt mehrsprachiger Schreibfähigkeiten, dessen sprach(en)übergreifende Indikatoren sowie deren empirische Operationalisierungen. In einem weiteren Schritt wird der Zusammenhang der Schreibfähigkeiten mit anderen sprachlichen Teilfähigkeiten betrachtet. Untersucht werden Korrelationen mit rezeptiven Fähigkeiten in Deutsch und den Herkunftssprachen, gemessen mit dem Lesegeschwindigkeits- und Verständnistest LGVT 5-12+, mit allgemeinen Sprachfähigkeiten in den untersuchten Fremdsprachen, gemessen mit C-Tests, sowie mit subjektiven Selbsteinschätzungen der eigenen Fähigkeiten in verschiedenen Sprachfähigkeitsbereichen in allen MEZ-Sprachen. Im Ergebnis zeigt sich, dass die sprachlichen Teilfähigkeiten positiv miteinander korrelieren, jedoch in unterschiedlichem Ausmaß.

**Abstract:**

This working paper presents the construct of multilingual writing skills developed within the study "Multilingual Development: A Longitudinal Perspective (MEZ)." The MEZ study design involved measuring the writing skills in various languages of secondary school students in Germany. The sample included adolescents with German-Russian, German-Turkish, and monolingual German heritage language backgrounds. The students' writing skills were assessed in four survey waves in German (majority language), Russian or Turkish (heritage languages of the families), and English, French, or Russian (foreign languages taught in school). The working paper presents the construct of multilingual writing skills, its cross-linguistic indicators, and their empirical operationalizations. In addition, the relationship between writing skills and other language competencies is examined. Here, correlations with receptive competence in German and the heritage languages, measured with the LGVT 5-12+ reading speed and comprehension test, with general language competence in the foreign languages studied, measured with C-tests, as well as with self-assessments of language skills in different domains of proficiency in all MEZ languages are analyzed. The results indicate positive correlations among language competencies, but to varying degrees.

## Inhalt

1	Einleitung.....	7
2	Kompetenzmodell.....	8
2.1	Beschreibung der Indikatoren und ihre Operationalisierung.....	9
3	MEZ-Schreibaufgaben für Jugendliche.....	12
3.1	Die Testversionen der Schreibaufgabe.....	14
3.2	Erhebungsdetails im Rahmen von MEZ.....	15
3.3	Abweichende Erhebungsmodi bei MEZ-2.....	16
3.4	Erhobene Daten.....	17
4	Auswertung.....	18
5	Zusammenhänge mit anderen Sprachdaten.....	19
5.1	Lesegeschwindigkeits- und Verständnistest für die Klassen 5–12 (LGVT 5-12+).....	20
5.2	C-Test.....	21
5.3	Selbsteinschätzung.....	21
6	Fazit.....	23
7	Literaturverzeichnis.....	26
	Anhang.....	32

## Tabellenverzeichnis

Tabelle 1: Operationalisierungen der im MEZ-Projekt verwendeten Indikatoren für Schreibkompetenzen (nach: Gogolin et al. 2022).....	11
Tabelle 2: Verteilung der in MEZ und MEZ-2 verwendeten Schreibaufgaben auf die sechs Messzeitpunkte.....	14
Tabelle 3: Anzahl der in MEZ und MEZ-2 elizitierten Schreibproben (Texte).....	17
Tabelle 4: Inter-Rater-Reliabilität nach Sprache und Impuls - Intraklassen-Korrelationskoeffizienten (ICCs).....	19
Tabelle 5: Korrelationen zwischen MEZ-Sprachdaten am 1. Messzeitpunkt.....	20
Tabelle 6: Korrelationen (Pearsons $r$ ) zwischen Schreibfähigkeiten und Selbsteinschätzungen am 1. Messzeitpunkt.....	22
Tabelle 7: Korrelationen zwischen Schreiben und Selbsteinschätzungen an den Messzeitpunkten 1, 3 und 6.....	32
Tabelle 8: Korrelationen zwischen Selbsteinschätzungen in unterschiedlichen Sprachen am 1. Messzeitpunkt.....	32



## 1 Einleitung

„Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“ ist eine vom Bundesministerium für Bildung und Forschung (BMBF) von 2014 bis 2019 geförderte Panelstudie über mehrsprachige Schülerinnen und Schüler in Deutschland. Die Studie wurde an der Universität Hamburg von einem interdisziplinären Forscherteam (Erziehungswissenschaft, Sprachwissenschaft) durchgeführt. Das Forschungsprojekt (Brandt et al., 2017; Gogolin et al., 2017; Klinger et al., 2022) sollte wesentliche Erkenntnisse über die individuelle Entwicklung mehrsprachiger Kompetenzen liefern und zielte darauf ab, diejenigen Bedingungen zu identifizieren, die sich - positiv oder negativ - auf die Entwicklung und Nutzung der sprachlichen Ressourcen und Potenziale junger Mehrsprachiger auswirken. Der Fokus liegt dabei nicht ausschließlich auf Jugendlichen mit Migrationshintergrund. Auch Jugendliche ohne Migrationshintergrund erwerben Formen der Mehrsprachigkeit, nicht zuletzt durch das Erlernen von Fremdsprachen in der Schule. Das MEZ-Projekt beschäftigt sich daher mit der Frage, wie sich mehrsprachige Kompetenzen, die in einem formalen Kontext in der Schule und/oder im außerschulischen Leben erworben werden, auf den Bildungserfolg von Lernenden mit ein- und mehrsprachigem familiärem Hintergrund auswirken. Die Follow-up-Studie „Mehrsprachigkeitsentwicklung an der Schwelle zum Beruf (MEZ-2)“ verfolgte bei einem Teil der MEZ-Stichprobe die Phase der Vorbereitung auf bzw. Einmündung in den Beruf (Gogolin et al., 2021). Auch diese Fortführung der Untersuchung wurde vom BMBF gefördert (2019 bis 2022). Während die Erhebungen im MEZ-Projekt im Gruppenkontext an allgemeinbildenden Schulen stattfanden (vgl. Abschnitt 3.2), wurden die Daten in MEZ-2 individuell online erhoben (Abschnitt 3.3 und Gogolin et al., 2021).

Die Schlüsselfragen der Studien lauten: (1) Welche sprachlichen, persönlichen und kontextuellen Bedingungen beeinflussen den Erwerb mehrsprachiger Fähigkeiten positiv oder negativ? (2) Wie verändern sich diese Bedingungen im Laufe der Zeit und wie stehen sie in Wechselwirkung zueinander? (3) Welche Rolle spielen andere Dimensionen des Bildungserfolgs (z. B. Übergänge im Bildungssystem, Bildungsentscheidungen)?

Zur Beantwortung dieser Fragen sollten die Sprachkompetenzen von Schüler(inne)n mit deutsch-russischem, deutsch-türkischem und einsprachig deutschem Sprachhintergrund untersucht werden. Dabei galt das Augenmerk der Untersuchungen an der Gesamtstichprobe der Entwicklung literaler Sprachfähigkeiten. Die Konzentration auf das Medium der Schriftlichkeit gründet auf der im Verlaufe der Bildungskarriere zunehmenden Bedeutung schriftsprachlicher Kompetenz im schulischen und im gesellschaftlichen Leben (Schründer-Lenzen, 2007, 14; vgl. auch Schneider und Tibken, 2018).

Einer der Schwerpunkte der Paneluntersuchung bestand – neben der Berücksichtigung von Leseverstehen und allgemeinen Sprachfähigkeiten – in der Analyse von *Schreibfähigkeiten* von Sekundarschüler(inne)n in verschiedenen Sprachen. Wie an anderer Stelle ausführlicher dargelegt (vgl. hierzu und zum folgenden Gogolin und Klinger, 2022) wurde in der Vergangenheit die Messung produktiver literaler Sprachfähigkeiten für die Untersuchung der Sprach- und Bildungsentwicklung älterer Lernender, Heranwachsender und Erwachsener selten in Betracht gezogen. Dies

gilt in besonderem Maße im Kontext von Mehrsprachigkeit, wenn also mindestens zwei der einer Person zur Verfügung stehenden Sprachen einzubeziehen sind.

Gegenstand des vorliegenden Arbeitspapiers ist zunächst eine zusammenfassende Darstellung des theoretischen Konstrukts, das den in MEZ und MEZ-2 verwendeten Schreibaufgaben zugrunde liegt, und die daraus abgeleiteten Auswertungskategorien. Dabei beziehen wir uns im Wesentlichen auf bereits publizierte Darstellungen (insbesondere auf Klinger et al., 2022; Gogolin et al., 2022). In einem weiteren Abschnitt werden die verschiedenen Versionen der Schreibaufgaben behandelt, die auf der Basis des bei Projektbeginn bereits vorliegenden Instruments „Fast Catch Bumerang“ für die Paneluntersuchung MEZ entwickelt und eingesetzt wurden. Dem Erhebungsdesign der Schreibaufgaben und einer Bestandsaufnahme der erhobenen Textdaten sind die darauffolgenden Kapitel gewidmet. Eine Darstellung des Auswertungs- und Kodierungsprozesses umfasst die allgemeinen Anweisungen und Prozeduren, die dabei zur Anwendung kamen. Abschließend widmen wir uns der Frage, wie die Ergebnisse der Schreibaufgaben mit anderen in MEZ erhobenen Sprachdaten zusammenhängen.

## 2 Kompetenzmodell

Das für die MEZ-Studie maßgebliche Konstrukt der Schreibkompetenz wurde ursprünglich im Rahmen eines Programms zur Förderung von Kindern und Jugendlichen mit Migrationshintergrund (FörMig, siehe Gogolin et al., 2011) entwickelt. Im Rahmen dieses Programms wurden Instrumente zur Messung des Sprachentwicklungsfortschritts benötigt, um die Angemessenheit der von den beteiligten Bildungseinrichtungen entwickelten und praktizierten Fördermaßnahmen zu evaluieren (Reich et al., 2009; Gantefort und Roth, 2010). Als theoretische Grundlage für die Konstruktion dieser Instrumente diente Ehlichs allgemeines Konzept der Erweiterung sprachlicher Basisqualifikationen im Verlauf der Sprachentwicklung (Ehlich, 2005; 2013a). Dieses Konzept basiert auf einer funktional-pragmatischen Perspektive auf Kommunikation und bietet ein komplexes Konstrukt der sprachlichen Gestaltung der kommunikativen Interaktion. In Bezug auf die Modalitäten umfasst es die Dimensionen Mündlichkeit und Schriftlichkeit sowie rezeptive und produktive Aktivitäten. Hinsichtlich der sprachlichen Strukturen werden die Dimensionen Morphosyntax, Semantik und Pragmatik einbezogen. Schließlich umfasst das Konzept auch gattungsspezifische Merkmale: diskursive und literarische Qualifikationen (Ehlich, 2013b). Auf der Grundlage dieses Ansatzes wird die Entwicklung der Kompetenzen der Lernenden durch die Komplexität ihrer (mündlichen oder schriftlichen) Äußerungen in den verschiedenen Dimensionen angezeigt.

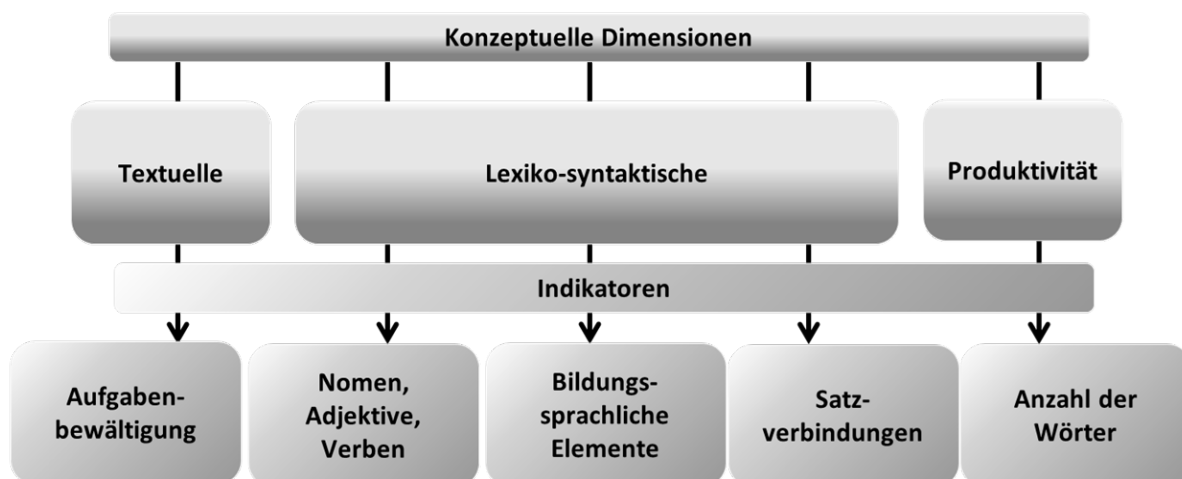
Das MEZ-Konstrukt mehrsprachiger Schreibkompetenz basiert auf diesem generischen Modell, das durch Ergebnisse der Schreibforschung ergänzt und weiter verfeinert wird (Puranik et al., 2008; Wagner et al., 2011). Diesem Ansatz zufolge sind morphosyntaktische, semantische, pragmatische und literale Qualifikationen für die Textproduktion fortgeschrittener Lerner(innen) in Bildungskontexten besonders wichtig (Becker-Mrotzek, 2014; Gantefort und Roth, 2010). Das theoretische Modell, das in MEZ verwendet wurde, konzentriert sich auf drei Kompetenzberei-



che: auf (1) textuell-pragmatische und (2) lexiko-syntaktische Qualifikationen sowie auf (3) Produktivität (Flüssigkeit) (siehe Abbildung 1). Die Kombination dieser Qualifikationen wird als einheitliches Konstrukt betrachtet (Mehta et al., 2005; Schnoor, 2019), das die aufgabenspezifische Gesamtqualität der mit MEZ-Instrumenten erhobenen Texte widerspiegelt (Klinger et al., 2019; Schnoor, 2019; Usanova, 2019). Im folgenden Abschnitt beschreiben wir das vollständige Set an empirischen Indikatoren für die Kompetenzdimensionen (textuell-pragmatisch, lexiko-syntaktisch sowie Produktivität), die das Modell abdeckt.

Im Hinblick auf das Längsschnitt-Messmodell der Schreibkompetenzen, das wir in diesem Beitrag vorstellen, haben unsere Analysen ergeben, dass hierfür nicht alle verfügbaren Indikatoren verwendet werden können: Einige von ihnen waren im Zeitverlauf nicht invariant. Dieses Ergebnis ist jedoch vernachlässigbar, da alle Kompetenzdimensionen des Modells durch die übrigen Indikatoren operationalisiert werden konnten. Aus Gründen der Transparenz sowie für die Nutzung der Variablen in Querschnittsuntersuchungen geben wir gleichwohl einen Einblick in die Gesamtheit der verfügbaren Indikatoren.

**Abbildung 1:** Das MEZ-Konstrukt der Schreibkompetenz: Qualifikationen und Indikatoren für die Textanalyse (nach: Gogolin et al., 2022)



## 2.1 Beschreibung der Indikatoren und ihre Operationalisierung

Für die Messung müssen die Qualifikationen, die das generische Konstrukt der Schreibkompetenz ausmachen, auf konkrete sprachliche und empirisch erfassbare Phänomene übertragen werden, d.h. auf Indikatoren für Qualifikationen, welche in Texten, die im definierten Testkontext produziert werden, tatsächlich vorkommen. Dabei sind aufgabenspezifische und sprachspezifische Aspekte ebenso zu wählen wie Faktoren, die die Erfahrung (z.B. Alter) der Schreibenden berücksichtigen (Entwicklungsfaktoren). Darüber hinaus müssen die Indikatoren an das Setting angepasst werden, das durch die Aufgabe, die den Befragten gestellt wird, evoziert wird.

Um den Entwicklungscharakter von Schreibfähigkeiten zu berücksichtigen, wurde die Auswahl der Indikatoren für unsere Analysen an die besonderen Entwicklungsmerkmale des Schreibens

im Jugendalter angepasst, wobei wir uns insbesondere auf frühere Untersuchungen zum Schreiben in ein- und zweisprachigen Kontexten beziehen. Studien zur Schreibkompetenz einsprachiger Schüler haben mehrere Entwicklungsmuster herausgearbeitet, die für erfahrenere oder jugendliche Lerner charakteristisch sind. Nach Nippold et al. (2005) gibt der angemessene Gebrauch von Syntax, Semantik und Pragmatik beim Schreiben appellativer Texte Aufschluss über die Kompetenz in der späteren Sprachentwicklung. Im Hinblick auf bestimmte Sprachphänomene beim Schreiben von Jugendlichen haben Studien die Verwendung von Konjunktionen (Crowhurst, 1987; Nippold et al., 2005) und die zunehmende Verwendung von Inhaltswörtern (Substantive, Adjektive, Verben) hervorgehoben (Johansson, 2008; Berman und Ravid, 2009; Ravid und Levie, 2010). Im schulischen Kontext haben sich Elemente der Bildungssprache - wie die Verwendung von Nominalisierungen oder Passivkonstruktionen - als relevante Entwicklungsmerkmale in der Textproduktion von Jugendlichen erwiesen (vgl. Galloway et al., 2020). Darüber hinaus wurde gezeigt, dass die Anzahl der Wörter in einem Text ein starker allgemeiner Prädiktor für die Textqualität ist (Ollinghouse und Leaird, 2009). In weiteren Studien wurden auch für zweisprachige Texte spezifische Merkmale hervorgehoben (Überblick in Usanova, 2019): die Verwendung von Konjunktionen (Verhoeven, 2006; Risse, 2014; Gantefort, 2013) und Inhaltswörtern (Lanauze und Snow, 1989; Usanova, 2019).

Folgende Indikatoren wurden für die MEZ-Aufgaben verwendet:

1. *Aufgabenbewältigung.* Bezogen auf die Dimension der textuell-pragmatischen Qualifikation ist der Grad der Vollständigkeit und Ausarbeitung der Handlung in einem elizitierten Text (Reich, 2009) ein valider Indikator für die pragmatische Qualität von MEZ-Aufgaben. Dieser Indikator ist unabhängig von der Sprache, in der ein Text präsentiert wird. Alle anderen Indikatoren, die wir für die Textauswertung in der MEZ-Schreibaufgabe verwendet haben, sind auf der konzeptionellen Ebene sprachübergreifend vergleichbar, nicht aber auf der Ebene der konkreten sprachlichen Äußerungen.
2. *Nomen, Adjektive und Verben.* Bezogen auf die Dimension der lexiko-syntaktischen Qualifikation spiegelt der differenzierte Einsatz lexikalischer Mittel die Textkomplexität wider und gilt als Indikator für Textqualität (vgl. Stavans et al., 2019). Die Verwendung von lexikalischen Mitteln wird von der Komplexität einer Aufgabe, den Spezifika des Genres (Ollinghouse und Wilson, 2013) und dem Thema (Harmon et al., 2005) beeinflusst. Direkte lexikalische Maße wie die Anzahl der produzierten Inhaltswörter müssen bei sprachübergreifenden Vergleichen mit Vorsicht behandelt werden (Berman und Verhoeven, 2002).
3. *Konjunktionen.* Als Indikator für lexiko-syntaktische Qualifikation verweisen Konjunktionen auf Konnektivität oder Kohäsion (Berman und Verhoeven, 2002). Die Anwendung von Parataxen und Hypotaxen in Texten spiegelt den Grad der Aneignung komplexer sprachlicher Mittel und deren Umsetzung in textuelle Strukturen wider. Ihre Verwendung ist sprachspezifisch (Berman und Verhoeven, 2002).
4. *Elemente der Bildungssprache.* Dieser Indikator bezieht sich auf lexikalische und morpho-syntaktische Elemente (Bailey und Butler, 2003; Schleppegrell, 2004), die in bildungsbezogenen

Kontexten mit konzeptioneller Schriftlichkeit auftreten. Diese lexikalischen und morphosyntaktischen Phänomene sind sprach- (Berman und Verhoeven, 2002) und aufgabenspezifisch.

5. *Anzahl der Wörter.* Die Länge eines schriftlichen Tests ist ein Indikator für Flüssigkeit oder Produktivität, und sie ist darüber hinaus ein Indikator für Textqualität (z.B. Grabowski et al., 2014). Auch dieser Indikator ist sprachspezifisch und aufgabenabhängig (Berman und Verhoeven, 2002).

Die folgende Tabelle 1 zeigt die Operationalisierungen der Schreibkompetenzen, wie sie in MEZ und MEZ-2 verwendet wurden.

**Tabelle 1:** Operationalisierungen der im MEZ-Projekt verwendeten Indikatoren für Schreibkompetenzen (nach: Gogolin et al. 2022)

Indikator	Operationalisierung
Aufgabenbewältigung	Die Punktevergabe umfasst neun Unterkategorien, die jeweils ein Bild als Teil des Handlungsablaufs repräsentieren. Jede Darstellung der neun Bildszenen wird nach einer vierstufigen Skala danach bewertet, ob eine Szene im Text nicht erwähnt wurde (0 Punkte) oder eine <i>angedeutete</i> (1 Punkt), <i>einfache</i> (2 Punkte) bzw. <i>ausführliche Beschreibung einer Sequenz</i> (3 Punkte) vorliegt. Es können maximal 27 Punkte erreicht werden.
Nomen, Adjektive und Verben	Gesamtzahl der <i>types</i> der verwendeten Nomen, Adjektive und Verben.
Satzverbindungen	Gesamtzahl der <i>types</i> von koordinierenden und subordinierenden Konjunktionen.
Bildungssprachliche Elemente	Gesamtzahl der kontextgemäßen lexikalischen und morpho-syntaktischen Elemente der Bildungssprache ( <i>tokens</i> ).
Anzahl der Wörter (Produktivität)	Gesamtzahl der Einzelwörter.

Im vorgegebenen Testkontext muss eine Messung produktiver Schreibqualifikationen sensitiv für den Prozesscharakter des Schreibens sein (Becker-Mrotzek, 2014; Hayes, 2012). Alle Indikatoren und ihre Operationalisierungen zielen darauf ab, die sprachliche Kompetenz oder Fähigkeit einer Person anhand von Texten mit Hilfe eines Bewertungsverfahrens zu beurteilen. Während das Potenzial einer Person durch ihre Qualifikationen, Kompetenzen oder Fertigkeiten bestimmt wird, hängt die Realisierung dieser Dispositionen in einer Testsituation und ihre Testleistung auch von anderen Faktoren ab, wie z. B. dem Konzentrationsniveau und der Tagesform, der Motivation oder unkontrollierten äußeren Faktoren. Um den Qualitätsstandards der Test- und Messtheorie gerecht zu werden, können solche Faktoren in der Regel bis zu einem gewissen Grad kontrolliert werden (Moosbrugger und Kelava, 2012). Insbesondere bei Panelstudien muss man sich zusätzlicher Probleme bewusst sein, z.B., dass der Unterschied zwischen Leistung und Kom-

petenz im Zeitverlauf systematisch verzerrt sein kann. Dies kann der Fall sein, wenn Parallelversionen von Prompts in ihrer Schwierigkeit nicht äquivalent sind oder Effekte der wiederholten Teilnahme (Test-Retest-, Lern- und Trainingseffekte) zu zunehmenden Diskrepanzen in der Testleistung über die Zeit führen (Schrauf, 2009; Schnell et al., 2005; Bortz und Döring, 2006, S. 559).

In der MEZ-Studie war auch die spezifische Konstellation der Mehrsprachigkeit bei unterschiedlichen Schriftsystemen zu berücksichtigen. Für das Türkische als Sprache mit lateinischer Schrift gibt es nur geringe Unterschiede zur Schrift des Deutschen oder Englischen. Im Fall des Russischen hingegen ist der Unterschied der kyrillischen zur lateinischen Schrift zu beachten. Im Allgemeinen kann dieser Unterschied zwischen den Schriften als zusätzliches Hindernis für die Entwicklung der Zweisprachigkeit betrachtet werden, während die Ähnlichkeit der Schriften als vorteilhaft angesehen wird (Bialystok et al., 2005). Die Frage, wie mehrsprachige Lernende mit der Herausforderung unterschiedlicher Schriften umgehen, ist jedoch noch nicht angemessen beantwortet worden. Dies gilt insbesondere für Kontexte, in denen die Schülerinnen und Schüler nicht systematisch im Unterricht mit den verschiedenen Systemen vertraut gemacht werden (Böhmer, 2015; Brehmer und Usanova, 2017; Usanova, 2019). Davon war für die MEZ-Studie auszugehen, da der Zugang, den Kinder und Jugendliche aus Migrantenfamilien zu Unterricht in den familiären Herkunftssprachen haben, nach Auskunft der Bundesländer im Bundesdurchschnitt bei nicht einmal 10 Prozent liegt (vgl. Mediendienst Integration, 2022).

Im Rahmen der MEZ-Studie wurde diesem Problem dadurch begegnet, dass den Probanden die Wahl der Schrift im Russisch-Schreibtest ermöglicht wurde. Für die in der MEZ-Studie gewählte Analysestrategie spielt es im Allgemeinen keine Rolle, welches Schriftsystem bei der Textproduktion verwendet wird, denn das Erkenntnisinteresse gilt der Möglichkeit, schriftsprachbezogene Konzepte aufzudecken, die in unterschiedlichen Kontexten des Spracherwerbs erworben werden - von der ausschließlich mündlichen Entwicklung in einer der Sprachen bis hin zur Alphabetisierung in einer oder mehreren anderen Sprachen im Rahmen der formalen Bildung. Deshalb werden russischsprachige Texte identisch analysiert, unabhängig davon, ob sie in lateinischer oder kyrillischer Schrift oder in einer Kombination aus beiden Systemen geschrieben sind. Um Fragen zu Konsequenzen der unterschiedlichen Aneignungskontexte nachzugehen, wird das gewählte Schriftsystem bei den Bewertungskriterien in den russischen Sprachproben vermerkt. Somit berücksichtigt das Verfahren das Phänomen der Mehrschriftigkeit sowohl im engeren Sinne als „Praktizieren verschiedener Schrift(typ)en und Schriftsysteme“, die mit oder ohne mehrsprachige Alphabetisierung erfolgen kann, als auch die Mehrschriftigkeit im weiteren Sinne als „mehrsprachige Textkompetenz, also die Fähigkeit zu kommunikativ bewusster, systematischer und von der situativen Einbettung weitgehend unabhängiger Verwendung von Schriftsprache“ (Böhm und Roll, 2022, S. 49-50).

### **3 MEZ-Schreibaufgaben für Jugendliche**

Das Untersuchungsdesign folgte dem Prinzip, die produktive literale Kompetenz in Deutsch und in der Herkunftssprache der Familie (Russisch oder Türkisch) mit funktional gleichwertigen

Schreibaufgaben zu messen. Für die in der Schule erlernten Fremdsprachen (Englisch, Französisch und Russisch) wurden Instrumente eingesetzt, die berücksichtigen, dass sie (in der Regel) vor allem im Kontext von Unterricht gelernt und gebraucht werden. Daher waren hier Instrumente einzusetzen, die sich hinsichtlich der dargestellten Themen in Inhalten und Komplexität am allgemeinen Lehrplan des Fremdsprachenunterrichts orientieren (vgl. Klinger et al. 2022).

Die Schreibkompetenzen in den Herkunftssprachen der Familien (Deutsch, Türkisch, Russisch) wurden mit Parallelformen des Instruments *Fast Catch Bumerang* erhoben (Reich et al. 2009; Dirim & Döll 2009; zu Details siehe auch Gogolin et al. 2022 und Schnoor 2022). Jede Parallelversion verwendet eine Serie von neun Fotos als visuelle Stimuli und eine darauf bezogene Aufgabenstellung. Die Parallelversionen des bereits bestehenden Instruments waren zu Beginn des Projekts noch nicht verfügbar und mussten im Laufe der MEZ-Studie sukzessive entwickelt werden. In jeder Welle wurde die gleiche Parallelform für die Beurteilungen in Deutsch und in den Herkunftssprachen Russisch und Türkisch verwendet.

Auch die lehrplanorientierten Parallelversionen für die Beurteilung der Schreibkompetenz in den Fremdsprachen (Englisch und zweiten Fremdsprachen Französisch bzw. Russisch) waren in jeder Welle identisch. Es wird davon ausgegangen, dass die parallele Anwendung in zwei Sprachen ein geringes Risiko für Lern- oder Trainingseffekte birgt, da diese Art von Instrumenten zur Beurteilung produktiver Sprachfähigkeiten in einem offenen Format eingesetzt wird und jeweils die von den Schreibenden aktuell verwendeten sprachlichen Mittel analysiert werden (Gogolin et al. 2011). Dies wird zudem durch die zeitliche Trennung gesichert: Die Testung im Deutschen, in der Fremdsprache Englisch, in den Herkunftssprachen und in der zweiten Fremdsprache wurden mit einem Abstand von ca. 1 Woche durchgeführt (vgl. Klinger et al. 2022).

Die Schreibaufgaben der ersten Erhebungswelle tragen folgende Bezeichnungen:

- Lebkuchenhaus (LH) für Deutsch und die Herkunftssprachen Russisch (Пряничный домик) bzw. Türkisch (Kurabiye Evi);
- Breakfast in Germany (BiG) für Englisch und die in der Schule erlernte zweite Fremdsprache Französisch (Le petit déjeuner en Allemagne) bzw. Russisch (Завтрак в Германии).

In der zweiten Welle kamen zum Einsatz:

- Lichterkette (LK) für Deutsch und die Herkunftssprachen Russisch (Гирлянда) bzw. Türkisch (Süs Lambaları);
- Picnic at the Park (PaP) für Englisch und die in der Schule erlernte zweite Fremdsprache Französisch (Un pique-nique au parc) bzw. Russisch (Пикник в парке).

In der dritten Welle wurden verwendet:

- Fast Catch Bumerang (BUM) für Deutsch und die Herkunftssprachen Russisch (Fast Catch бумеранг) bzw. Türkisch (Fast Catch Bumerang);

- A Trip to Hamburg (TtH) für Englisch und die in der Schule erlernte zweite Fremdsprache Französisch (Une excursion à Hambourg) bzw. Russisch (Поездка в Гамбург).

In der vierten Welle wurden alle Versionen für Deutsch und die Herkunftssprachen sowie alle Versionen für die Fremdsprachen in einem parallel rotierenden System eingesetzt, um ihre Vergleichbarkeit zu messen und aufgabenspezifische Motivationseffekte zu erfassen (Schnoor 2022).

**Tabelle 2:** Verteilung der in MEZ und MEZ-2 verwendeten Schreibaufgaben auf die sechs Messzeitpunkte

Sprachen	MEZ				MEZ-2	
	MZP 1	MZP 2	MZP 3	MZP 4	MZP 5	MZP 6
Deutsch und Herkunftssprachen (Russisch/Türkisch)	LH	LK	BUM	LK	BUM	BUM
Fremdsprachen (MEZ: Englisch, Französisch, Russisch; MEZ-2: Englisch)	BiG	PaP	TtH	PaP	TtH	TtH


Lebkuchenhaus (LH), Lichterkette (LK), Bumerang (BUM), Breakfast in Germany (BiG), Picnic at the Park (PaP), A Trip to Hamburg (TtH) bzw. die jeweiligen herkunfts- und fremdsprachlichen Entsprechungen

### 3.1 Die Testversionen der Schreibaufgabe

**Fast Catch Bumerang/Fast Catch бумеранг/Fast Catch Bumerang:** Als Ausgangsmodell für alle Schreibaufgaben diente das Instrument *Fast Catch Bumerang*, das im Rahmen des Modellprojekts FörMig entwickelt wurde. Es erfasst sprachproduktive Schreibfähigkeiten von Jugendlichen am Ende der Sekundarstufe I (Dirim & Döll 2009; Döll 2012). Als Arbeitsprobe im Rahmen einer Bewerbung soll ein Artikel für eine Jugendzeitschrift verfasst werden, in dem dargestellt wird, wie ein bestimmtes Bumerang-Modell gebaut wird. Vorlage dafür ist eine Folge von neun Fotografien, die die einzelnen Schritte seiner Produktion zeigen (vgl. Reich et al. 2009, 231ff.).

Nach diesem Muster wurden während der Projektlaufzeit sukzessive Parallelversionen entwickelt, die jeweils ebenfalls mit neun Fotografien vorgestellt werden, in denen eine konsistente Abfolge von Aktivitäten abgebildet ist. Um eine weitgehende Vergleichbarkeit der Versionen zu gewährleisten, ist für alle Aufgaben vorgegeben, dass sie als Arbeitsproben im Rahmen einer fiktiven Bewerbung dienen sollen. Der Wortlaut der Anweisungen ist (mit Ausnahme des Themas) identisch:

Schreibe nun bitte den Artikel. Beschreibe dazu die Abbildungen, die erklären, wie *#impulsspezifische Formulierung* - z.B.: „ein Lebkuchenhaus gebaut wird“#.

- Dein Artikel muss ohne Abbildungen verständlich sein.
- Bitte schreibe mindestens bis zu diesem Zeichen<sup>1</sup>: 
- Beschreibe alle Arbeitsschritte so gut du kannst.

**Lebkuchenhaus/ Пряничный домик/ Kurabiye Evi:** In diesem Impuls umfasst die Arbeitsprobe einen Artikel über den Bau eines Lebkuchenhauses, der in neun Arbeitsschritten dargestellt ist.

**Lichterkette/ Гирлянда/ Süs Lambaları:** Hier werden neun Fotografien vorgegeben, die den Zusammenbau einer Lichterkette mithilfe von elektrischen Lampen, Wollfäden, Leim und Luftballons darstellen.

**Breakfast in Germany/ Le petit déjeuner en Allemagne/ Завтрак в Германии:** Hier besteht die Arbeitsprobe aus einem Artikel für einen Reiseführer für englisch-/französisch-/russischsprachigen Touristen, in dem in neun Schritten erklärt wird, wie ein „typisch deutsches“ Frühstück zubereitet wird.

**Picnic at the Park/ Un pique-nique au parc/ Пикник в парке:** Der Artikel soll für ein englisch-/französisch-/russischsprachiges Jugendmagazin geschrieben werden und die Vorbereitungen für ein Picknick im Park mit Freunden erklären.

**A Trip to Hamburg/ Une excursion à Hambourg/ Поездка в Гамбург:** Der Artikel soll in einem Reiseführer für englisch-/französisch-/russischsprachige Touristen einen Ausflug mit Freunden nach Hamburg beschreiben.

### 3.2 Erhebungsdetails im Rahmen von MEZ

Die Organisation der Feldarbeit für die MEZ-Hauptstudie, ihre Durchführung und ein Teil der Datenverarbeitung wurden von einem externen Institut, der *International Association for the Evaluation of Educational Achievement* (IEA Hamburg), durchgeführt (IEA Hamburg 2017a, 2017b, 2018a, 2018b). Die Testsitzungen wurden in den teilnehmenden Schulen von geschulten Testleiterinnen und Testleitern angeleitet. Die Tests und die schriftlichen Befragungen wurden an zwei verschiedenen Tagen während der regulären Unterrichtszeit durchgeführt (vgl. hierzu und zum folgenden Klinger et al. 2022). Die Aufteilung auf zwei Tage diente dazu, Transfereffekte zwischen den Tests in mehreren Sprachen zu verringern. Der Abstand zwischen dem ersten und dem zweiten Testtag betrug eine Woche. Am ersten Testtag betrafen die Tests alle Schüler(innen) der Stichprobe. Der zweite Testtag betraf die Testabschnitte für die Herkunftssprachen und die zweite Fremdsprache, daran nahmen nur Schülerinnen und Schüler mit einem deutsch-russischen oder deutsch-türkischen Sprachhintergrund oder solche mit den Fächern Französisch oder Russisch als Fremdsprache teil (IEA Hamburg 2017a, 23). Zu den schriftlichen Tests der Herkunfts-

---

<sup>1</sup> Das „Daumen hoch“ Zeichen wurde an einer bestimmten Stelle auf der ersten für die Texte vorgesehenen linierten Seite in allen Testheften abgebildet, um nach Möglichkeit eine Mindesttextlänge von ca. 100 Wörtern zu erhalten.

sprachen Türkisch und Russisch wurden nur Teilnehmende zugelassen, die in Welle 1 Mindestanforderungen an die jeweiligen Sprachkenntnisse erfüllten. Dies wurde mit einem als Hörverständnistest konzipierten Locator-Test (NEPS) ermittelt.

Alle Sitzungen wurden in Testgruppen durchgeführt (IEA Hamburg 2017a, 23). Eine Testgruppe bestand aus maximal 30 Schülerinnen und Schülern. Da nicht mehr als 30 Schülerinnen und Schüler gleichzeitig getestet werden sollten, gab es an einigen Schulen parallele Testgruppen. In Schulen mit wenigen teilnehmenden Schüler(inne)n wurden beide Klassenstufen gemeinsam getestet.

Der zeitliche Ablauf der Testsitzungen richtete sich nach den vorgegebenen Zeitvorgaben für die Einführung und Bearbeitung der Instrumente und Fragebögen und sah feste Pausen vor. Die Testleiterinnen und Testleiter kontrollierten die korrekte Durchführung der Testsitzungen.

Für die hier behandelten Schreibaufgaben galten folgende Regeln: Alle Instruktionen waren in Testheften abgedruckt und wurden bei den in deutscher Sprache vorgegebenen Aufgaben für Deutsch und die Schulfremdsprachen zusätzlich von den Testleiterinnen und -leitern vorgelesen. Für die Schreibaufgaben in den Herkunftssprachen waren besondere Anweisungen vorgesehen, da davon auszugehen war, dass etliche Schüler(innen) keinen formalen Unterricht in diesen Sprachen erhalten hatten. Um zusätzliche Impulse für das Schreiben in der Herkunftssprache zu geben, wurden die Anweisungen für die Schreibaufgaben in gedruckter Form auf Russisch oder Türkisch vorgelegt und zusätzlich auf einer Audio-CD präsentiert, die mit „Muttersprachler(inne)n“ dieser Sprachen aufgenommen wurden, da die Testleiterinnen und Testleiter diese Sprachen in der Regel nicht sprachen. Bei allen Schreibaufgaben wurden die Teilnehmenden darüber informiert, dass die Orthographie für die Analysen keine Rolle spielt. Die Anweisungen für das Schreiben in Russisch erlaubten ausdrücklich die Verwendung der lateinischen Schrift zusammen mit oder anstelle der kyrillischen. Die Zeit für das Schreiben der Texte in Deutsch und in der Herkunftssprache betrug jeweils 30 Minuten. Wenn alle Teilnehmer einer Testgruppe die Aufgabe früher beendeten, konnte die Testzeit beendet werden. Für die Texte in Englisch und der zweiten Fremdsprache wurde in der ersten Welle die gleiche Testzeit angesetzt. Da sich herausstellte, dass die tatsächlich benötigte Zeit in diesen Sprachen durchweg niedriger war, wurde die Bearbeitungszeit ab Welle 2 auf 20 Minuten reduziert.

### **3.3 Abweichende Erhebungsmodi bei MEZ-2**

Vor dem Ende der MEZ-Studie wurden die Teilnehmenden um ihr Einverständnis zur Adressspeicherung für eine Folgebefragung gebeten, das von 1133 Proband(inn)en erklärt wurde (IEA Hamburg 2020). Da die Teilnehmenden zu einem großen Teil nicht mehr im früheren Schulkontext angetroffen werden konnten, erfolgte die Datenerhebung in MEZ-2 in zwei Erhebungswellen in individualisierter Form. Die Datenerhebung wurde mit einem gemischten Erhebungsdesign (Multi-Mode-Survey) aus onlinebasierten und postalischen Erhebungsverfahren realisiert (vgl. hierzu und zum folgenden Gogolin et al. 2021). Die Datenerhebung erfolgte computergestützt im Rahmen eines Online-Testmoduls, in das sich die Teilnehmenden mit persönlichen Login-Da-



ten einwählen konnten und das sie anschließend durch die einzelnen Bestandteile der Testsitzung (Sprachtests und Fragebogen) führte. Die nach wie vor *paper-and-pencil*-basierten Schreibaufgaben wurden durch Einblenden der Aufgaben und Bildimpulse zeitlich in die Online-Testsitzungen integriert. Die Texte wurden von den Teilnehmenden handschriftlich in ein Testheft geschrieben, das per Post an das Erhebungsinstitut zurückgeschickt wurde.<sup>2</sup> Es wurden wie in MEZ pro Welle zwei Testsitzungen durchgeführt. Die deutschsprachigen Texte wurden in der ersten Testsitzung geschrieben, in der möglichst eine Woche später erfolgenden<sup>3</sup> zweiten Testsitzung wurden die Texte in Herkunftssprachen Russisch bzw. Türkisch sowie auf Englisch erhoben.<sup>4</sup>

In MEZ-2 wurde in beiden Erhebungswellen im Deutschen sowie in den Herkunftssprachen Russisch bzw. Türkisch ausschließlich das Instrument *Fast Catch Bumerang* und im Englischen *A Trip to Hamburg* eingesetzt. Die Erhebungszeit wurde in allen Versionen auf 20 Minuten begrenzt.

### 3.4 Erhobene Daten

Tabelle 3 liefert eine nach Erhebungswellen und Sprachversionen gegliederte Übersicht über das im Rahmen von MEZ und MEZ-2 gewonnene Gesamtkorpus, das alle Texte umfasst, die von den MEZ-Schüler(inne)n in den sechs Erhebungswellen geschrieben wurden.

**Tabelle 3:** Anzahl der in MEZ und MEZ-2 elizitierten Schreibproben (Texte)

	Deutsch	Herkunftssprache Russisch	Herkunftssprache Türkisch	Fremdsprache Englisch	Fremdsprache Französisch	Fremdsprache Russisch	Gesamt
<b>Welle 1*</b>	1.751	248	421	1.741	661	58	4.880
<b>Welle 2*</b>	1.716	235	469	1.680	652	60	4.812
<b>Welle 3*</b>	1.591	182	398	1.584	558	55	4.368
<b>Welle 4*</b>	1.083	123	243	1.048	269	33	2.799
<b>Welle 5**</b>	656 +49 digit.	108	121	635	-	-	1.569
<b>Welle 6**</b>	597 +58 digit.	97	111	562	-	-	1.425
<b>Gesamt</b>	7.501	993	1.763	7.250	2.140	206	19.853

\* Gruppenerhebung im Schulkontext (MEZ), \*\* online administrierte Individualerhebung (MEZ-2), teilweise digital (digit.) erhoben.

<sup>2</sup> In beiden Erhebungswellen von MEZ-2 wurden an einer Zufallsauswahl von je 60 Teilnehmenden die Schreibaufgabe auf Deutsch als computerbasierte Version administriert (IEA Hamburg 2020, 2021).

<sup>3</sup> „Dies ließ sich in dem individualisierten Erhebungsdesign der MEZ -2-Studie nicht immer realisieren, weil die Teilnehmerinnen und Teilnehmer selbst den Zeitpunkt bestimmten, zu dem sie die Erhebungssitzung 1 beendeten bzw. mit der Erhebungssitzung 2 begannen“ (IEA Hamburg 2020, S.7).

<sup>4</sup> Die Schulfremdsprachen Französisch und Russisch wurden in MEZ-2 nicht berücksichtigt.

## 4 Auswertung

Die Texte wurden von geschulten Kodierer(inne)n unter Supervision des Forschungsteams der Universität Hamburg ausgewertet. Im Projekt MEZ selbst standen für die Textanalysen durch Kodierer(innen) ausschließlich Bilddateien (Scans) der handschriftlichen Texte aus den Testheften zur Verfügung; Transkriptionen waren während der Projektlaufzeit nicht verfügbar. Die Datenauswertung erfolgte anhand von sprachversionenspezifischen Auswertungshinweisen, in denen die in Abschnitt 2 beschriebenen Indikatoren der Auswertungskategorien sowie die darauf bezogene Subsumierung von Texten und Textbestandteilen beschrieben und anhand von Beispielen aus Pretests und früheren Untersuchungen erläutert wurden.

Die Auswahl des Kodierungspersonals erfolgte nach Feststellung der erforderlichen Sprachkenntnisse in der jeweiligen Testsprache und der Beherrschung der für die Analysen relevanten linguistischen Grundkenntnisse. Die Kodierungsschulung erfolgte in zwei je vierstündigen Blöcken. Im ersten Block wurde der Kontext der Untersuchung vorgestellt sowie das Auswertungsverfahren und die Auswertungsindikatoren erläutert. In Kleingruppen wurden dann Beispieltex-te analysiert und die Ergebnisse im Plenum vorgestellt und diskutiert. Jede(r) Kodierer(in) erhielt weitere Texte zur Analyse bis zum zweiten Schulungsblock, in dem diese Analysen präsentiert sowie Problemfälle und Fragen diskutiert wurden. Im mehrjährigen Verlauf der Studie waren bei Ausscheiden von Kodierungspersonal Neurekrutierungen nötig, für die gleichartige Schulungen individuell oder in Kleingruppen durchgeführt wurden.

Regelmäßige Qualitätskontrollen für jede Kodiererin und für jeden Kodierer umfassten eine kontinuierliche Überprüfung der Eindimensionalität und internen Konsistenz der auf den Auswertungskategorien basierenden Gesamtskalen. Bei Abweichungen von der Norm wurde eine individuelle Nachschulung durchgeführt.

Die Inter-Rater-Reliabilität wurde für jede Parallelf orm und jede Sprachversion der Schreibaufgabe in der letzten Phase der Untersuchung überprüft. Dazu wurden Doppelkodierungen an Textstichproben aus allen Wellen vorgenommen und mit den vorhandenen Analysen verglichen. Durchweg wurde bei allen Impulsversionen und in jeder der untersuchten Sprachen eine hohe Inter-Rater-Reliabilität erzielt (Tabelle 4, S. 19).

**Tabelle 4:** Inter-Rater-Reliabilität nach Sprache und Impuls - Intraklassen-Korrelationskoeffizienten (ICCs)

Sprachversion	Impuls 1*	Impuls 2**	Impuls 3***
<b>Deutsch</b>	.869	.935	.944
<i>n</i>	50	45	50
<b>Russisch</b>	.988	.970	.966
<i>n</i>	35	33	39
<b>Türkisch</b>	.971	.968	.865
<i>n</i>	48	46	47
<b>Englisch</b>	.887	.983	.945
<i>n</i>	50	49	50
<b>Französisch</b>	.969	.937	.957
<i>n</i>	48	49	44

\* Deutsch, Russisch, Türkisch: "Lebkuchenhaus", Englisch, Französisch: "Breakfast in Germany"

\*\* Deutsch, Russisch, Türkisch: "Lichterkerle", Englisch, Französisch: "Picnic at the park"

\*\*\* Deutsch, Russisch, Türkisch: "Fast Catch Bumerang", Englisch, Französisch: "A trip to Hamburg"

## 5 Zusammenhänge mit anderen Sprachdaten

Die Sprachfähigkeit eines Individuums besteht aus einer Vielzahl komplexer sprachlicher Teilfähigkeiten. Da diese Teilfähigkeiten auf mehreren sozialen und kognitiven Faktoren beruhen, können - abgesehen von direkten Zusammenhängen - für jede Teilfähigkeit eigenständige Erklärungsmechanismen eine Rolle spielen (de Bot, Lowie & Verspoor 2007). So kann eine dissoziierte Entwicklung unterschiedlicher Teilfähigkeiten vorkommen, wobei Personen in bestimmten Sprachfähigkeitsbereichen im Vergleich zu Fähigkeiten in anderen Bereichen überproportional fortgeschritten sein können (Berninger, Abbott, Cook, & Nagy, 2017). Insbesondere das Schreiben ist eine komplexe Fähigkeit, die sich über längere Zeit unter unterschiedlichen Lernbedingungen entwickelt und individuell unterschiedlich ausgeprägt ist (Bazerman et al., 2017). Um Schreiben als komplexe, sich über die Zeit entwickelnde Fähigkeit besser verstehen und einordnen zu können, wird in diesem Kapitel der Zusammenhang zu anderen sprachlichen Teilfähigkeiten untersucht und auf offene Fragen und Forschungsdesiderate hingewiesen. Es ist zu erwarten, dass die Teilfähigkeiten positiv miteinander korrelieren, jedoch in unterschiedlichem Ausmaß (vgl. Klinger et al. 2019; Jude 2008).

In diesem Kapitel werden Daten vom ersten Messzeitpunkt (MZP 1) verwendet, weil die Unterschiede zwischen den Messzeitpunkten relativ gering sind und hier eine hinreichend große Teilnehmerzahl gewährleistet ist. In MEZ wurden rezepptive Fähigkeiten in Deutsch und Herkunftssprachen mithilfe des Lesegeschwindigkeits- und Verständnistests LGVT 5-12+ (Schneider et al. 2017; Schlagmüller et al. 2022) sowie allgemeine Sprachfähigkeiten in den untersuchten Fremdsprachen mittels teilweise selbstentwickelter C-Tests erhoben (vgl. Klinger et al. 2022). Außer-

dem liegen in der Sprache Deutsch sowie in den Fremdsprachen und Herkunftssprachen für jeden Messzeitpunkt Selbsteinschätzungen der eigenen Fähigkeiten in verschiedenen Sprachfähigkeitsbereichen vor, jeweils gemessen auf einer 6-Punkte Skala.

### 5.1 Lesegeschwindigkeits- und Verständnistest für die Klassen 5–12 (LGVT 5-12+)

Der standardisierte Test LGVT 5-12+ misst Leseverständnis und Lesegeschwindigkeit im Deutschen für die Klassen 5 bis 12. Beim Lesen eines Fließtextes sollen an 47 Textstellen unter jeweils drei Alternativen das in den Textzusammenhang passende Wort unterstrichen werden. Wir verwenden für diesen Beitrag den Rohwert für das Leseverständnis, der eine Ratekorrektur vorsieht (Schneider et al. 2017). Im MEZ-Projekt wurden russisch- bzw. türkischsprachige Adaptionen erarbeitet, für Russisch wurde wahlweise neben einer kyrillischen auch eine in lateinischer Schrift transliterierte Version angeboten (Schlagmüller et al. 2022). Die Reliabilität, gemessen an der internen Konsistenz (Cronbachs  $\alpha$ ), beträgt am ersten Messzeitpunkt von MEZ für den deutschsprachigen Test .82 und kann damit als gut bewertet werden; auch die russischsprachige Adaption erreicht für die kyrillische Version einen guten  $\alpha$ -Wert mit 0.86, aber eine sehr geringe Konsistenz für die lateinisch transliterierte Fassung ( $\alpha = 0.15$ ); die türkische Version weist mit einem  $\alpha$  von 0.64 eine mäßige Konsistenz auf (Schlagmüller et al. 2022; Klinger et al. 2019).

**Tabelle 5:** Korrelationen zwischen MEZ-Sprachdaten am 1. Messzeitpunkt

Schreiben	LGVT Deutsch	LGVT Türkisch	LGVT Russisch <sup>+</sup>		C-Test Englisch	C-Test Französisch
			KYR	TRLT		
<b>Deutsch</b>	.38**	.19**	.08	.14	.46**	.41**
<i>n</i>	1748	417	156	105	1747	667
<b>FS Englisch</b>	.45**	.31**	.19*	.11	.74**	.57**
<i>n</i>	1737	406	158	105	1739	663
<b>FS Französisch</b>	.40**	.26**	.24	.32	.75**	.76**
<i>n</i>	635	163	56	38	635	659
<b>HS Russisch</b>	.14*	--	.61**	.35**	.34**	.31**
<i>n</i>	242	--	148	94	242	89
<b>HS Türkisch</b>	.21**	.58**	--	--	.36**	.19*
<i>n</i>	402	421	--	--	402	169

Die Korrelationen sind auf dem Niveau von \*  $p < .05$  und \*\*  $p < .001$  (2-seitig) signifikant.

<sup>+</sup>Die Schreibaufgaben Russisch durften in kyrillischen (KYR) oder lateinisch transliterierten Buchstaben (TRLT) verfasst werden (vgl. oben, S.13).

Wie aus Tabelle 5 ersichtlich, hängen - den Erwartungen entsprechend - die Rohwerte für das Leseverständnis in allen Sprachen positiv und signifikant auf einem Signifikanzniveau von  $p < .001$  mit den Gesamtscores für die Schreibfähigkeiten zusammen. Dabei sind die Zusammenhänge für die Herkunftssprachen stärker (Pearsons  $r$  zwischen .58 und .61) als für das Deutsche ( $r = .38$ ). Der auffällige Unterschied zwischen den bei kyrillischem und lateinischem Alphabet erzielten Ergebnissen spiegelt die größere Konsistenz der Testergebnisse der kyrillischen wider. Dabei sind jedoch auch die geringeren Fallzahlen bei der lateinisch transliterierten Version zu

beachten. Über die Zeit, an Messzeitpunkt 3 und 6, bleiben die Zusammenhänge zwischen dem Schreiben und dem Leseverständnis für Deutsch gleich, während sie sich in den Herkunftssprachen etwas abschwächen. Jedoch sind die Korrelationen auch an Messzeitpunkt 6 noch etwas stärker in der deutschen Sprache als in den anderen Sprachen (siehe Tabelle 7 im Anhang).

## 5.2 C-Test

Zur Messung der allgemeinen Sprachkompetenz in den im Schulunterricht erworbenen Fremdsprachen Englisch, Französisch und Russisch entwickelte das MEZ-Team insgesamt acht C-Tests auf der Grundlage von Lehrplänen für den Fremdsprachenunterricht in der Sekundarstufe (Schnoor et al. 2023), zwei Texte wurden aus dem KESS-Projekt (Bos et al. 2009; Bos & Gröhlich 2010) übernommen. C-Tests messen die allgemeine Sprachkompetenz anhand der Korrektheit der Rekonstruktion von systematisch bei jedem zweiten bzw. dritten Wort in einem Text gelöschten Worthälften. In jeder Erhebungswelle von MEZ, so auch am hier berücksichtigten 1. Messzeitpunkt, bestand jeder C-Test aus vier Texten.

Die erzielten C-Test-Scores in den Fremdsprachen korrelieren im Vergleich zu den Rohwerten für das Leseverständnis stärker mit den Gesamtscores für Schreibfähigkeiten ( $r$  zwischen 0.74 und 0.83). Da C-Tests sowohl produktive als auch rezeptive Sprachfähigkeiten erfordern und sich somit von den Anforderungen her teilweise mit der Schreibaufgabe überschneiden, ist dieses Ergebnis erwartungsgemäß (vgl. Gogolin, Schnoor & Usanova, 2021).

## 5.3 Selbsteinschätzung

Ein weiteres Instrument, das in MEZ an allen sechs Messzeitpunkten eingesetzt wurde, erfasst die Selbsteinschätzung der Sprachfähigkeiten in den jeweils zutreffenden MEZ-Sprachen. Als Teil der Schüler(innen)befragung schätzten die Teilnehmer(innen) die Beherrschung von verschiedenen sprachlichen Teilfähigkeiten auf einer sechs-Punkte-Skala ein (1 = sehr gut bis 6 = ungenügend). Die pro Sprache erfragten sechs Teilfähigkeiten umfassen Aussprache; Texte schreiben; Rechtschreibung und Zeichensetzung; Grammatik beim Sprechen; Grammatik beim Schreiben; und Wortschatz bzw. Vokabeln. Am Messzeitpunkt 3 wurde zusätzlich die Beherrschung des Hörverstehens erfragt; an Messzeitpunkt 5 und 6 wurde Texte lesen und verstehen erfragt. Die Selbsteinschätzungen an den Messzeitpunkten 1, 3 und 6 zeichnen sich durch hohe Cronbach's  $\alpha$ -Werte von mindestens .85 aus. Hauptkomponentenanalysen zeigten außerdem eine überwiegend starke interne Validität, d.h. die gemessenen Fähigkeiten gehören zu einem latenten Konstrukt. Einzige Ausnahme bildet die Herkunftssprache Russisch an Messzeitpunkt 1, hier ergaben sich zwei latente Konstrukte aufgrund einer niedrigeren Ladung der Items „Texte schreiben“ und „Grammatik beim Schreiben“. Da sich keine unterschiedlichen Tendenzen für das Konstrukt Schreiben gegenüber dem Gesamtkonstrukt zeigten, werden hier Analysen mit der durchschnittlichen Gesamteinschätzung pro Sprache vorgestellt.

Die Gesamt-Selbsteinschätzungen in den verschiedenen Sprachen korrelieren positiv und signifikant auf einem Signifikanzniveau von  $p < .05$  oder  $p < .001$  miteinander (2-seitig, siehe Tabelle 8 im Anhang), mit einer Ausnahme: Trotz signifikanter Zusammenhänge zwischen *gemessenen* Sprachfähigkeiten in diesen Sprachen (Tabelle 5) besteht kein Zusammenhang zwischen den

Selbsteinschätzungen in der Herkunftssprache Russisch und der Fremdsprache Französisch ( $r = .07$ , n.s.).

Inwieweit Selbsteinschätzungen der eigenen Sprachfähigkeiten mit gemessenen Sprachleistungen zusammenhängen, ist eine eigenständige Forschungsfrage (Heath, 1995; Klinger 2022, S. 109), die hier nicht im Detail verfolgt werden kann. Wir gehen davon aus, dass zwischen den Einschätzungen und den Schreibkompetenzen ein positiver, aber unvollkommener Zusammenhang besteht, der unterschiedlichen Einflüssen ausgesetzt ist.

Aus Tabelle 6 ist ersichtlich, dass die Selbsteinschätzungen in allen Sprachen positiv und signifikant auf einem Signifikanzniveau von  $p < .001$  mit den Gesamtscores für das Schreiben in der jeweiligen Sprache korrelieren. Die Korrelation für das Deutsche ist erneut niedriger (.29) als die Korrelationen zwischen Schreiben und den Selbsteinschätzungen in den anderen Sprachen, insbesondere niedriger als in den Herkunftssprachen, wo die Korrelationen zwischen .50 und .59 liegen, und niedriger als in der Fremdsprache Englisch ( $r = .48$ ). Diese Rangordnung findet sich im Prinzip auch an den späteren Messzeitpunkten 3 und 6 wieder.<sup>5</sup>

**Tabelle 6:** Korrelationen (Pearsons  $r$ ) zwischen Schreibfähigkeiten und Selbsteinschätzungen am 1. Messzeitpunkt

Schreiben	Selbsteinschätzungen					
	Deutsch	FS Englisch	FS Französisch	FS Russisch	HS Russisch	HS Türkisch
<b>Deutsch</b>	.29**	.24**	.09*	.17	.12	.02
<i>n</i>	1708	1697	647	59	266	401
<b>FS Englisch</b>	.29**	.48**	.18**	.20	.14*	.04
<i>n</i>	1701	1688	648	59	267	390
<b>FS Französisch</b>	.26**	.28**	.37**	--	.07	.05
<i>n</i>	622	623	643	--	97	158
<b>FS Russisch</b>	.17	.20	--	.58**	.49**	--
<i>n</i>	53	51	--	53	36	2
<b>HS Russisch</b>	.17*	.15*	-.03	.64**	.59**	--
<i>n</i>	238	235	86	37	238	--
<b>HS Türkisch</b>	.16*	.24**	-.04	--	--	.50**
<i>n</i>	392	386	164	2	--	402

Die Korrelationen sind auf dem Niveau von \*  $p < .05$  und \*\*  $p < .001$  (2-seitig) signifikant.

Es ist anzunehmen, dass für die geringere Korrelation im Deutschen ein Deckeneffekt in der Selbsteinschätzung verantwortlich ist: Als dominante Alltagssprache im Bildungssystem wird Deutsch von allen MEZ-Teilnehmenden genutzt. Die vorhandene Varianz wird in Messungen der Sprachfähigkeit besser erfasst als vermittelt Selbsteinschätzungen. Ähnliche Ergebnisse zu dem Deckeneffekt in den Selbsteinschätzungen im Deutschen wurden für eine Erwachsenen-Stichprobe in einer Studie von Klinger (2022) berichtet. Für die größere Treffsicherheit in den Schulfremdsprachen bietet sich die Erklärung an, dass Schüler(innen) sich bei der Selbsteinschätzung

<sup>5</sup> An Messzeitpunkt 3 ist allerdings die Korrelation zwischen Schreiben und Selbsteinschätzung für die Fremdsprache Französisch am höchsten ( $r = .56$ ).

an den Schulnoten oder anderen Formen von Feedback zu ihren Leistungen im Unterricht orientieren. Dass sich das Schulfach Deutsch nicht in gleicher Weise auf die Selbsteinschätzung überträgt, könnte dem Tatbestand geschuldet sein, dass Deutsch als alltägliche Umgebungssprache und Medium fast allen Unterrichts nicht so stark mit dem Fach Deutsch assoziiert wird. Wenig erklärlich ist die relativ gute Übereinstimmung zwischen Selbsteinschätzung und Testwerten. Bei den Herkunftssprachen dürfte vor allem das familiäre und weitere soziale Umfeld als Orientierung dienen, weil diese Sprachen nur in relativ geringem Maß durch Unterricht im Regelsystem vermittelt werden (Mediendienst Integration, 2022). Möglich ist, dass Reaktionen auf alltägliche Sprachhandlungen eine realistische Einschätzung der eigenen Sprachfähigkeiten unterstützen.

Die Selbsteinschätzungen für Deutsch und für Englisch korrelieren mit den Gesamtscores für das Schreiben in allen Sprachen, außer der Fremdsprache Russisch. Dies entspricht der Erwartung, dass Sprachfähigkeiten positiv miteinander korrelieren (Schnoor und Usanova 2023) und dass sich dies auch in Selbsteinschätzungen abbilden kann (Klinger et al. 2022). Zugleich zeigen sich zwischen Selbsteinschätzungen und Testergebnissen in anderen Sprachen unterschiedliche, teilweise schwache Übereinstimmungen. Die Orientierungsmechanismen, die den Selbsteinschätzungen zugrunde liegen, und die Frage, unter welchen Umständen Selbsteinschätzungen und messbare Sprachfähigkeiten einander besser oder schlechter entsprechen, können in diesem Arbeitspapier nicht weiterverfolgt werden. Die MEZ-Daten bieten aber weitere Möglichkeiten, diesem Problem nachzugehen, etwa unter Einbeziehung von Information zu Bildungsaspirationen, Wohlbefinden oder Akkulturationsmerkmalen.

## 6 Fazit

Im Rahmen der MEZ-Studie wurde ein theoretisches Konstrukt für mehrsprachige Schreibkompetenz entwickelt, auf dessen Grundlage ein Messmodell zur Erfassung der mehrsprachigen Schreibfähigkeiten von Jugendlichen erstellt und in fünf verschiedenen Sprachen über vier Untersuchungswellen hinweg angewendet wurde (vgl. Gogolin et al., 2022). Die Analyse schriftlicher Texte mithilfe dieses im Rahmen der MEZ-Studie entwickelten Messmodells für mehrsprachige Schreibkompetenz zeigt, dass dieses Modell sowohl für querschnittliche als auch für längsschnittliche Analysen geeignet ist und über mehrere Sprachen hinweg Einblicke in die Schreibfähigkeiten in den untersuchten Sprachen bietet.

Ein Ziel dieses Arbeitspapiers bestand darin, den Zusammenhang zwischen den mehrsprachigen Schreibfähigkeiten, die mit dem MEZ-Konstrukt gemessen wurden, und anderen sprachlichen Teilfähigkeiten zu untersuchen, die mithilfe unterschiedlicher Tests erfasst wurden. Unsere Analysen zeigen hin, dass die sprachlichen Teilfähigkeiten positiv miteinander korrelieren, jedoch in unterschiedlichem Ausmaß. Die positiven und signifikanten Korrelationen zwischen den Schreibfähigkeiten und den anderen Sprachdaten in verschiedenen Sprachen unterstreichen die Validität unserer Messungen. Gleichzeitig deuten die relativ niedrigen Korrelationsstärken darauf hin, dass die verwendeten Instrumente eigenständige Konstrukte messen, die spezifische Fähigkeiten reflektieren und auf individuelle Weise mit den Schreibfähigkeiten interagieren.

Das in MEZ und weiteren Vorgängerstudien entwickelte Kompetenzmodell eröffnet die Möglichkeit für komplexe Analysen mehrsprachiger Schreibfähigkeiten, die Wechselwirkungen zwischen verschiedenen Sprachen und Teilfähigkeiten aufdecken sowie Einflussfaktoren untersuchen und somit Erklärungen für die beobachteten Phänomene liefern können. Die nachfolgende Liste bisheriger Untersuchungen mit dem MEZ-Kompetenzmodell gibt einen Eindruck von der Vielfalt der Fragestellungen, für deren Beantwortung sich der Ansatz bereits bewährt und als fruchtbar erwiesen hat.

Untersuchungsthema	Publikation
Kompetenzmodell mehrsprachiger Schreibfähigkeiten	<p>Gogolin, I., Klinger, T., Schnoor, B. &amp; Usanova, I. (2022). The Competence Model of Writing Skills in the project "Multilingual Development: A Longitudinal Perspective (MEZ). In H. Brandt, M. Krause, &amp; I. Usanova (Hrsg.), <i>Language Development in Diverse Settings. Interdisziplinäre Ergebnisse aus dem Projekt "Mehrsprachigkeitsentwicklung im Zeitverlauf"</i> (MEZ) (S. 35–72). Springer VS.</p> <p>Klinger, T., Usanova, I., Akgün, G., Minh, T.T. (2022). Die Messung produktiver Sprachfähigkeiten im Kontext von Mehrsprachigkeit: Die Herangehensweise in LiPS. In: Klinger, T., Gogolin, I., Schnoor, B. (Hrsg.) <i>Sprachentwicklung im Kontext von Mehrsprachigkeit</i>. Springer VS, Wiesbaden. <a href="https://doi.org/10.1007/978-3-658-36770-1_2">https://doi.org/10.1007/978-3-658-36770-1_2</a></p>
Vorteile der Entwicklung mehrsprachiger Schreibfähigkeiten für die Schreibfähigkeit in Fremdsprache Englisch	<p>Klinger, T., Schnoor, B. &amp; Usanova, I. (2022). Developing Multilingual Writing Skills: Bilingual Advantage in Foreign Language Learning? In H. Brandt, M. Krause, &amp; I. Usanova (Hrsg.), <i>Language Development in Diverse Settings. Interdisziplinäre Ergebnisse aus dem Projekt "Mehrsprachigkeitsentwicklung im Zeitverlauf"</i> (MEZ) (S. 235–261). Springer VS.</p>
Zusammenhänge zwischen Lese- und Schreibfähigkeiten im Deutschen	<p>Klinger, T., Usanova, I., &amp; Gogolin, I. (2019). Entwicklung rezeptiver und produktiver schriftsprachlicher Fähigkeiten im Deutschen. <i>Zeitschrift für Erziehungswissenschaft</i>, 22(1), 75–103. <a href="https://doi.org/10.1007/s11618-018-00862-0">https://doi.org/10.1007/s11618-018-00862-0</a></p>
Messinvarianz der „MEZ-Schreibaufgabe Jugendliche“	<p>Klinger, T., &amp; Schnoor, B. (2020). <i>Prüfung der Messinvarianz über die Zeit: Das Konstrukt Schreibfähigkeit im Deutschen (MEZ Arbeitspapier Nr. 6)</i>. Hamburg: Universität Hamburg. <a href="https://www.mez.uni-hamburg.de/bilder/pdf/mezarbeitspapier06.pdf">https://www.mez.uni-hamburg.de/bilder/pdf/mezarbeitspapier06.pdf</a></p> <p>Klinger, T., Schnoor, B. (2022). Die vergleichende Messung produktiver sprachlicher Fähigkeiten in verschiedenen Alterskohorten. In: Klinger, T., Gogolin, I., Schnoor, B. (Hrsg.) <i>Sprachentwicklung im Kontext von Mehrsprachigkeit</i>. Springer VS, Wiesbaden. <a href="https://doi.org/10.1007/978-3-658-36770-1_3">https://doi.org/10.1007/978-3-658-36770-1_3</a></p>
Zusammenhänge zwischen mehrsprachigen Schreibfähigkeiten über die Zeit	<p>Schnoor, B. &amp; Usanova, I. (2022). Multilingual writing development: Relationships between writing proficiencies in German, heritage language and English. <i>Reading and Writing</i>, 36(3), 599–623. <a href="https://doi.org/10.1007/s11145-022-10276-4">https://doi.org/10.1007/s11145-022-10276-4</a></p>



Untersuchungsthema	Publikation
Vergleichbarkeit der Impulsversionen der „MEZ-Schreibaufgabe Jugendliche	Schnoor, B. (2022). Vergleichbarkeit der Impulsversionen der „MEZ-Schreibaufgabe Jugendliche“. In H. Brandt, M. Krause & I. Usanova (Hrsg.), <i>Language development in diverse settings: Interdisziplinäre Ergebnisse aus dem Projekt "Mehrsprachigkeitsentwicklung im Zeitverlauf" (MEZ)</i> (S.73–98). Wiesbaden: Springer VS.
Zusammenhang zwischen mehrsprachigen Schreibfähigkeiten und digitalen literalen Praktiken	Usanova, I., Schnoor, B. & Gogolin, I. (2023). Mehrsprachigkeit, digitale Praxis und Schreibfähigkeit. <i>Unterrichtswissenschaft</i> . <a href="https://doi.org/10.1007/s42010-023-00165-2">https://doi.org/10.1007/s42010-023-00165-2</a>
Sprachübergreifende mehrsprachige Schreibkompetenz als Konstrukt	Usanova, I., & Schnoor, B. (2022). Approaching the Concept of Multiliteracies: Multilingual Writing Competence as an Integrated Model. <i>Canadian Journal of Applied Linguistics</i> , 23(3), 144–165. <a href="https://doi.org/10.37213/cjal.2022.32598">https://doi.org/10.37213/cjal.2022.32598</a>
Profile mehrsprachiger Schreibfähigkeiten	Usanova, I. & Schnoor, B. (2021). Exploring multiliteracy in multilingual students – multilingual profiles of writing skills. <i>Bilingual Research Journal</i> , 44(1), 56–73. <a href="https://doi.org/10.1080/15235882.2021.1890649">https://doi.org/10.1080/15235882.2021.1890649</a>

## 7 Literaturverzeichnis

- Bailey, A. L., & Butler, F. A. (2003). *An evidentiary framework for operationalizing academic language for broad application to K–12 education: A design document (CSSE Report No. CSE-R-611)*. Los Angeles: Center for the Study of Evaluation, University of California.
- Bazerman, C., Graham, S., Applebee, A. N., Matsuda, P. K., Berninger, V. W., Murphy, S., ... & Schleppegrell, M. (2017). Taking the long view on writing development. *Research in the Teaching of English*, 351–360.
- Becker-Mrotzek, M. (2014). Schreibkompetenz. In J. Grabowski (Hrsg.), *Sinn und Unsinn von Kompetenzen. Fähigkeitskonzepte im Bereich von Sprache, Medien und Kultur* (S. 51–71). Opladen: Budrich.
- Berninger, V., Abbott, R., Cook, C. R., & Nagy, W. (2017). Relationships of attention and executive functions to oral language, reading, and writing skills and systems in middle childhood and early adolescence. *Journal of learning disabilities*, 50(4), 434–449.
- Berman, R. A., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language and Literacy*, 5(1), 1–43.
- Berman, R. A., & Ravid, D. (2009). Becoming a literate language user: Oral and written text construction across adolescence. In D. R. Olson, & N. Torrance (Hrsg.), *Cambridge handbook of literacy* (S. 92–111). Cambridge: Cambridge University Press.
- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy and learning to read. Interactions among languages and writing systems. *Scientific Studies of Reading*, 9(1), 43–61.
- Böhmer, J. (2015). *Biliteralität. Eine Studie zu literalen Strukturen in Sprachproben von Jugendlichen im Deutschen und Russischen*. Münster: Waxmann.
- Bortz, J., Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4th ed.). Berlin: Springer.
- Bos, W., Bonsen, M. & Gröhlich, C. (Hrsg.) (2009). KESS 7: Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7. *HANSE - Hamburger Schriften zur Qualität im Bildungswesen, Band 5*. Münster: Waxmann.
- Bos, W. & Gröhlich, C. (Hrsg.) (2010). KESS 8: Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8. *HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Band 6*. Münster: Waxmann.
- Böhm, M., & Roll, H. (2022). Nicht nur mehrsprachig, sondern auch mehrschriftig! Argumente für eine mehrsprachige Literalität in der Migrationsgesellschaft. *OBST 100 (Osnabrücker Beiträge zur Sprachtheorie)*, 43–64. Online unter: <https://journals.uni-marburg.de/obst/article/view/8467>
- Brandt, H., Lagemann, M., & Rahbari, S. (2017). Multilingual development – a longitudinal perspective: Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ). *European Journal of Applied Linguistics* 5(2). 347–357.
- Brehmer, B., & Usanova, I. (2017). Biscrptality and heritage language maintenance: Russian in Germany. In H. Peukert, & I. Gogolin (Hrsg.), *Dynamics of linguistic diversity (Hamburg*

- Studies on Linguistic Diversity* 6, (S. 99–121). Amsterdam: John Benjamins.  
<https://doi.org/10.1075/hslid.6.06bre>.
- Colón-Quintana, N., Polo, A. J., & Smith Carter, J. (2020). The reciprocal effects of language proficiency and depression among low income Latinx youth. *Journal of Clinical Child & Adolescent Psychology*, 1–14. <https://doi.org/10.1080/15374416.2020.1731818>
- Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the Teaching of English*, 21(2), 185–201.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and cognition*, 10(1), 7-21.
- Dirim, I., & Döll, M. (2009). ‚Bumerang‘ – Erfassung der Sprachkompetenzen im Übergang von der Schule in den Beruf – vergleichende Beobachtungen zum Türkischen und Deutschen am Beispiel einer Schülerin. In D. Lengyel, H. H. Reich, H.-J. Roth, & M. Döll (Hrsg.), *Von der Sprachdiagnose zur Sprachförderung* (S. 139–146). Münster: Waxmann.
- Ehlich, K. (2005). Sprachaneignung und deren Feststellung bei Kindern mit und ohne Migrationshintergrund – Was man weiß, was man braucht, was man erwarten kann. In K. Ehlich (Hrsg.), *Anforderungen an Verfahren der regelmäßigen Sprachstandsfeststellung als Grundlage für die frühe und individuelle Förderung von Kindern mit und ohne Migrationshintergrund* (S. 11–63). BMBF.
- Ehlich, K. (2013a). Sprach(en)aneignung – mehr als Vokabeln und Sätze. *Osnabrücker Beiträge zur Sprachtheorie (OBST)*, 83, 21–37.
- Ehlich, K. (2013b). Sprachliche Basisqualifikationen, ihre Aneignung und die Schule. *Die Deutsche Schule (DDS)*, 105(2), 199–209.
- Galloway, P. E., Qin, W., & Uccelli, P. (2020). The role of cross-disciplinary academic language skills in disciplinary, source-based writing: investigating the role of core academic language skills in science summarization for middle grade writers. *Reading and Writing*, 33, 13–44.  
<https://doi.org/10.1007/s11145-019-09942-x>
- Gantefort, C. (2013). *Schriftliches Erzählen mehrsprachiger Kinder – Entwicklung und sprachübergreifende Fähigkeiten*. Münster: Waxmann.
- Gantefort, C., & Roth, H. (2010). Sprachdiagnostische Grundlagen für die Förderung bildungssprachlicher Fähigkeiten. *Zeitschrift für Erziehungswissenschaften (ZfE)*, 13, 573–591.  
<https://doi.org/10.1007/s11618-010-0163-2>
- Gogolin, I., Klinger, T. (2022). Die LiMA-Panelstudie (LiPS): Pilotierung einer Längsschnittuntersuchung zur Sprachentwicklung im Kontext von Mehrsprachigkeit. In T. Klinger, I. Gogolin, & B. Schnoor (Hrsg.) *Sprachentwicklung im Kontext von Mehrsprachigkeit*. Springer VS, Wiesbaden. [https://doi.org/10.1007/978-3-658-36770-1\\_1](https://doi.org/10.1007/978-3-658-36770-1_1)
- Gogolin, I., Klinger, T., Schnoor, B. & Usanova, I. (2022). The competence model of writing skills in MEZ. In H. Brandt, M. Krause & I. Usanova (Hrsg.), *Language development in diverse settings: Interdisziplinäre Ergebnisse aus dem Projekt "Mehrsprachigkeitsentwicklung im Zeitverlauf" (MEZ)*. Wiesbaden: Springer VS.

- Gogolin, I., Klinger, T., Lagemann, M., & Schnoor, B. (2017). *Indikation, Konzeption und Untersuchungsdesign des Projekts Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)* (MEZ Arbeitspapiere No. Nr. 1) Hamburg. Available at: [http://www.pedocs.de/volltexte/2017/14825/pdf/Gogolin\\_et\\_al\\_2017\\_Indikation\\_Konzeption\\_Untersuchungsdesign.pdf](http://www.pedocs.de/volltexte/2017/14825/pdf/Gogolin_et_al_2017_Indikation_Konzeption_Untersuchungsdesign.pdf).
- Gogolin, I., Schnoor, B., Usanova, I. (2021). Crossing the bridge to literacy in foreign languages: C-test as a measure of language development. *Multilingua*. Advance online publication. <https://doi.org/10.1515/multi-2021-0018>.
- Gogolin, I., Klinger, T., Schnoor, B. & Usanova, I. (2021). *Mehrsprachigkeit an der Schwelle zum Beruf. Die Funktion sprachlicher Fähigkeiten für Berufsqualifizierung und Berufseinkommensmündung von Jugendlichen mit und ohne Migrationshintergrund (MEZ-2)* (MEZ Arbeitspapiere Nr. 10). Hamburg: Universität Hamburg. [https://www.pedocs.de/frontdoor.php?source\\_opus=21988](https://www.pedocs.de/frontdoor.php?source_opus=21988).
- Gogolin, I., Dirim, İ., Klinger, T., Lange, I., Lengyel, D., Michel, U., Neumann, U., Reich, H. H., Roth, H.-J., & Schwippert, K. (2011). *Förderung von Kindern und Jugendlichen mit Migrationshintergrund FörMig: Bilanz und Perspektiven eines Modellprogramms*. Waxmann.
- Grabowski, J., Becker-Mrotzek, M., Knopp, M., Jost, J., & Weinzierl, C. (2014). Comparing and combining different approaches to the assessment of text quality. In D. Knorr, C. Heine & J. Engberg (Hrsg.), *Methods in writing process research* (S. 147–165). Lang.
- Harmon, J. M., Hedrick, W. B., & Wood, K. D. (2005). Research on vocabulary instruction in the content areas: Implications for struggling readers. *Reading and Writing Quarterly*, 21, 261–280.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388. <https://doi.org/10.1177/0741088312451260>.
- Heath, N. L. (1995). Distortion and deficit: Self-perceived versus actual academic competence in depressed and nondepressed children with and without learning disabilities. *Learning Disabilities Research & Practice*.
- IEA Hamburg. (2017a). *Methodenbericht. MEZ – Mehrsprachigkeitsentwicklung im Zeitverlauf. Erhebung in den Jahrgangsstufen 7 und 9*. 1. Messzeitpunkt – Januar bis März 2016. Vorgelegt von M. Hellrung, N. Hugk, G. Turezkiy, U. Sievers. Hamburg 2017. IEA Hamburg.
- IEA Hamburg (2017b). *Methodenbericht. MEZ – Mehrsprachigkeitsentwicklung im Zeitverlauf. Erhebung in den Jahrgangsstufen 8 und 10*. 2. Messzeitpunkt – Oktober bis Dezember 2016. Vorgelegt von M. Hellrung, N. Hugk, G. Turezkiy, U. Sievers. Hamburg 2017. IEA Hamburg.
- IEA Hamburg (2018a): *Methodenbericht. MEZ – Mehrsprachigkeitsentwicklung im Zeitverlauf. Erhebung in den Jahrgangsstufen 8 und 10*. 3. Messzeitpunkt – Mai bis Juli 2017. Vorgelegt von M. Hellrung, N. Hugk, G. Turezkiy, U. Sievers. Hamburg 2018. IEA Hamburg. <https://www.mez.uni-hamburg.de/bilder/pdf/mezmethodenbericht3-iea-pdf.pdf>
- IEA Hamburg (2018b): *Methodenbericht. MEZ – Mehrsprachigkeitsentwicklung im Zeitverlauf. Erhebung in den Jahrgangsstufen 9 und 11*. 4. Messzeitpunkt – Mai bis Juni 2018. Vorgelegt von M. Hellrung, N. Hugk, G. Turezkiy, U. Sievers. Hamburg 2018. IEA Hamburg. <https://www.mez.uni-hamburg.de/bilder/pdf/mezmethodenbericht4-iea-pdf.pdf>

- IEA Hamburg (2020). Feld- und Methodenbericht der IEA Hamburg für das Projekt „Mehrsprachigkeit an der Schwelle zum Beruf“ (MEZ-2). 1. Messzeitpunkt. Feldstart: März 2020. Vorgelegt von M. Hellrung, P. Hillen. Hamburg 2021. IEA Hamburg.
- IEA Hamburg (2021). Feld- und Methodenbericht der IEA Hamburg für das Projekt „Mehrsprachigkeit an der Schwelle zum Beruf“ (MEZ-2). 2. Messzeitpunkt. Feldstart: März 2021. Vorgelegt von M. Hellrung, P. Hillen. Hamburg 2021. IEA Hamburg.
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. Working paper, Lund University, *Dept. of Linguistics and Phonetics*, 53, 61–79.
- Jude, N. (2008). *Zur Struktur von Sprachkompetenz*. Dissertation. Frankfurt am Main: Johann Wolfgang Goethe-Universität. <https://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/6694>.
- Klinger, T., Usanova, I., & Gogolin, I. (2019). Entwicklung rezeptiver und produktiver schriftsprachlicher Fähigkeiten im Deutschen. *Zeitschrift für Erziehungswissenschaft (ZfE)*, 22(1), 75–103. <https://doi.org/10.1007/s11618-018-00862-0>.
- Klinger, T. (2022). Die Selbsteinschätzung von Sprachfähigkeiten: Eine Skala zur differenzierten Erfassung. In T. Klinger, I. Gogolin, & B. Schnoor (Hrsg.). *Sprachentwicklung im Kontext von Mehrsprachigkeit* (S. 79-112). Springer VS, Wiesbaden.
- Klinger, T., Usanova, I., Akgün, G., Minh, T.T. (2022). Die Messung produktiver Sprachfähigkeiten im Kontext von Mehrsprachigkeit: Die Herangehensweise in LiPS. In T. Klinger, I. Gogolin, & B. Schnoor (Hrsg.). *Sprachentwicklung im Kontext von Mehrsprachigkeit*. Springer VS, Wiesbaden. [https://doi.org/10.1007/978-3-658-36770-1\\_2](https://doi.org/10.1007/978-3-658-36770-1_2)
- Klinger, T., Usanova, I. & Gogolin, I. (2019). Entwicklung rezeptiver und produktiver schriftsprachlicher Fähigkeiten im Deutschen. *Zeitschrift für Erziehungswissenschaft*, 22, 75–103 (2019). <https://doi.org/10.1007/s11618-018-00862-0>.
- Lanauze, M., & Snow, C. (1989). The relation between first- and second-language writing skills: Evidence from Puerto Rican elementary school children in bilingual programs. *Linguistics and Education*, 1(4), 323–339. [https://doi.org/10.1016/S0898-5898\(89\)80005-1](https://doi.org/10.1016/S0898-5898(89)80005-1).
- Lengyel, D. & Neumann, U. (2016). *Herkunftssprachlicher Unterricht in Hamburg – Eine Studie zur Bedeutung des herkunftssprachlichen Unterrichts aus Elternsicht (HUBE)*. Projektbericht. Universität Hamburg. URL: <https://www.ew.uni-hamburg.de/einrichtungen/ew1/vergleichende/diver/forschung/-files/08122016-bericht-hube-ev.pdf>; Zugriffsdatum: 9.10.2023.
- Mediendienst Integration (2022). *Wie verbreitet ist herkunftssprachlicher Unterricht?* Mediendienst Recherche, Berlin. URL: [https://mediendienst-integration.de/fileadmin/Daten/Factsheet\\_Herkunftssprachlicher\\_Unterricht\\_2022.pdf](https://mediendienst-integration.de/fileadmin/Daten/Factsheet_Herkunftssprachlicher_Unterricht_2022.pdf); Zugriffsdatum: 01.02.2024
- Mehta, P. D., Foorman, B. R., Branum-Martin, L., & Taylor, W. P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading*, 9(2), 85–116.

- Moosbrugger, H., Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion (2nd ed.)* (S. 7–26). Berlin: Springer.
- Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: a study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools*, 36(2), 125–138.
- Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing*, 22(5), 545–565. <https://doi.org/10.1007/s11145-008-9124-z>.
- Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing*, 26, 45–65. <https://doi.org/10.1007/s11145-012-9392-5>.
- Puranik, C. S., Lombardino, L., & Altmann, L. (2008). Assessing the microstructure of written language using a retelling paradigm. *American Journal of Speech Language Pathology*, 17, 107–120.
- Ravid, D., & Levie, R. (2010). Adjectives in the development of text production: Lexical, morphological and syntactic analyses. *First Language*, 30, 27–55. <https://doi.org/10.1177/0142723709350529>.
- Reich, H. H. (2009). Aufbauende Sprachförderung unter Nutzung der FörMig-Instrumente. In D. Lengyel, H. H. Reich, H.-J. Roth, & M. Döll (Hrsg.), *Von der Sprachdiagnose zur Sprachförderung*. FörMig Edition, Band 5 (S. 25–34). Münster: Waxmann.
- Reich, H. H., Roth, H.-J., & Döll, M. (2009). Fast Catch Bumerang – Auswertungshinweise, Schreibimpuls und Auswertungsbogen. In D. Lengyel, H. H. Reich, H.-J. Roth, & M. Döll, (Hrsg.), *Von der Sprachdiagnose zur Sprachförderung*. FÖRMIG Edition, Band 5 (S. 209–241). Münster: Waxmann.
- Risse, S. (2014). Deutsch als Zweitsprache im mehrsprachigen Kontext. Zum Erwerb von Konjunktionen als Indikator für Textqualität. *Zeitschrift für Literaturwissenschaft und Linguistik*, 174, 86–95. <https://doi.org/10.1007/BF03379518>.
- Schlagmüller, M., Ennemoser, M., & Usanova, I. (2022). Diagnostics of reading speed, reading comprehension, and reading accuracy using the LGVT 5–12+. In H. Brandt, M. Krause & I. Usanova (Hrsg.). *Language Development in Diverse Settings: Interdisziplinäre Ergebnisse aus dem Projekt " Mehrsprachigkeitsentwicklung im Zeitverlauf "(MEZ)* (S. 99–132). Wiesbaden: Springer Fachmedien Wiesbaden.
- Schleppegrell, M. J. (2004). Teaching academic writing to English learners. *University of California Linguistic Minority Research Institute (UCLMRI) Newsletter*, 13(2), 1–2.
- K. (2015). Negative covariance in CFA? Discussion via researchgate. net. [https://www.researchgate.net/post/negative\\_covariance\\_in\\_CFA](https://www.researchgate.net/post/negative_covariance_in_CFA).
- Schneider, W., & Tibken, C. (2017). Schriftspracherwerb. In C. Titz, S. Geyer, A. Ropeter, H. Wagner, S. Weber & M. Hasselhorn (Hrsg.), **Entwicklung von Konzepten zur Sprachund Schriftsprachförderung** (S. 69–86). Kohlhammer.

- Schneider, W., Schlagmüller, M., & Ennemoser, M. (2017). *LGVT 5–12+. Lesegeschwindigkeits- und Verständnistest für die Klassen 5–12*. Hogrefe.
- Schnell, R., Hill, P.B., & Esser, E. (2005). *Methoden der empirischen Sozialforschung* (7th ed.). München: Oldenbourg.
- Schnoor, B. (2019). *Soziale Herkunft und Bildungssprache. Humankapitalinvestitionen in deutschen, türkischen und vietnamesischen Familien*. Wiesbaden: Springer VS.
- Schnoor, B. (2022). Vergleichbarkeit der Impulsversionen der „MEZ-Schreibaufgabe Jugendliche“. In H. Brandt, M. Krause & I. Usanova (Hrsg.), *Language development in diverse settings: Interdisziplinäre Ergebnisse aus dem Projekt "Mehrsprachigkeitsentwicklung im Zeitverlauf" (MEZ)* (S.73–98). Wiesbaden: Springer VS.
- Schnoor, B., Hartig, J., Klinger, T., Naumann, A., & Usanova, I. (2023). Measuring the development of general language skills in English as a foreign language—Longitudinal invariance of the C-test. *Language Testing*, 40(3), 796–819. <https://doi.org/10.1177/02655322231159829>.
- Schnoor, B., Usanova, I. (2023). Multilingual writing development: Relationships between writing proficiencies in German, heritage language and English. *Reading and Writing*, 36, 599–623 <https://doi.org/10.1007/s11145-022-10276-4>.
- Schrauf, R.W. (2009). Longitudinal designs in studies of multilingualism. In K. de Bot, & R. W. Schrauf (Hrsg.), *Language Development over the Lifespan* (S. 245–270). New York: Routledge.
- Schründer-Lenzen, A. (2007). *Schriftspracherwerb und Unterricht. Bausteine professionellen Handlungswissens*. Leske & Budrich.
- Stavans, A., Seroussi, B., & Zadunaisky Ehrlich, S. (2019). Literacy-related abilities' effects on argumentative text quality structure. *Journal of Literacy Research*, 51(3), 315–335. <https://doi.org/10.1177/1086296X19859515>
- Usanova, I. (2019). *Biscriptuality: Writing skills among German-Russian adolescents* (Hamburg Studies on Linguistic Diversity 8). Amsterdam: John Benjamins.
- Verhoeven, L. (2006). Sociocultural variation in literacy achievement. *British Journal of Educational Studies*, 54(2), 189–211. <https://doi.org/10.1111/j.1467-8527.2006.00341.x>
- Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing*, 24(2), 203–220. <https://doi.org/10.1007/s11145-010-9266-7>

## Anhang

**Tabelle 7:** Korrelationen zwischen Schreiben und Selbsteinschätzungen an den Messzeitpunkten 1, 3 und 6

Schreiben	Selbsteinschätzungen		
	MZP1	MZP3	MZP6
<b>Deutsch</b>	.29**	.28**	.20**
<i>n</i>	1708	1558	541
<b>FS Englisch</b>	.48**	.40**	.34**
<i>n</i>	1688	1553	530
<b>FS Französisch</b>	.37**	.56**	--
<i>n</i>	643	549	--
<b>HS Russisch</b>	.59**	.47*	.50**
<i>n</i>	238	180	79
<b>HS Türkisch</b>	.50**	.41**	.41**
<i>n</i>	402	388	98

Gemessen wurde jeweils der Zusammenhang zwischen den Testergebnissen der Schreibaufgabe mit den Selbsteinschätzungen in derselben Sprache. Die Korrelationen sind auf dem Niveau von \*  $p < .05$  und \*\*  $p < .001$  (2-seitig) signifikant.

**Tabelle 8:** Korrelationen zwischen Selbsteinschätzungen in unterschiedlichen Sprachen am 1. Messzeitpunkt

Selbsteinschätzung	Selbsteinschätzungen			
	Deutsch	FS Englisch	FS Französisch	FS Russisch
<b>FS Englisch</b>	.47**	--	--	
<i>n</i>	1689	--	--	
<b>FS Französisch</b>	.33**	.38**	--	
<i>n</i>	637	638	--	
<b>FS Russisch</b>	.37*	.35*	--	
<i>n</i>	59	58	--	
<b>HS Russisch</b>	.22**	.21**	.07	.98**
<i>n</i>	266	261	102	42
<b>HS Türkisch</b>	.18**	.26**	.17*	--
<i>n</i>	392	385	167	2

Die Korrelationen sind auf dem Niveau von \*  $p < .05$  und \*\*  $p < .001$  (2-seitig) signifikant.