

Zur diagnostischen Eignung von deutschen Rechtschreibnormen in der Deutschschweiz

Exemplarische Studie zur Hamburger-Schreib-Probe aus dem Kanton Freiburg

Erich Hartmann, Julia Winkes, Felix Studer
Universität Freiburg/Schweiz

Zusammenfassung: In der vorliegenden Studie wird die Frage der diagnostischen Angemessenheit der neu normierten Hamburger-Schreib-Probe (HSP; May, 2013) für Schweizer Schülerinnen und Schüler exemplarisch in Deutschfreiburg untersucht. Ein Vergleich der Freiburger Stichprobe mit der HSP-Eichstichprobe hinsichtlich quantitativer und qualitativer Testindikatoren ergab eine Reihe von Unterschieden, wobei die deutsche Stichprobe insgesamt mehr Leistungsvorteile verzeichnete als die Freiburger. Die Schweizer Schülerinnen und Schüler weisen auf mehreren Klassenstufen bessere alphabetische Schreibfertigkeiten auf, sind den deutschen Kindern jedoch im orthographischen Schreiben fast durchwegs unterlegen. Die Anwendung von HSP-Originalnormen führt zu einer Unterschätzung der Rechtschreibleistung von Freiburger Primarschulkindern und generiert sehr unterschiedliche stufenspezifische Anteile von als rechtschreibschwach eingestuften Schulkindern. Die überprüften HSP-Normen passen folglich nur bedingt zu den Rechtschreibleistungen der Freiburger Stichprobe. Eine zusätzliche Mehrebenenanalyse konnte die prädiktive Bedeutung von einigen Individualfaktoren für die Schreibkompetenz belegen, jedoch nicht für Merkmale des Klassenkontextes.

Schlüsselbegriffe: Rechtschreibschwierigkeiten, Diagnostik, Test, Hamburger-Schreib-Probe, Deutschschweiz

On the Appropriateness of Applying German Norms of Spelling Skills in German-Speaking Switzerland: An Exemplary Study with the Hamburger-Schreib-Probe in the Canton of Fribourg

Summary: To investigate the diagnostic appropriateness of the newly standardized Hamburger-Schreib-Probe (HSP; May, 2013), we applied it to Swiss-German students in the canton of Fribourg. In respect to quantitative and qualitative indicators the comparison of the Fribourg sample with the original sample resulted in a range of differences. Regarding most aspects, the school children from the original sample perform better than the Swiss children. While Fribourg students across several grades show better alphabetic writing skills, they lack behind in the development of orthographic writing skills in comparison to the German sample. Applying the original norms of the HSP leads to an underestimation of the spelling skills of the primary school children from German-speaking Switzerland, resulting in very different, grade-specific proportions of students with spelling difficulties. Therefore, the original norms are of limited suitability for the spelling skills of the Fribourg sample. An additional multilevel modeling revealed the predictive relevance of several individual variables, whereas characteristics of the classroom had no effect on the writing skills.

Keywords: Spelling difficulties, assessment, test, Hamburger-Schreib-Probe, German-speaking Switzerland

1 Problemaufriss und Ziele des Freiburger HSP-Projektes

Trotz vermehrter Bemühungen um die Prävention von Lese-Rechtschreibschwierigkeiten (LRS) sind Lehrpersonen und andere Fachleute weiterhin mit Schülerinnen und Schülern mit relevanten Schriftsprachproblemen konfrontiert. Unabhängig davon, ob es sich hierbei um eine umschriebene Entwicklungsbeeinträchtigung oder um eine allgemeine Lese-Rechtschreibschwäche handelt, bergen signifikante Schwierigkeiten beim Schriftspracherwerb ein Risiko für die schulische, berufliche und psychosoziale Entwicklung der Betroffenen (Klicpera, Schabmann & Gasteiger-Klicpera, 2013; Kohn, Wyschkon, Ballaschk, Ihle & Esser, 2013; Kohn, Wyschkon & Esser, 2013; Endlich, Dummett, Schneider & Schwenck, 2014). Dies stellt eine erhebliche Herausforderung für das Bildungssystem dar, welches ein Angebot an möglichst frühen und effektiven Interventionen sicherzustellen hat.

Laut aktuellen Forschungssynthesen können Schülerinnen und Schüler mit LRS von gezielten Interventionen zum Lesen und Rechtschreiben profitieren (Galuschka, Ise, Krick & Schulte-Körne, 2014; Ise, Engel & Schulte-Körne, 2012). Grundlage für die Planung aussichtsreicher Maßnahmen ist eine fundierte Diagnostik. Diese erweist sich als komplex, zumal neben der Kernsymptomatik im Lesen und Schreiben auch Informationen zu kognitiven und linguistischen Lernvoraussetzungen sowie zur lernbezogenen bzw. schulischen und psychischen Situationen der betroffenen Schulkinder zu erfassen sind, um Ansatzpunkte für individualisierte Interventionen gewinnen zu können. Eine mehrdimensionale Diagnostik unter Einbezug unterschiedlicher Methoden erfolgt in der Regel interdisziplinär, wobei in der Schweiz vorrangig Logopädinnen, Sonderpädagoginnen und Schulpsychologen in den diagnostischen Prozess involviert sind.

Entsprechend aktuellen Empfehlungen bzw. Leitlinien zur LRS-Diagnostik bilden standardisierte Tests zum Lesen und Rechtschreiben einen unverzichtbaren Bestandteil des diagnostischen Methodeninventars. Solche Verfahren dienen nicht nur der quantitativen Messung und Einordnung von kindlichen Schriftsprachleistungen, sondern ermöglichen vorteilhaft auch qualitative Fehleranalysen (Schulte-Körne, 2010; Klicpera et al., 2013; DGKJP, 2015).

Für die Diagnostik von Rechtschreibleistungen, die im Fokus dieses Beitrags stehen, existieren verschiedene Tests, „die sehr gut etabliert und weit verbreitet sind, unter Umständen aber veraltete Normen und aus der Mode gekommene Wörter umfassen. Für die Auswahl geeigneter Verfahren empfiehlt es sich, jeweils die Passung mit der didaktischen Herangehensweise zu überprüfen und Verfahren zu verwenden, die innerhalb der letzten 10 Jahre normiert wurden. Dies ist leider bei den (...) erhältlichen standardisierten Verfahren nur selten der Fall (...)“ (Lenhard & Lenhard, 2017, S. 192).

Ein vor wenigen Jahren neu normierter deutscher Rechtschreibtest ist die *Hamburger Schreib-Probe* (HSP 1 – 10; May, 2013). Dieses Instrument, das in der DGKJP-Leitlinie zur LRS-Diagnostik aufgeführt wird, gelangt in der Praxis – auch in der Schweiz – häufig zur Anwendung. Dies dürfte nicht nur der praktikablen Gestaltung des Testmaterials und der Auswertung sowie der breiten Altersspanne (Schuleintritt bis mittlere Reife) geschuldet sein, sondern auch der Tatsache, dass die HSP neben quantitativen Ergebnissen auch qualitative bzw. förderrelevante Informationen liefert (Lenhard & Lenhard, 2017).

Eine Besonderheit bzw. Unsicherheit bei der diagnostischen Nutzung – nicht nur – der HSP in der Deutschschweiz ergibt sich daraus, dass dieser Test in Deutschland entwickelt und nor-

miert wurde. In Anbetracht des evidenten Mangels an ländervergleichender Forschung bezüglich kindlicher Rechtschreib(test)leistungen stellt sich in der Praxis immer wieder die Frage nach der Übertragbarkeit von in Deutschland (oder Österreich) gewonnenen Testnormen auf die LRS-Diagnostik in der Schweiz. Diese Frage ist von praktischer Relevanz, zumal im Falle von systematischen Leistungsunterschieden zwischen diesen Ländern die diagnostische Aussagekraft des jeweiligen Verfahrens bei Schweizer Kindern infrage gestellt wäre bzw. die Gefahr bestünde, dass deren Rechtschreibleistungen anhand der Originalnormen unter- oder überschätzt werden – mit möglichen Konsequenzen für die Zuweisung von sonderpädagogischen Betreuungsressourcen.

Dass die Passung von Normen aus umliegenden deutschsprachigen Ländern nicht einfach als gegeben vorausgesetzt werden kann, zeigen Erfahrungen mit dem *Salzburger Lese- und Rechtschreibtest II (SLRT)*, für den seit 2014 gesonderte Schweizer (Berner) Normen vorliegen (Moll & Landerl, 2014). Diese zusätzliche Normierung wurde initiiert aufgrund der Beobachtung, „dass die Normen des SLRT-II in der Berner Praxis zu streng ausfallen und verhältnismäßig zu viele Schülerinnen und Schüler als rechtschreibschwach klassifiziert werden“ (Moll & Landerl, 2014, S. 89). Die Schweizer Normierung erfolgte anhand einer repräsentativen Stichprobe (N = 5808) von Berner Schülerinnen und Schülern (1. – 4. Klasse). In der Tat zeigt ein Vergleich der Originalnormen und der Berner Normen, dass die Schweizer Kinder aller Altersklassen mit den deutschen Normen tendenziell zu streng eingestuft würden.

Was die HSP betrifft, erfolgte im Jahr 2000 eine partielle Überprüfung der damaligen deutschen Normen (1998) für die Schweiz im Kanton Solothurn. Daraus ging hervor, dass die Passung der deutschen Normen für die Solothurner Primarschulkinder je nach Klassenstufe variierte: Während sie in den Klassen 1, 2 und 4 „prak-

tisch gleich gut“ abschnitten wie die deutsche Eichstichprobe, erreichten die Schweizer Schulkinder der Klassenstufen 3 und 5 durchschnittlich bessere Werte (Meyer, 2000, S. 21).

Die 2012 erfolgte Neunormierung der HSP gab den Anlass für das darzustellende Forschungsprojekt aus dem deutschsprachigen Teil des Kantons Freiburg/CH. Die Studie verfolgt in erster Linie das Ziel, die diagnostische Angemessenheit der Anwendung der neuen HSP-Normen auf Deutschschweizer Primarschulkinder exemplarisch zu evaluieren. Darüber hinaus sollen anhand der Stichprobendaten Zusammenhänge zwischen individuellen bzw. klassenstrukturellen Faktoren und der kindlichen Rechtschreibleistung exploriert werden.

Der empirischen Studie (Kap. 5) gehen Hintergrundinformationen zur HSP (Kap. 2) voraus, bevor relevante Faktoren für potenzielle schriftsprachliche Leistungsunterschiede zwischen Schweizer Kindern und Schülern aus deutschsprachigen Nachbarländern erörtert (Kap. 3) und individuelle und unterrichtliche Einflüsse auf den Orthografieerwerb beleuchtet (Kap. 4) werden.

2 Die Hamburger-Schreib-Probe – ein verbreitetes Instrument zur Rechtschreibdiagnostik

2.1 Ziel, Testversionen und Untersuchungsmaterial

Die HSP zielt darauf ab, „das Rechtschreibkönnen und die grundlegenden Rechtschreibstrategien“ (May, 2013, S. 7) von Schülerinnen und Schülern der Klassenstufen 1 – 10 einzuschätzen und dabei insbesondere im unteren Leistungsbereich zu differenzieren, gleichzeitig aber auch genügend hohe Anforderungen an fortgeschrittene Rechtschreiber zu stellen. Hierzu existieren sechs Versionen mit steigen-

dem Schwierigkeitsgrad: HSP 1+, HSP 2, HSP 3, HSP 4 – 5, HSP 5 – 10 B (Basisanforderungen) und HSP 5 – 10 EK (Erweiterte Kompetenz). Alle Tests, die in Gruppen oder mit einzelnen Schülern durchgeführt werden können, bestehen aus Einzelwörtern und Sätzen, deren Anzahl mit steigender Klassenstufe zunimmt. Da den Schulkindern die Wörter und Sätze seitenweise vorgelesen werden, sind die zu schreibenden Elemente als Erinnerungsstütze mit Illustrationen versehen. Die Testwörter beziehen sich nicht auf einen bestimmten Grund- oder Übungswortschatz. Vielmehr wurden Schreibwörter gewählt, welche die wichtigsten orthografischen Elemente enthalten bzw. grundlegende Rechtschreibstrategien überprüfen (May, 2009; 2013).

2.2 Testauswertung

Dem Testanwender stehen zwei quantitative und sechs qualitative Kennwerte zur Verfügung: Die *quantitative Auswertung* erfolgt anhand der Anzahl *richtig geschriebener Wörter* und der *Graphemtreffer*. Der letztgenannte Kennwert bildet den Hauptindikator der HSP, der „das Maß für den erreichten Stand des Rechtschreibkönnens (bündelt)“ und eine differenziertere Einschätzung kindlicher Rechtschreibkompetenzen erlaubt als die wortbezogene Auswertung (May, 2013, S. 18).

Die *qualitative Analyse* bezieht sich auf das Konzept der Rechtschreibstrategien. Laut May (2013, S. 22) erfolgt der Orthografieerwerb „als ein schrittweiser Aufbau immer komplexer werdender Zugriffsweisen auf die Schriftsprache“. Vier unterschiedliche Rechtschreibstrategien werden mittels sogenannter Lupenstellen erfasst: *alphabetische* (lautorientiertes Verschriften), *orthografische* (Verschriften von Länge- und Kürzezeichen u. a.), *morphematische* (Verschriften von Vorsilben, Umlautableitung, Auslautverhärtung u. a.) und *wortübergreifende Strategie* (Satzzeichen, Groß-/Kleinschreibung u. a.). Zusätzlich können *überflüssige orthografische*

Elemente (Länge-/Kürzezeichen u. a.) und *Oberzeichenfehler* (Umlautzeichen u. a.) analysiert werden. Für alle diese Indikatoren stellt die HSP klassenstufenspezifische Normwerte (Prozentrang, T-Wert) zur Verfügung (May, 2013, S. 17ff.).

2.3 Normierungen

Eine bundesweite Normierung der HSP mit 23'000 Schülerinnen und Schülern aller Klassenstufen (1–9) erfolgte erstmals im Jahre 2001. Von verschiedenen Seiten wurden in der Folge Zweifel an der Angemessenheit der Normen geäußert. Deimel (2002) bemerkte, dass bei vielen Anwendern der Eindruck entstand, die HSP sei um ca. 5 T-Punkte zu leicht. Auch Tacke, Völker und Lohmüller (2001, S. 135) kritisierten, der Test sei zu einfach und würde nicht mit anderen Rechtschreibtests übereinstimmen. Die Konsequenz daraus sei, dass mit der HSP „viele rechtschreibschwache Schüler nicht erkannt werden“.

Aufgrund von „Veränderungen in der deutschen Gesellschaft und in der Bildungslandschaft“ (May, 2013, S. 5) fand im Jahr 2012 eine Neunormierung statt, im Rahmen derer auch der Altersbereich auf die zehnte Klassenstufe ausgeweitet wurde. Die Eichung der verschiedenen HSP-Versionen erfolgte mittels eines Online-Auswertungssystems, wobei lediglich die Testdaten aus vollständigen Klassen bzw. Lerngruppen herangezogen wurden. Über alle Klassenstufen hinweg umfasst die Eichstichprobe etwa 270'000 Kinder und Jugendliche, zu denen sich im Testhandbuch keine weiteren Angaben (Alter, Geschlechtsverteilung, Anteil mehrsprachiger Kinder u. a.) finden. Es werden sowohl bundesweite Vergleichswerte als auch gesonderte Normen für Stadtstaaten und großstädtische Ballungsräume angeboten. Während für die HSP 1+, HSP 2, HSP 3 und HSP 4–5 Normen für zwei Testzeitpunkte (Mitte/Ende Schuljahr) existieren, werden für

die HSP 5 – 10 lediglich Normen für das Schuljahresende ausgewiesen. Die Daten für die einzelnen Klassenstufen wurden mit der bundesweiten IGLU-Studie (Kowalski & Voss, 2009) und der KESS-Längsschnittstudie (May, 2009) abgeglichen, wobei sich übereinstimmende Werteverteilungen ergaben. May (2013, S. 13) kommt daraufhin zu dem Schluss, die HSP messe die Leistungsverteilung „präzise und korrekt“.

Nach Lenhard & Lenhard (2017, S. 193) ist unklar, ob das Problem einer zu positiven Leistungsbeurteilung mit der Neunormierung beseitigt werden konnte, zumal es sich bei der aktuellen Eichstichprobe um keine repräsentative Stichprobe handle. Ungeachtet dieses Einwandes gelangt die HSP in der Praxis häufig zur Anwendung.

2.4 Gütekriterien

Die *Reliabilität* wurde durch verschiedene Verfahren ermittelt. Die interne Konsistenz (Cronbach's Alpha) für das Kernkriterium Graphemtreffer beträgt zwischen .92 und .99, was als sehr hoch einzustufen ist. Dasselbe gilt für die Stabilität der Graphemtreffer-Ergebnisse. So beträgt beispielsweise der Korrelationskoeffizient zwischen den Testzeitpunkten Mitte 4. Klasse und Ende 4. Klasse $r_{tt} = .92$. Die übrigen Ergebnisse für die Stabilität liegen je nach Abstand der Messzeitpunkte zwischen $r_{tt} = .52$ und .93. Die Zuverlässigkeit der HSP ist insgesamt als recht hoch einzuschätzen (May, 2013, S. 97ff.).

Um die *Gültigkeit* zu überprüfen wurden u. a. Untersuchungen zur theoretischen Plausibilität des Testkonzepts und zur Übereinstimmung mit Expertenurteilen und mit anderen Rechtschreibtests durchgeführt. Die Übereinstimmungsvalidität zwischen der HSP (Ende 4. Klasse) und dem Weingartener Grundwortschatz Rechtschreib-Test (4/5) beträgt $r = .87$. Die Kor-

relationen zwischen der HSP und der Rechtschreibleistung in Deutschaufsätzen bewegen sich zwischen $r = .78$ und .81. Diese Befunde indizieren, dass die durch die HSP ermittelten Leistungen eine gute Validität aufweisen bzw. „in hohem Maße verallgemeinerbar sind“ (May, 2013, S. 54).

Zur *Durchführungsobjektivität* werden im Handbuch keine Angaben gemacht. In den Hinweisheften der einzelnen HSP-Versionen sind jedoch detaillierte Durchführungsanweisungen beschrieben. Somit ist die Objektivität der Testdurchführung bei korrekter Umsetzung gegeben. Die *Auswertungsobjektivität* wurde anhand der Auswertung von 120 Tests durch drei Auswerter überprüft. Für die wortbezogene Analyse wurde eine Interrater-Übereinstimmung von 99 % erreicht, für die Graphemstellen resultierten weniger als 0.2 % abweichende Bewertungen (May, 2013, S. 106).

3 Annahmen und Erklärungen zu möglichen Unterschieden im Orthografieerwerb von Schweizer und deutschen Schulkindern

Das vorliegende Forschungsprojekt impliziert, dass es in den Rechtschreibkompetenzen von Schweizer und deutschen (bzw. österreichischen) Primarschulkindern gewisse Unterschiede gibt, eine Annahme, die durch bisherige Praxisbeobachtungen, Erfahrungen mit der Schweizer Normierung des SLRT II sowie vereinzelte Studien begründet wird. Daraus folgt die Frage, welche Faktoren zu allfälligen rechtschreibbezogenen Leistungsunterschieden zwischen Schweizer und deutschen Schulkindern beitragen könnten. Geht man davon aus, dass methodische Gründe für Leistungsunterschieden in ländervergleichenden Studien auszuschließen sind (Unterschiede bezüglich Stichproben, Testzeitpunkten u. a.), lassen sich

zur Erklärung insbesondere *sprachbiografische* und *unterrichtlich-curriculare Einflussfaktoren* in Betracht ziehen.

In Zusammenhang mit der Anwendung von deutschen Rechtschreibtests (und Sprachtests) in der Deutschschweiz interessiert zunächst die Frage, inwieweit die besondere (diglossische) Sprachsituation einen Einfluss auf den Rechtschreiberwerb von Schweizer Kindern hat. Denn nicht nur in Deutschfreiburg, sondern auch in der restlichen Deutschschweiz „verfügen die einheimischen Sprecherinnen und Sprecher über zwei Sprachformen des Deutschen, über einen schweizerdeutschen Dialekt und das Hochdeutsche, die strikt auseinandergehalten werden“ (Christen, Glaser & Friedli, 2013, S. 22). Während in der mündlichen Alltagskommunikation fast durchgehend schweizerdeutsche Dialekte – in Deutschfreiburg Senslerdeutsch oder Berndeutsch – verwendet werden, hat das (Schweizer-)Hochdeutsche „seinen Platz in der Schule, in der Kirche oder in bestimmten massenmedialen Sendeformaten, und es wird für Vorträge und Reden gewählt“ (Christen et al., 2013, S. 24). Was die schriftsprachliche Kommunikation angeht, erfolgt diese bis auf Ausnahmen (SMS, Chat u. a.) auf Hochdeutsch.

Das in der Schweiz gesprochene (*Schweizer-*) *Hochdeutsch* (Shdt.) ist aus linguistischer Sicht eine Varietät der deutschen Standardsprache. Sie weicht von der binnen- oder bundesdeutschen Varietät (Hdt.) in bestimmten linguistischen Eigenschaften ab. Unterschiede zeigen sich im Wortschatz (z. B. Shdt.: Metzger/Hdt.: Fleischer), in der Grammatik (z. B. Shdt.: Sie ruft ihrem Hund/Hdt.: Sie ruft ihren Hund; Shdt.: Bögen/Hdt.: Bogen), in der Prosodie bzw. Wortbetonung (Shdt.: *Motor*/Hdt.: *Motor*) und in der Aussprache. Diesbezügliche Besonderheiten des Schweizerhochdeutschen betreffen mitunter die Artikulation von a) Kurzvokalen (Shdt.: geschlossen, z. B. Bett = [bet]/Hdt.: offen, [bet]), b) unbetonten

Vokalen im Auslaut (Shdt.: Vokal zwischen [e], [ɛ] und [ə]; Hdt.: zentrales Schwa [ə]), c) auslautenden Obstruenten (Shdt.: keine Verhärtung, z. B. Lob = [lo:b])/Hdt.: Auslautverhärtung, [lo:p]) sowie d) des R-Lautes vor Konsonanten oder im Auslaut (Shdt.: nicht vokalisiert, z. B. nur = [nu:r])/Hdt.: vokalisiert, [nu:ʀ]) (Hove, 2002; 2017; Landert, 2007).

Rezeptiv haben Deutschschweizer Kinder schon vor dem Schuleintritt vielfältige potenzielle Möglichkeiten, Kontakte mit dem Hochdeutschen zu knüpfen (in Vorlesesituationen, beim Fernsehen u. a.), und ihr Verständnis von standardsprachlichen Äußerungen ist schon beachtlich entwickelt. Produktiv hingegen zeigen die Kinder nur in wenigen, ausgewählten Situationen (z. B. Rollenspiel, Nachahmen von Filmszenen) hochdeutsche Sprechanteile, die auf bemerkenswerte Fertigkeiten schließen lassen (Häcki Buhofer & Burger, 1998; Landert, 2007). In der Schule, wo Standarddeutsch seit einigen Jahren von Anfang an konsequent die Unterrichtssprache ist und der Hochdeutscherwerb der Kinder schulisch gesteuert erfolgt, bevorzugen viele Kinder die schweizerhochdeutsche Varietät, obgleich sie „dank dem regelmäßigen Konsum deutscher Fernsehsender ohne weiteres eine dem deutschen Standard angenäherte Aussprache beherrschen würden“ (Bickel, 2001, S. 20).

Was nun den Orthografierwerb betrifft, ist wesentlich, dass vor allem die *alphabetische Phase* zu Beginn der Schreibentwicklung (lautorientiertes Verschriften) dadurch beeinflusst wird, wie nahe die kindliche Aussprache des Standarddeutschen an die geschriebene Sprache heranreicht. Starke Dialektsprecher halten beide Sprachformen klar auseinander und lernen gezielt, sich beim Rechtschreiben an der Standardlautung zu orientieren. Wo hingegen ein Kontinuum zwischen Hochlautung und regionaler Aussprache besteht, ist das Prinzip

„Schreibe, wie Du sprichst“ oft eine irreführende Strategie (May, Malitzky & Vieluf, 2001; Hove, 2002). Tatsächlich zeigte sich dieses Bild bei Auswertungen von HSP-Schreibproben hinsichtlich des Anteils und der Art von alphabetischen Fehlern bei Kindern der drei dialektalen Hauptregionen des Deutschen (May, 2013). In Gebieten mit starken Dialekteinflüssen schneiden Schülerinnen und Schüler in der alphabetischen Strategie demnach tendenziell besser ab, was auch eine Untersuchung von Schmidlin (2003) mit südwestdeutschen, norddeutschen und Schweizer Schreibanfängern belegt. Für Kinder aus der Deutschschweiz bzw. Deutschfreiburg sind vor diesem Hintergrund gewisse Vorteile beim Erwerb des lautorientierten Schreibens zu erwarten, da sich die schweizerhochdeutsche Aussprache enger an der geschriebenen deutschen Sprache orientiert als das Bundeshochdeutsche – z. B. durch fehlende Vokalisation des R-Lautes und durch Artikulation auslautender Vokale und Konsonanten (Häcki Buhofer & Burger, 1998; Landert, 2007).

Zu möglichen Unterschieden zwischen deutschen und Schweizer Schulkindern im Erwerb von orthografischem und morphematischem Wissen lassen sich weniger eindeutige Annahmen formulieren als zur alphabetischen Strategie. Für die Aneignung dieser Aspekte der Orthografie dürften unterrichtlich-curriculare Faktoren eine wichtigere Rolle spielen als der Dialekteinfluss. Dabei ist es plausibel, dass allfällige Leistungsunterschiede im Rechtschreiben von Schweizer und deutschen Schulkindern insbesondere (mit-)bedingt werden durch mögliche Unterschiede in Bezug auf die in den jeweiligen Ländern zur Anwendung gelangenden (heterogenen) Curricula, Ansätze/Methoden und Lehrmittel zum Thema Rechtschreiben. Wie zu dem Faktorenbündel Unterricht anzumerken bleibt, stehen vergleichende Studien aus, die über Umfang, Art und Qualität des Rechtschreibunterrichts in Schweizer und deutschen Schulen zuverlässig Auskunft geben könnten.

4 Individuelle und unterrichtliche Einflüsse auf den Orthografieerwerb

Der Erwerb von schulischen Fertigkeiten bzw. Schriftsprachkompetenzen erfolgt eingebettet in ein komplexes Bedingungsgefüge familiärer, schulischer und individueller Wirkfaktoren (Deimel, Ziegler & Schulte-Körne, 2005; Klicpera et al., 2013; Brügelmann, 2015). Die nachfolgenden Ausführungen fokussieren auf Einflussvariablen aus den Bereichen *Individuum* und *Unterricht*, die für die eigene Studie von Relevanz sind.

4.1 Individuelle Faktoren

Hierzu gehören neben motivationalen Aspekten vor allem spezifische kognitive und linguistische Voraussetzungen bzw. Verarbeitungsfähigkeiten für das Lesen- und Schreibenlernen, deren Bedeutung über die internationale Forschung (vgl. Kamhi & Catts, 2012) hinaus auch für den deutschsprachigen Raum empirisch belegt ist (z. B. Ennemoser, Marx, Weber & Schneider, 2012). Während gut entwickelte (*meta-*)sprachliche Kompetenzen den Schriftspracherwerb begünstigen, bergen sprachliche Defizite in Zusammenhang mit Sprachentwicklungsstörungen (z. B. Hartmann, 2017) oder Migrationshintergrund ein erhöhtes Risiko für Lernprobleme: „Insgesamt verdeutlichen die Ergebnisse empirischer Studien (...), dass Kinder und Jugendliche, die Deutsch als Zweitsprache erwerben, in ihren Lese- und Rechtschreibleistungen oftmals hinter ihren monolingualen deutschsprachigen Peers zurückbleiben“ (Marx, 2017, S. 332). So zeigte sich auch anhand der HSP-Eichstichprobe aus dem Jahr 2001, dass zweisprachige Lernende insgesamt deutlich geringere Rechtschreibleistungen aufweisen als Schülerinnen und Schüler deutscher Muttersprache. Wie ein Vergleich zwischen zweisprachig aufwachsenden Kindern

mit europäischen bzw. außereuropäischen Herkunftssprachen indes nahelegt, „ist die Zweisprachigkeit nicht generell Ursache für schwache Rechtschreibleistungen“, die vielmehr auch von ungünstigen individuellen und/oder sozio-kulturellen Faktoren begünstigt werden (May, 2013, S. 76).

Mit Blick auf individuelle Einflüsse auf das Rechtschreibenlernen interessieren weiter *Geschlechterspezifika*. Laut Hattie (2014) sind die geschlechtsspezifischen Unterschiede im Lernen insgesamt gering ($ES = 0.12$). Die bei Philipp (2017) referierten spezifischen Studien zum (Recht-)Schreiben liefern empirische Evidenz für „eindeutige Vorsprünge zugunsten der Mädchen, die nicht nur orthografisch korrektere Texte verfassen, sondern auch in besserer Qualität“ (ebd., S. 328). Entsprechend belegen auch die Daten der HSP-Normierungsstichprobe von 2001, dass Mädchen ab der zweiten Klasse bessere orthografische Leistungen zeigen. Diese Unterschiede werden bis zur neunten Klasse ausgeprägter, sodass Mädchen im Laufe der Schulzeit einen größeren Lerngewinn als Jungen aufweisen (May, 2013). Eine solche Zunahme des Geschlechtseffekts fanden auch Deimel et al. (2005). May (2013, S. 70) weist aber ebenfalls darauf hin, dass die Streuung innerhalb der beiden Gruppen wesentlich größer ist als die Streuung zwischen den Gruppen, weshalb „im individuellen Fall die Geschlechtszugehörigkeit relativ wenig über die Rechtschreibfähigkeit aussagt“.

4.2 Unterrichtlich-curriculare Einflüsse

Ein häufig und kontrovers diskutierter Faktor in Zusammenhang mit dem Klassenkontext als Rahmenbedingung für konkreten Unterricht ist die Klassengröße. Wie die Forschungssynthese von Hattie (2014) dokumentiert, bewirken strukturelle Maßnahmen wie die Reduktion der Klassengröße für sich keine nennenswerten Lernvorteile ($ES = 0.21$). Der Einfluss der Klas-

sengröße auf den Schriftspracherwerb ist auch abhängig von anderen Merkmalen wie der Kompetenz der Lehrperson, der Unterrichtsform bzw. -methode oder der Klassenstufe (Martschinke & Kammermeyer, 2003; Treutlein, Roos & Schöler, 2012). In der Untersuchung von Treutlein et al. (2012) beeinflussten weder die Klassengröße noch der Klassenanteil an Kindern mit unzureichenden Deutschkenntnissen das orthografische Leistungsniveau von Drittklässlern. Individuelle Deutschkompetenzen hingegen erklärten über 40 % der Varianz der Rechtschreibleistung, wobei dieser Einfluss (anders als beim Lesen) nicht durch Klassenmerkmale moderiert wurde.

Bereits zu Schulbeginn unterscheiden sich Klassen hinsichtlich der sprachlich-kognitiven und schriftsprachlichen Voraussetzungen der Schulkinder (Klicpera & Gasteiger-Klicpera, 1993; Martschinke & Kammermeyer, 2003). Entsprechend fanden sich schon in der Wiener Längsschnittstudie von Klicpera und Gasteiger-Klicpera (1993) teilweise deutliche Leistungsunterschiede im Lesen und Rechtschreiben zwischen verschiedenen Klassen, selbst wenn diese demselben Schulhaus angehörten. Da diese Unterschiede nur bedingt durch Individualfaktoren erklärbar sind, liegt es nahe, dass vor allem die konkrete Gestaltung bzw. die Qualität des Unterrichts einen wesentlichen Einfluss auf die kindlichen Schriftsprachkompetenzen ausübt.

Obgleich „der Erwerb der Orthographie zu komplex (ist), als dass nur ein didaktischer Zugang diesen fördern könnte“ (Lindauer & Schmellentin, 2017, S. 33; vgl. Brügelmann, 2015), scheint unter Experten für Schriftsprachdidaktik und LRS-Förderung überwiegend Konsens zu bestehen, dass ein (direkter) Unterricht, in dem Rechtschreibkompetenzen gezielt und systematisch aufgebaut, erweitert und gefestigt werden, für die meisten Lernenden, vor allem aber für schwächere Kinder, besonders hilfreich ist (z. B. Klicpera et al., 2013; Lindauer & Schmellentin, 2017; Sturm, 2017).

Abschließend bleibt zu betonen: Individuelle, unterrichtsbezogene und familiäre Faktoren entfalten ihren Einfluss auf den Schriftspracherwerb letztlich gemeinsam und sollten daher nicht isoliert betrachtet werden. Dies gestaltet sich jedoch de facto als schwierig. Obgleich besonders im englischsprachigen Raum zahlreiche empirische Arbeiten zu Determinanten von Schulleistungen vorliegen, beschränkt sich der aktuelle Erkenntnisstand überwiegend auf Einzelfaktoren, wohingegen Aussagen über Interaktionen zwischen verschiedenen Bedingungen „auf der Strecke bleiben“ (Deimel et al., 2005, S. 5).

5 Empirische Studie

5.1 Fragestellungen

Aus der einleitend skizzierten Zielsetzung des HSP-Projektes ergeben sich die folgenden Forschungsfragen:

1. Unterscheiden sich die HSP-Rechtschreibleistungen von Freiburger Primarschulkindern signifikant von denjenigen von deutschen Schülerinnen und Schülern (HSP-Normierungsstichprobe)?
2. Differiert das individuelle Testergebnis in Abhängigkeit von der Referenzstichprobe (Freiburger Stichprobe vs. HSP-Eichstichprobe)?
3. Werden je nach Bezugsgruppe unterschiedliche Anteile an Kindern mit schwacher Rechtschreibleistung identifiziert?
4. Inwieweit bestehen Zusammenhänge zwischen individuellen sowie klassenbezogenen Merkmalen und der kindlichen Rechtschreibleistung?

Während die Beantwortung der ersten drei Fragen der Beurteilung der Angemessenheit der aktuellen HSP-Normen für die Diagnostik bei Deutschschweizer Schulkindern dient,

adressiert die letzte Fragestellung mögliche relevante Einflussfaktoren auf die Rechtschreibkompetenz der Schulkinder.

5.2 Methode

5.2.1 Design, Stichprobenrekrutierung und -merkmale

Diese Forschungsfragen wurden im Rahmen einer Querschnittstudie mit einer Gesamtstichprobe von $N = 1641$ Schulkindern der Klassenstufen 1–6 bearbeitet. Die 95 untersuchten Primarschulklassen (Regelschule) entstammen 41 Schulen aus 20 (von 35) Gemeinden aus dem gesamten deutschsprachigen Teil des Kantons Freiburg und wurden von den dortigen Logopädinnen des Regionalen Schuldienstes Freiburg rekrutiert. Voraussetzung für die Studienteilnahme war die freiwillige Zustimmung der jeweiligen Lehrperson, welche die Eltern über die Untersuchung informierte.

Tabelle 1 zeigt die Merkmale dieser angefallenen Klumpenstichprobe. Demnach konnten pro Klassenstufe 13 bis 18 Klassen bzw. zwischen 267 und 281 Kinder getestet werden. Die Gesamtstichprobe umfasst etwas mehr Knaben als Mädchen, was mit Angaben des Amtes für deutschsprachigen obligatorischen Unterricht des Kantons Freiburg korrespondiert (2013, Klassenstufen 1–6: Knaben: 51.01 %; Mädchen 48.99 %, mündliche Auskunft). Mit Blick auf die einzelnen Stufen überwiegen die Jungen mehrheitlich, wobei die Häufigkeitsverteilung keinen signifikanten Unterschied aufzeigt ($\chi^2 [5, N = 1641] = 3.69, p = .59$).

Etwas über 80 % der untersuchten Schulkinder haben Deutsch (vorwiegend Schweizerdeutsch) als Erstsprache, die restlichen eine nicht-deutsche Sprache. Im Vergleich zu den bildungsstatistischen Angaben (2013, Klassenstufen 1–6: 23.67 %, mündliche Auskunft) sind Kinder nicht-deutscher Erstsprache (18.5 %) in vorliegender Stichprobe leicht unterrepräsentiert.

Tab. 1 Merkmale der Deutschfreiburger Stichprobe

Stufen	Schüler N	Klassen N	Alter in Mt. M (SD)	Geschlecht ♀ : ♂ (in %)	Erstsprache (in %)		Sonder- maßnahme (in %)
					Deutsch	Nicht- Deutsch	
1. Klasse	272	18	90.8 (4.4)	47.4:52.6	81.3	18.8	21.0
2. Klasse	274	16	103.1 (5.1)	43.4:56.6	70.8	29.2	25.2
3. Klasse	272	17	114.5 (4.9)	47.8:52.2	84.6	15.4	28.3
4. Klasse	275	13	128.4 (5.6)	45.8:54.2	81.8	18.2	23.6
5. Klasse	281	16	141.3 (6.1)	50.9:49.1	86.8	13.2	18.9
6. Klasse	267	15	153.7 (5.9)	49.1:50.9	83.5	16.5	13.9
Alle	1641	95	122.5 (22.1)	47.4:52.6	81.5	18.5	21.8

Jedes fünfte Kind (21.8 %) erhielt zum Untersuchungszeitpunkt Logopädie (7.98 %), heilpädagogische Förderung (8.78 %) oder beide Zusatzmaßnahmen (5.06 %). Den meisten dieser Schulkinder (81.84 %) kam primär eine Unterstützung im Bereich der Schriftsprache zu (Gesamtstichprobe: 17.9 %), die restlichen waren wegen anderer Probleme (Lautsprache, Mathematik u. a.) betreut. Während der Anteil speziell geförderter Kinder von der ersten Klasse (21 %) bis zur dritten Klasse (28 %) zunimmt, geht er danach zurück und beträgt am Ende der Primarschulzeit noch 14 %.

Ein Abgleich dieser Stichprobenmerkmale mit bildungsstatistischen Angaben für Deutschfreiburg ist aus verschiedenen Gründen (Datenlücken und -struktur u. a.) nur eingeschränkt möglich. Anhand des Tätigkeitsberichts 2014 des Regionalen Schuldienstes Freiburg (Bless, 2015) lässt sich zumindest für die Logopädie eine Einschätzung vornehmen. Demnach hatten im Berichtsjahr 10.02 % aller Deutschfreiburger Schulkinder (Kindergarten, Primarschule, einschl. Kleinklassen, Orientierungsschule Sek. I, N = 8815) ein „offenes Dossier Logopädie“. Weitere 2.7 % erhielten Logopädie und eine andere Maßnahme (Schulpsychologie, Psychomotorik). Der auch für das Jahr

2013 ausgewiesene Anteil an Schülern mit logopädischer Betreuung fällt analog aus (insgesamt 12.7 %). Diese Angaben korrespondieren mit dem Befund für die Freiburger Stichprobe, wonach 13 % der Kinder logopädisch betreut waren.

Ogleich lediglich Informationen zur Logopädie und heilpädagogischen Förderung erfasst wurden, erscheint der Anteil an Schulkindern mit besonderer Maßnahme beachtlich – wenn auch nicht außergewöhnlich hoch für die Schweiz. Zwar fehlen gesamtschweizerische Zahlen zum sonderpädagogischen Bedarf bzw. Angebot in der Regelschule; gleichwohl wird in den Medien und im politischen Diskurs moniert, dass auffällig viele Primarschulkinder – je nach Expertenschätzung 20 – 50 % – niederschwellige Maßnahmen wie schulische Heilpädagogik, Logopädie oder Schulpsychologie in Anspruch nehmen. Als Erklärung für die häufige bzw. zunehmende Inanspruchnahme solcher Interventionen werden mitunter folgende Faktoren in Betracht gezogen: gesellschaftlicher Wandel, verbesserte Früherkennung und Prävention, bildungspolitische Entwicklungen bzw. verstärkte Bemühungen um die inklusive Unterrichtung und Förderung von Kindern mit Beeinträchtigungen in der Regelschule (Schmid, 2014; Birrer, 2015; Wimmerberger, Bächler &

Lüdin, 2017). Tatsächlich hat sich die Schweizer Aussonderungsquote seit 2004/2005 deutlich verringert. Wurden damals 6.2% der Schulkinder ausgesondert, waren es 2013/14 nur noch 3.7% (Bless, 2017 a, S. 221). Für 2014/15 wird eine noch etwas tiefere Quote von 3.4% ausgewiesen, die im internationalen Vergleich durchschnittlich ist (Bless, 2017 b, S. 43). Mit diesen Fortschritten einher geht eine gewisse Zunahme von niederschweligen sonderpädagogischen Maßnahmen im regelpädagogischen Kontext, um Kinder mit Entwicklungsbeeinträchtigungen oder -risiken gezielt zu unterstützen.

Bezüglich des Schriftsprachunterrichts in den untersuchten Klassen sei schließlich angemerkt, dass in Deutschfreiburg seit einigen Jahren das Lehrmittel *Die Sprachstarken* (Lindauer & Senn, 2016–2018) obligatorisch zur Anwendung kommt. Es deckt den Bereich der Schulsprache Deutsch für die Klassenstufen 2–9 ab, orientiert sich an den Lehrplänen der Deutschschweizer Kantone (Lehrplan 21) und berücksichtigt die geforderten curricularen Inhalte (Sprechen, Hören, Lesen, Schreiben, Orthografie u. a.) weitgehend.

Ein Abgleich der in diesem Abschnitt erörterten Merkmale der Freiburger Stichprobe (Alter, Geschlecht, Mehrsprachigkeit, Sondermaßnahme) mit der aktuellen HSP-Eichstichprobe ist wegen fehlender entsprechender Angaben im HSP-Testmanual (May, 2013) nicht möglich.

5.2.2 Instrumente

Die sechs Klassenstufen wurden mit den korrespondierenden HSP-Versionen (1+, 2, 3, 4–5, 5–10 B) getestet. Zusätzlich wurden anhand eines informellen Erfassungsbogens einige relevante Angaben über die einzelnen Schulkinder erhoben (Alter, Geschlecht, Erstsprache, Sondermaßnahmen u. a.).

5.2.3 Datenerhebung, Testauswertung und Datenerfassung

Die Untersuchung erfolgte am Ende des Schuljahres 2013/2014, was dem Testzeitpunkt *Ende Klasse* gemäß HSP-Testhandbuch entspricht. Die Tests wurden klassenweise durch die Logopädinnen im Beisein der Lehrperson durchgeführt. Obwohl die meisten Fachleute bereits mit der HSP vertraut waren, wurde die Testung in einer Veranstaltung speziell vorbereitet. Um die Durchführungsobjektivität zu gewährleisten, wurde den Testleiterinnen eine genaue Dokumentation zur Durchführung der HSP abgegeben.

Zur Auswertung der Tests nach den Vorgaben des HSP-Manuals kamen zwei geschulte wissenschaftliche Hilfskräfte (Studierende der Logopädie) zum Einsatz. Für sämtliche HSP-Indikatoren wurden neben den ermittelten Rohwerten auch die korrespondierenden Normwerte (Prozentränge, T-Werte) erfasst. Dabei kamen die *bundesweiten Normen* (May, 2013) zur Anwendung, was mit dem diagnostischen Vorgehen in der Praxis übereinstimmt. Den wissenschaftlichen Hilfskräften oblagen auch die digitale Datenerfassung und die systematische Überprüfung des Datensatzes. Aufbereitung und statistische Analysen der Daten (SPSS, MLwiN) erfolgten durch das Autorenteam.

5.2.4 Statistische Analysen

Die vergleichende Analyse der Rechtschreibleistungen von der Freiburger Stichprobe und der HSP-Normierungsstichprobe (Fragestellung 1) erfolgte aufgrund von rohwertbasierten Stichprobenkennwerten (M, SD, N). Diese Statistiken wurden für die Freiburger Stichprobe anhand des Datensatzes berechnet und für die Eichstichprobe dem HSP-Testmanual entnommen (May, 2013, S. 85ff.). Die statistischen Mittelwertvergleiche (t-Tests für unabhängige Stichproben, zweiseitig) für die

Indikatoren *Richtige Wörter*, *Graphemtreffer*, *alphabetische*, *orthografische*, *morphematische* und *wortübergreifende Lupenstellen* sowie all-fällige nachgeordnete Effektstärkeberechnungen (Hedges g) erfolgten mittels Online-Tools (SISA, Polytechnic University Hong Kong). Da für jedes der sechs HSP-Kriterien mehrere stufenspezifische Vergleiche – insgesamt 32 – vorgenommen wurden, erfolgte eine Alphafehler-Korrektur ($\alpha = 5\%$, $p_{\text{korrt}} = .00156$). Für die Interpretation der Effektstärken gelten die Angaben von Cohen (1988), wonach 0.2 als kleiner, 0.5 als mäßiger und 0.8 als hoher Effekt gilt.

Betreffend Fragestellung 2 wurden die *individuellen Prozentränge* verglichen, welche die Freiburger Kinder aufgrund ihrer eigenen Referenzgruppe (nachfolgend FR-PR, berechnet aufgrund des Datensatzes) oder der HSP-Eichstichprobe (nachfolgend HSP-PR, gemäß Normtabellen) erzielten. Die statistischen Vergleiche (Gesamtstichprobe) obengenannter Indikatoren erfolgten mittels Wilcoxon-Test, weil Voraussetzungen für parametrische Tests verletzt waren. Da multipel getestet wird, wurde das Alphafehlerniveau korrigiert ($\alpha = 5\%$, $p_{\text{korrt}} = .0083$). Bei signifikanten Unterschieden wurde anhand des Online-Tools Psychometrica die Effektstärke r berechnet, deren Interpretation nach Cohen (1988) erfolgt: 0.1 bis 0.3: kleiner Effekt; 0.3 bis 0.5: mittlerer Effekt; 0.5 und höher: starker Effekt.

Mit Blick auf Fragestellung 3 wurde anhand der Indikatoren *Richtige Wörter* und *Graphemtreffer* überprüft, ob und inwieweit bei Anwendung von Freiburger Normen (FR-PR) und deutschen Normen (HSP-PR) unterschiedliche Anteile an rechtschreibschwachen Kindern resultieren. Als leistungsbezogenes Kriterium für „schwache Schreiber“ wurde nach verbreiteter Konvention ein $PR < 16$ festgelegt (z. B. Schulte-Körne, 2010). Der Vergleich der je nach

Referenzstichprobe ermittelten Anteile rechtschreibschwacher Schulkinder (insgesamt und pro Klassenstufe) erfolgte deskriptiv. Zur Einschätzung der jeweiligen Übereinstimmung der Klassifikation ‚rechtschreibschwach‘ vs. ‚unauffällig‘ ($PR \geq 16$) basierend auf Freiburger Normen oder HSP-Normen wurde Cohens Kappa (k) berechnet und wie folgt interpretiert: bis 0.40: geringe Übereinstimmung, 0.40 bis 0.60: mäßige Übereinstimmung, 0.60 bis 0.75: gute Übereinstimmung, 0.75 und mehr: sehr gute Übereinstimmung.

Bezüglich Fragestellung 4 legte die hierarchisch strukturierte Stichprobe (Schulkinder in Klassen) eine *Mehr- bzw. Zweiebenenanalyse* nahe (Ebene 1: Schüler; Ebene 2: Klassen), um den Zusammenhang zwischen Klassen- und Individualfaktoren (Prädiktoren) und der kindlichen Rechtschreibleistung (abhängige Variable) schätzen zu können. Eine solche Vorgehensweise ist „in vielen Fällen die Methode der Wahl zur Analyse von Daten mit Mehrebenenstruktur“ und eine sinnvolle Alternative zu traditionellen Analysemethoden (z. B. Standard-Regressionsanalyse), die „zu Ungenauigkeiten und Fehlern führen (können)“ (Nezlek, Schröder-Abé & Schütz, 2006, S. 215). Dabei erweist sich eine Mehrebenenmodellierung als besonders nützlich, wenn ein relevanter Anteil der Gesamtvarianz der abhängigen Variable auf Merkmale (Unterschiede) der höheren bzw. zweiten Ebene zurückgeht. Dieser Anteil entspricht dem *Intraklassen-Korrelationskoeffizienten* (ICC), der anhand des sogenannten Nullmodells berechnet wird. Dieses rudimentäre (unkonditionierte) Modell ohne Prädiktoren ist der erste Schritt einer Mehrebenenanalyse. Es dient dazu, das Ausmaß der Heterogenität zwischen den Einheiten der Ebene 2 bzw. eines *clustering effects* zu bestimmen, um entscheiden zu können, ob eine mehrebenenanalytische Auswertung angezeigt ist (Langer, 2004; Nezlek et al., 2006; Nezlek, 2008). Es gibt zwar keine fixe Regel, ab wann von einem ausreichend hohen ICC gesprochen werden kann,

doch sind „(...) viele Autoritäten in Bildung und anderen Disziplinen der Meinung, dass ein ICC von 5 % eine substanzielle Evidenz für einen Clusteringeffekt (ist)“ (Glaser & Hastings, 2011, S. 880; Orig. in Engl.). Für den Freiburger Datensatz (N = 1641) resultierte aus dem berechneten Nullmodell (Ebene 1: Schüler; Ebene 2: Klassen) mit der abhängigen Variable *Freiburger T-Wert für Graphemtreffer* ein ICC von 5 % (exakt: 4.95 %) (vgl. Tab. 4 im Ergebnisteil). Dieser eher kleine, wenngleich signifikante Varianzanteil auf Klassenebene rechtfertigte somit eine weiterführende mehr-ebenenanalytische Auswertung. So wurde im Anschluss an das Nullmodell ein *Gesamtmodell* mit insgesamt sieben Prädiktoren der Ebene 1 und der Ebene 2 und derselben Zielvariable gerechnet. Als Prädiktoren auf Klassenebene dienten die folgenden Variablen, für die ein möglicher Zusammenhang mit der Rechtschreibleistung angenommen wird: 1) *Klassengröße* (um grand mean zentriert), 2) *Anteil an Schülern nicht-deutscher Erstsprache pro Klasse* (um grand mean zentriert) und 3) *Anteil an Schülern mit Sondermaßnahme zur Schriftsprache pro Klasse* (um grand mean zentriert). Als individuelle Prädiktoren (Schülerebene) gingen das *Alter in Monaten* (um grand mean zentriert), das *Geschlecht* (m vs. w, Referenzkategorie m), die *Erstsprache* (deutsch vs. nicht-deutsch, Referenzkategorie nicht-deutsch) und die Variable *Sonderpädagogische Maßnahme im Bereich Schriftsprache* (ja vs. nein, Referenzkategorie nein) in die Auswertung ein. Schließlich wurde unter Eliminierung der nicht signifikanten Prädiktoren des Gesamtmodells ein *Endmodell* mit vier verbleibenden Individualprädiktoren gerechnet. Die Berechnung der drei Modelle erfolgte im Programm MLwiN (Version 2.32). Für jedes Modell wurde zusätzlich das Akaike-Informationskriterium ($AIC = -2LL + 2 \cdot k$) ermittelt, um beurteilen zu können, ob sich die sukzessiven Modelle hinsichtlich ihrer Passung unterscheiden (Nezlek, 2008; Glaser & Hastings, 2011).

5.3 Ergebnisse

5.3.1 Vergleich der Rechtschreibleistungen von Freiburger Stichprobe und HSP-Eichstichprobe (Fragestellung 1)

Aus dem Leistungsvergleich der beiden Gruppen auf der Basis von Stichprobenkennwerten resultierten bei allen HSP-Indikatoren Unterschiede. Von 32 Vergleichen fielen 19 signifikant aus, 13 zugunsten der HSP-Stichprobe und sechs zugunsten der Freiburger Stichprobe. Abbildung 1 repräsentiert für jedes Kriterium die klassenspezifischen Effektstärken (Hedges g). Aus Darstellungsgründen werden Effektstärken für signifikante (*) wie für nicht signifikante Unterschiede ausgewiesen.

Mit Blick auf *Richtige Wörter* fallen die Freiburger Kinder ab der dritten Klasse hinter die HSP-Eichstichprobe zurück, wobei die praktische Bedeutsamkeit der Unterschiede (Effektstärke g) gering bis moderat ausfällt. Für die *Graphemtreffer* als Kernindikator der HSP ergibt sich ein uneinheitlicheres Bild. Hier resultiert nur für die dritte und die vierte Klasse ein signifikanter, geringer bis moderater Vorteil der deutschen Kinder, während die Freiburger Erstklässler bei diesem Kriterium klar besser abschneiden. Für alle anderen Stufen sind keine bedeutsamen Differenzen zugunsten der einen oder der anderen Stichprobe auszumachen.

Auffällig wie interessant sind die divergierenden Ergebnisse für die alphabetische und die orthografische Schreibstrategie. Während die Freiburger Stichprobe beim lautorientierten Verschriften auf vier von sechs Klassenstufen überzufällig besser abschneidet als die HSP-Stichprobe (geringe bis moderate Effekte), ist sie bei den orthografischen Lupenstellen ab der 2. Klasse klar schwächer als die deutschen Kinder (geringe bis mäßige Effekte). Die Befunde für die restlichen Strategien fallen wieder heterogener aus; (nur) zweimal klar zugunsten der HSP-Stichprobe (morphematische Lupenstellen) und einmal zum Vorteil der Freiburger

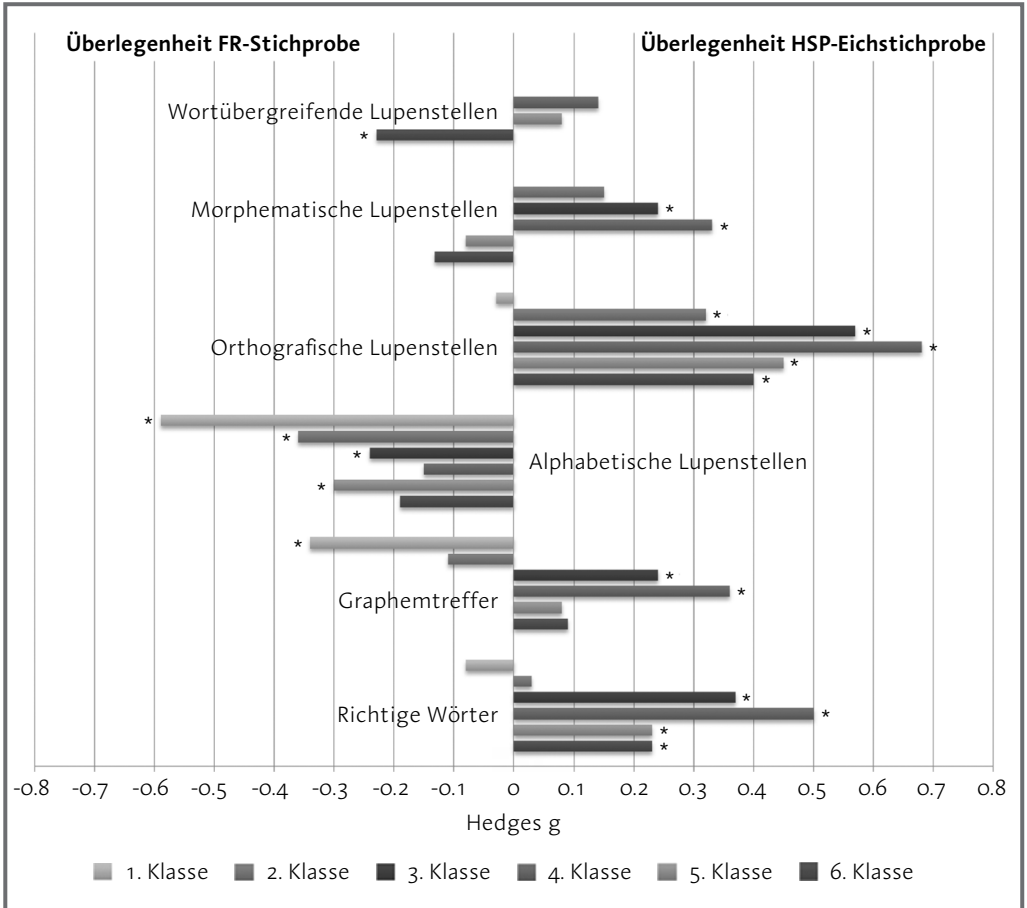


Abb. 1 Effektstärken (Hedges g) für Leistungsunterschiede zwischen Freiburger Stichprobe und HSP-Eichstichprobe bei verschiedenen HSP-Indikatoren

Tab. 2 Ergebnisse des statistischen Vergleichs (Wilcoxon-Test) von individuellen HSP-Leistungen (Prozentrang) basierend auf Freiburger Normen und deutschen Normen

HSP-Kennwerte	N	FR PR Median	HSP PR Median	Z	Sig.	ES r
Richtige Wörter	1641	53	42	31.26	<.001	.55
Graphemtreffer	1641	54	46	20.71	<.001	.36
Alphabetische Lupenstellen	1641	74	56	22.95	<.001	.40
Orthografische Lupenstellen	1641	57	36	34.06	<.001	.59
Morphematische Lupenstellen	1369	56	50	25.88	<.001	.49
Wortübergreifende Lupenstellen	823	57	49	19.04	<.001	.47

Anmerkungen: Interpretation der Effektstärke r nach Cohen (1988): 0.1 bis 0.3: kleiner Effekt; 0.3 bis 0.5: mittlerer Effekt; 0.5 und höher: starker Effekt

Kinder (wortübergreifende Lupenstellen). Soweit es zu statistisch bedeutsamen Unterschieden kommt, sind die Effektstärken gering.

5.3.2 Vergleich individueller Rechtschreibleistungen basierend auf Freiburger Normen und HSP-Normen (Fragestellung 2)

Gemäß den in Tabelle 2 ausgewiesenen Daten erzielt die untersuchte Stichprobe bei allen HSP-Indikatoren einen hoch signifikant besseren FR-PR als HSP-PR (Median). Die Freiburger Primarschulkinder werden demnach insgesamt günstiger eingestuft, wenn statt der deutschen Eichstichprobe die eigene Bezugsgruppe zur Leistungsbeurteilung herangezogen wird. Die Effektstärken sind je nach Kriterium moderat (Graphemtreffer, alphabetische, morphematische, wortübergreifende Lupenstellen) oder hoch (Richtige Wörter, orthografische Lupenstellen).

5.3.3 Vergleich der Anteile rechtschreibschwacher Schulkinder basierend auf Freiburger Normen und HSP-Normen (Fragestellung 3)

Zur Klärung, ob je nach Referenzstichprobe unterschiedliche Anteile an rechtschreibschwachen Kindern (PR < 16) identifiziert werden,

dienen die Daten in Tabelle 3. Für die Gesamtstichprobe zeigt sich, dass die beiden Indikatoren *Richtige Wörter* und *Graphemtreffer* insofern konservativ testen, als unter Verwendung sowohl des FR-PR als auch des HSP-PR etwas weniger als 16% aller Kinder als leistungsschwach eingestuft werden. Während der Anteil schwacher Rechtschreiber basierend auf den Originalnormen für *Richtige Wörter* leicht höher ausfällt als mit Freiburger Normen, ergeben sich für die *Graphemtreffer* identische Anteile an als rechtschreibschwach klassifizierten Schulkindern. Mit Kappa $\kappa = .70$ ($p < .001$, $N = 1641$) für *Graphemtreffer* fällt die Übereinstimmung der Klassifikation (schwache vs. unauffällige Schreiber) anhand von FR-PR bzw. HSP-PR gut, wenn auch nicht optimal aus. Für das Kriterium *Richtige Wörter* resultiert ein analoger Befund ($\kappa = 0.68$, $p < .001$, $N = 1641$).

Betrachtet man hingegen die diagnostisch primär relevanten Ergebnisse für die einzelnen Klassenstufen, so fallen die Anteile schwacher Rechtschreiber bei Anwendung der deutschen Normen weit heterogener aus als mit Freiburger Normen, was für *Richtige Wörter* (HSP 2.9–28.4% vs. FR 8.5–14.2%) und für *Graphemtreffer* (HSP 2.6–26.2% vs. FR 10.3–15.3%) in ähnlicher Weise gilt.

Tab. 3 Stufenspezifische und totale Anteile an rechtschreibschwachen Freiburger Kindern basierend auf Freiburger Normen und deutschen Normen (Prozentrang) für *Richtige Wörter* und *Graphemtreffer*

	FR-PR Richtige Wörter	HSP-PR Richtige Wörter	FR-PR Graphemtreffer	HSP-PR Graphemtreffer
1. Klasse (N = 272)	8.5 %	2.9 %	10.3 %	2.6 %
2. Klasse (N = 274)	13.1 %	8.0 %	12.0 %	4.0 %
3. Klasse (N = 272)	13.2 %	25.4 %	14.3 %	21.3 %
4. Klasse (N = 275)	14.2 %	28.4 %	15.3 %	26.2 %
5. Klasse (N = 281)	12.8 %	8.9 %	14.6 %	9.3 %
6. Klasse (N = 267)	13.1 %	18.4 %	15.0 %	18.4 %
Alle (N = 1641)	12.5 %	15.3 %	13.6 %	13.6 %

5.3.4 Zusammenhänge zwischen klassenstrukturellen und individuellen Variablen und der Rechtschreibleistung (Fragestellung 4)

Die Ergebnisse der Mehrebenenanalyse sind in Tabelle 4 ausgewiesen. Wie das Nullmodell zeigt, erzielte die Freiburger Stichprobe (N = 1641) einen durchschnittlichen T-Wert von 49.98 (Intercept), der minimal vom statistisch erwarteten Wert 50 abweicht. Die Varianz auf Klassenebene beträgt 4.94 (SE = 1.54, $p < .01$), diejenige auf Schülerebene 94.73 (SE = 3.40, $p < .001$) und die totale Varianz 99.67. Da der Intraklassen-Korrelationskoeffizient (ICC) des Nullmodells 5% beträgt und die Varianz auf Klassenebene signifikant ist, wurde anschließend ein Gesamtmodell mit allen sieben E1- und E2-Prädiktoren und derselben Zielvariable gerechnet. Dieses Modell umfasst wegen einiger fehlender Daten (Alter) lediglich 1611 Fälle. Wie die Ergebnisse zeigen, resultierte für die auf Ebene 2

angesiedelten Variablen *Klassengröße*, *Anteil Fremdsprachige pro Klasse* und *Anteil sonderpädagogisch Betreute im Bereich Schriftsprache pro Klasse* kein signifikanter Effekt auf das Rechtschreiben. Im Unterschied dazu erwiesen sich alle individuellen Prädiktoren – *Alter*, *Geschlecht*, *Erstsprache*, *Maßnahme im Bereich Schriftsprache* – als signifikant. Wie sich anhand des AIC zeigen lässt, ist das Gesamtmodell (11550) dem Nullmodell (12191) hoch signifikant überlegen (Differenz 641, $df = 7$, $p < .001$). Nach Eliminierung der im Gesamtmodell nicht signifikanten Variablen verblieben alle vier Prädiktoren der Schülerebene auch im Endmodell (hoch) signifikant, dessen Betakoeffizienten und Standardfehler nicht nennenswert von denjenigen des Gesamtmodells differieren.

Gemäß dem Endmodell erzielen ältere Schulkinder innerhalb der einzelnen Klassen (pro Monat) einen um 0.03 Punkte tieferen T-Wert

Tab. 4 Ergebnisse der Mehrebenenanalyse zur Vorhersage der HSP-Rechtschreibleistung (Grammfehler) durch Klassen- und Individualfaktoren

Variablen/Kennwerte	Nullmodell (ohne Prädiktoren) N = 1641 B (SE)	Gesamtmodell (7 Prädiktoren) N = 1611 B (SE)	Endmodell (4 Prädiktoren) N = 1611 B (SE)
Intercept	49.978*** (0.335)	51.738*** (0.65)	51.850*** (0.635)
E1: Alter in Mt.		-0.030* (0.015)	-0.031* (0.015)
E1: Geschlecht (Ref.: Mädchen)		-2.747*** (0.429)	-2.760*** (0.429)
E1: Erstsprache (Ref.: nicht-deutsch)		2.099*** (0.605)	1.895** (0.579)
E1: Maßnahme Schriftsprache (Ref.: keine)		-11.072*** (0.574)	-11.018*** (0.565)
E2: Klassengröße		0.061 (0.070)	
E2: Fremdsprachige pro Klasse (%)		0.024 (0.021)	
E2: Sonderpäd. Betreute pro Klasse (%)		0.027 (0.034)	
Varianzen:			
Klassenebene (E2)	4.935** (1.537)	5.933*** (1.499)	6.162** (1.533)
Schülerebene (E1)	94.732*** (3.404)	71.438*** (2.592)	71.474** (2.593)
ICC	5.0%	7.7%	7.9%
AIC	12191	11550	11547

Anmerkungen: * $p < .05$; ** $p < .01$; *** $p < .001$

B: nicht standardisierter Regressionskoeffizient; SE: Standardfehler; ICC: Intraklassen-Korrelationskoeffizient; AIC: Akaike Informationskriterium

für Graphemtreffer als jüngere Kameraden. Jungen ($M = 48.38$; $SD = 10.55$) schneiden um 2.76 T-Wert-Punkte schwächer ab als Mädchen ($M = 51.80$; $SD = 8.98$), und Schüler und Schülerinnen mit deutscher Erstsprache ($M = 50.43$; $SD = 9.89$) erreichen rund 2 T-Wert-Punkte mehr als Kinder nicht-deutscher Sprache ($M = 48.13$; $SD = 10.19$). Am deutlichsten – nämlich um 11 T-Wert-Punkte – unterscheiden sich indes Schulkinder mit sonderpädagogischer Schriftsprachförderung ($M = 40.56$; $SD = 11.33$) von Schulkindern ohne solche Maßnahme ($M = 52.05$; $SD = 8.37$).

Nicht überraschend ist auch das Endmodell dem Nullmodell überlegen. So betragen das AIC des Nullmodells 12191 und das AIC des Modells mit den vier signifikanten Prädiktoren 11547. Die Differenz von 644 ist hoch signifikant ($df = 4$, $p < .001$). Demgegenüber erweist sich die Differenz von 3 der AIC-Werte des Endmodells (11547) und des Gesamtmodells (11550) als nicht bedeutsam ($df = 3$, $p > .10$). Die analysierten Daten passen folglich unwesentlich besser zum sparsameren Endmodell als zum Gesamtmodell.

6 Diskussion

Das Hauptanliegen dieser Studie war es, die diagnostische Eignung der Anwendung von deutschen Rechtschreibnormen auf Freiburger Primarschulkinder anhand der HSP exemplarisch zu überprüfen. In nachgeordneten Datenanalysen wurde weiter die prädiktive Bedeutung einiger individueller und klassenstruktureller Faktoren für die kindliche Rechtschreibleistung exploriert. Nachfolgend werden zuerst die Befunde zur Hauptfragestellung und sodann die Ergebnisse hinsichtlich der Zusatzfragestellung erörtert. Hinweise auf Einschränkungen der Studie, die im Hinblick auf die Reichweite der Forschungsergebnisse zu berücksichtigen sind, runden den Diskussions teil ab.

6.1 Eignung der HSP-Normen

Um zu beurteilen, wie gut die deutschen Normen auf die Freiburger Schulkinder passen, werden die Befunde zu deren Rechtschreibleistungen unter folgenden drei Gesichtspunkten resümiert und diskutiert:

1. Bezüglich der Frage, ob die Freiburger Stichprobe und die HSP-Eichstichprobe hinsichtlich quantitativer und qualitativer Rechtschreibindikatoren signifikant differieren, dokumentieren die Analyseergebnisse bei allen HSP-Kriterien mehrere signifikante Leistungsunterschiede, deren Richtung und Ausmaß (geringe und moderate Effektstärken) je nach Indikator und Klassenstufe variieren können. Während die deutsche Stichprobe insgesamt mehr Leistungsvorteile (13) verzeichnet als die der Freiburger Schulkinder (6) und dieser insbesondere im orthografischen Schreiben fast durchwegs klar überlegen ist, schneidet umgekehrt die Deutschschweizer Stichprobe bei der alphabetischen Strategie mehrheitlich deutlich besser ab. Dieser beobachtete Leistungsvorsprung der Freiburger Kinder korrespondiert mit früheren Befunden von May (2013) oder von Schmidlin (2003), wonach Schülerinnen und Schüler aus Gebieten mit starkem Dialekteinfluss relative Stärken im lautorientierten Schreiben verzeichnen. Dabei ist allerdings zu betonen, dass die präsentierten Analysen und Ergebnisse keine Aussagen über den tatsächlichen Einfluss von dialektalen bzw. linguistischen Faktoren auf die Rechtschreibkompetenzen von Freiburger (bzw. deutschen) Schulkindern erlauben. Hierzu wären zumindest weiterführende Auswertungen auf Itemebene erforderlich, die den Rahmen des Beitrags gesprengt hätten – vor allem aber direkte ländervergleichende Studien mit geeignetem Untersuchungsmaterial (vgl. Ausblick). Unklar bleibt auch, inwieweit die festgestellte alphabetische Überlegenheit der Freiburger Kinder zusätzlich durch unterrichtliche Faktoren bzw. Unterschiede zwischen den Ländern begünstigt wird.

Der Befund, dass die Freiburger Kinder demgegenüber beim Verschriften von orthografischen (und teilweise auch morphematischen) Lupenstellen insgesamt schwächer sind als die deutschen Schulkinder, weist in dieselbe Richtung wie Erfahrungen bei der Schweizer Normierung des SLRT II (Moll & Landerl, 2014). Über die Gründe für die (relative) orthografische Schwäche der Freiburger Stichprobe im Vergleich zur HSP-Stichprobe lassen sich nur Vermutungen anstellen. Da der Dialekteinfluss per se im Hinblick auf das orthografisch korrekte Schreiben wenig plausibel erscheint, sind erklärend eher curriculare und methodische Unterrichtsunterschiede zwischen der Schweiz und Deutschland in Betracht zu ziehen, über welche diese Studie indes keine Auskunft zu geben vermag.

2. In einem weiteren Schritt wurde analysiert, ob die Bewertungen der individuellen HSP-Testleistungen der Freiburger Schulkinder in Abhängigkeit von der Referenzstichprobe differieren. Bei allen HSP-Indikatoren erzielt die Gesamtstichprobe einen deutlich besseren Prozentrang basierend auf Freiburger Normen im Vergleich zu den deutschen Originalnormen, mit denen die Freiburger Kinder insgesamt strenger beurteilt werden. Die ermittelten Effektstärken verweisen je nach Kennwert auf eine mäßige oder hohe praktische Bedeutsamkeit der statistischen Unterschiede.

3. Zur Frage, ob sich die Anteile an Freiburger Kindern mit schwacher Rechtschreibleistung (definiert als $PR < 16$) je nach Bezugsgruppenormen unterscheiden, zeigen die Ergebnisse das Folgende: Anhand der deutschen Normen werden 13.6% (*Richtige Wörter*) bzw. 15.3% (*Graphemtreffer*) aller Kinder (Gesamtstichprobe) als rechtschreibschwach klassifiziert, was auf eine leicht konservative Tendenz der Originalnormen hinweist. Ebenfalls etwas konservativ fallen die Ergebnisse für die Freiburger Normen aus, mit denen 12.5% (*Richtige Wörter*) bzw. 13.6% (*Graphemtreffer*) aller Schul-

kinder als leistungsschwach eingestuft werden. Diagnostisch weit wichtiger sind die Befunde für die einzelnen Klassenstufen, wonach die HSP-Normen sehr unausgeglichene – zu tiefe wie zu hohe – Anteile an rechtschreibschwachen Schülerinnen und Schülern (*Richtige Wörter* 2.9 – 28.4%; *Graphemtreffer* 2.6 – 26.2%) generieren und somit weniger befriedigend ausfallen als die Ergebnisse für die Freiburger Normen (*Richtige Wörter* 8.5 – 14.2%; *Graphemtreffer* 10.3 – 15.3%).

In diagnostischer Hinsicht primär bedeutsam sind zusammenfassend die folgenden Befunde: a) Freiburger Schulkinder werden anhand der HSP-Normen im Allgemeinen strenger beurteilt bzw. etwas unterschätzt im Vergleich zu den Leistungen (Normen) der eigenen Bezugsgruppe. b) Die deutschen Normen für diagnostische Kernkriterien generieren in der untersuchten Stichprobe verzerrte, d. h. je nach Klassenstufe stark variierende Anteile an rechtschreibschwachen Kindern. Diese Bilanz liefert insgesamt Evidenz, dass die in dieser Studie analysierten HSP-Normen nicht bzw. nur bedingt zu Freiburger Primarschulkindern passen, was ihre Eignung für die hiesige Rechtschreibdiagnostik in einem kritischen Licht erscheinen lässt.

6.2 Prädiktoren der Rechtschreibleistung

Nach den Ergebnissen der Mehrebenenanalyse stehen strukturelle Variablen (Klassengröße, Anteile fremdsprachiger bzw. sonderpädagogisch betreuter Schulkinder pro Klasse) in keinem signifikanten Zusammenhang mit der kindlichen Rechtschreibleistung, ein Befund, der u. a. mit Ergebnissen etwa von Treutlein et al. (2012) konvergiert. Demgegenüber erweisen sich alle berücksichtigten Individualfaktoren als statistisch bedeutsame, wenn auch unterschiedlich gewichtige Prädiktoren. In Übereinstimmung mit dem Forschungsstand belegen die Daten, dass Knaben im Recht-

schreiben insgesamt schwächer sind als Mädchen (May, 2013; Philipp, 2017). Weiter weisen Kinder mit nicht-deutscher Erstsprache im Allgemeinen geringere Rechtschreibleistungen auf als Kinder mit deutscher Erstsprache, was mit Angaben in der Literatur übereinstimmt (May, 2013; Marx, 2017). Schließlich unterscheiden sich die Rechtschreibleistungen von Kindern mit sonderpädagogischer Maßnahme zum Bereich der Schriftsprache erheblich – d. h. durchschnittlich um 11 T-Wert-Punkte – von den Leistungen von Kindern ohne eine solche Förderung. Dieser Sachverhalt lässt sich positiverweise dahingehend interpretieren, dass in der untersuchten Stichprobe vorrangig Kinder mit schwächeren Rechtschreibleistungen zusätzliche Förderangebote erhalten.

6.3 Einschränkungen

In Übereinstimmung mit früheren Beobachtungen (zum SLRT II) verweist die präsentierte Studie aus dem Kanton Freiburg auf mögliche Probleme der Anwendung von deutschen Testnormen zur Beurteilung von Schriftsprachleistungen von Deutschschweizer Schulkindern. Sie lässt im Speziellen gewisse Vorbehalte gegenüber der diagnostischen Eignung der neuen HSP-Normen in der Schweiz aufkommen. Dabei ist zu betonen, dass unsere Forschungsergebnisse keine abschließende und generalisierbare Beurteilung der Angemessenheit der Nutzung der HSP in der Deutschschweiz erlauben, was methodischen Einschränkungen der Studie geschuldet ist.

So ist zunächst darauf hinzuweisen, dass anhand der Freiburger Stichprobe lediglich bundesweite HSP-Normen für die Klassenstufen 1–6 überprüft wurden, und dies ausschließlich für den Testzeitpunkt „Ende Schuljahr“. Die Untersuchung erlaubt somit keine Aussagen zur Tauglichkeit der HSP-Normen für die Klassenstufen 7–10 sowie für den zusätzlichen Testzeitpunkt „Mitte Schuljahr“.

Darüber hinaus basieren die präsentierten Ergebnisse und Schlussfolgerungen auf Daten einer relativ umfangreichen, angefallenen Stichprobe, die nicht als (strikt) repräsentativ gelten kann. Da es sich bei der HSP-Eichstichprobe ebenfalls um keine repräsentative Stichprobe handeln soll (Lenhard & Lenhard, 2017), sind die in diesem Beitrag referierten Befunde und Aussagen zu den Rechtschreibleistungen von „Freiburger Schulkindern“ und „deutschen Schulkindern“ mit gebührender Vorsicht zu interpretieren bzw. nicht verallgemeinerbar. Selbstredend lassen sich die Ergebnisse des Freiburger HSP-Projekts nicht ungeprüft auf Schulkinder aus anderen Deutschschweizer Kantonen übertragen.

7 Forschungsausblick

In Anbetracht der beschränkten Reichweite der erörterten Studienbefunde ist es angezeigt, die in diesem Beitrag behandelten Fragen in weiterführenden Untersuchungen empirisch zu bearbeiten. So wäre es sinnvoll, dass die diagnostische Eignung der aktuellen HSP-Normen auch noch in anderen Kantonen überprüft wird, in denen die HSP routinemäßig zur Anwendung gelangt. Durch einen Vergleich der zusätzlich gewonnenen Daten mit den hier referierten Befunden ließe sich klären, ob in anderen Deutschschweizer Regionen analoge oder abweichende Befundmuster resultieren wie im Freiburger Projekt. Sollten sich die HSP-Normen in der weiteren Forschung in ähnlicher Weise als kritisch bzw. nur bedingt passend erweisen für die Rechtschreibdiagnostik in der Deutschschweiz, wäre zu überlegen, ob eine umfassende Schweizer Normierung der HSP in Betracht gezogen werden sollte.

Über die Überprüfung der Eignung von Normen von deutschen Schriftsprachttests für die LRS-Diagnostik in der Schweiz hinaus wäre es schließlich interessant(er) und wissenschaftlich innovativ, den Erwerb von Rechtschreibkompe-

tenzen von Deutschschweizer, deutschen und österreichischen Schulkindern in ländervergleichenden (Längsschnitt-)Studien systematischer und eingehender zu untersuchen, und zwar auch unter Berücksichtigung des jeweiligen schulisch-unterrichtlichen und sozialen Kontextes, in dem die Schülerinnen und Schüler sukzessive Rechtschreibfertigkeiten erwerben.

Literatur

- Bickel, H. (2001). Schweizerhochdeutsch: kein minderwertiges Hochdeutsch! Das Deutsche als plurizentrische Sprache aus Schweizer Sicht. *Babylonia*, 2(1), 19–22.
- Birrer, R. (2015, 7. Mai). Häufig klappt die Integration in der Oberstufe nicht mehr. *Tages Anzeiger (online)*. Abgerufen am 13.9.2017 von <https://www.tagesanzeiger.ch/schweiz/standard/Haeufig-klappt-die-Integration-in-der-Oberstufe-nicht-mehr/story/13761183>
- Bless, G. (2015). *Regionaler Schuldienst Freiburg: Tätigkeitsbericht 2014*. Freiburg/CH: Heilpädagogisches Institut der Universität Freiburg.
- Bless, G. (2017a). Integrationsforschung: Entwurf einer Wissenskarte. *Zeitschrift für Heilpädagogik*, 68(5), 216–227.
- Bless, G. (2017b). Irreguläre Schullaufbahnen im Rahmen der obligatorischen Schule und Forschungsergebnisse zur Klassenwiederholung. In M. P. Neuenschwander & C. Nägele (Hrsg.), *Bildungsverläufe von der Einschulung bis in den ersten Arbeitsmarkt*, 39–57. Wiesbaden: Springer.
- Brügelmann, H. (2015). Wie wirken verschiedene Methoden des Rechtschreibunterrichts? In E. Brinkmann (Hrsg.), *Rechtschreiben in der Diskussion – Schriftspracherwerb und Rechtschreibunterricht*, 185–193. Frankfurt: Grundschulverband.
- Christen, H., Glaser, E. & Friedli, M. (Hrsg.) (2013). *Kleiner Sprachatlas der Schweiz*. 5., überarbeitete und erweiterte Auflage. Frauenfeld/Stuttgart/Wien: Huber.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum Associates.
- Deimel, V. W. (2002). *Testverfahren zur Diagnostik der Lese-Rechtschreibstörung – eine Übersicht*. Abgerufen am 22.4.2016 von https://www.kjp.med.uni-muenchen.de/download/Testverfahren_zurDiagnostik1.pdf
- Deimel, W., Ziegler, A., & Schulte-Körne, G. (2005). *Modell Schriftsprach-Moderatoren (MSM). Zwischenbericht der wissenschaftlichen Begleitung nach zwei Jahren*. Abgerufen am 22.4.2016 von <http://www.kjp.med.uni-muenchen.de/download/MSM-Zwischenbericht.pdf>
- DGKJP/Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie (2015). *Diagnostik und Behandlung von Kindern und Jugendlichen mit Lese- und/oder Rechtschreibstörung. Evidenz- und konsensbasierte Leitlinie*. Abgerufen am 10.9.2016 von http://www.awmf.org/uploads/tx_szleitlinien/028-044l_S3_Lese-Rechtschreibstorgen_Kinder_Jugendliche_2015-06.pdf
- Endlich, D., Dummert, F., Schneider, W. & Schwenck, C. (2014). Verhaltensprobleme bei Kindern mit umschriebener und kombinierter schulischer Minderleistung. *Kindheit und Entwicklung*, 23(1), 61–69. <https://doi.org/10.1026/0942-5403/a000128>
- Ennemoser, M., Marx, P., Weber, J. & Schneider, W. (2012). Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(2), 53–67. <https://doi.org/10.1026/0049-8637/a000057>
- Galuschka, K., Ise, E., Krick, K. & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. *PLoS One*, 9(2). <https://doi.org/10.1371/journal.pone.0089900>
- Glaser, D. & Hastings, R. H. (2011). An introduction to multilevel modeling for anesthesiologists. *Anesthesia & Analgesia*, 113(4), 887–887. <https://doi.org/10.1213/ANE.0b013e3182198a01>
- Häcki Buhofer, A. & Burger, H. (1998). *Wie Deutschschweizer Kinder Hochdeutsch lernen. Der ungesteuerte Erwerb des gesprochenen Hochdeutschen durch Deutschschweizer Kinder zwischen sechs und acht Jahren*. Stuttgart: Franz Steiner Verlag.
- Hartmann, E. (2017). Kinder und Jugendliche mit spezifischer Sprachentwicklungsstörung. In M. Philipp (Hrsg.), *Handbuch Schriftspracherwerb und weiterführendes Lesen*, 361–375. Weinheim: Beltz.

- Hattie, J. A. (2014). *Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“*. 3., erweiterte Auflage. Hohengehren: Schneider.
- Hove, I. (2002). *Die Aussprache der Standardsprache in der deutschen Schweiz*. Tübingen: M. Niemeyer. <https://doi.org/10.1515/9783110919936>
- Hove, I. (2017). Konfrontiert mit Schweizerdeutsch, Schweizerhochdeutsch und bundesdeutschem Hochdeutsch. *SAL-Bulletin*, 165 (3), 5–13.
- Ise, E., Engel, R. & Schulte-Körne, G. (2012). Was hilft bei der Lese-Rechtschreibstörung? Ergebnisse einer Metaanalyse zur Wirksamkeit deutschsprachiger Förderansätze. *Kindheit und Entwicklung*, 21 (2), 122–136. <https://doi.org/10.1026/0942-5403/a000077>
- Kamhi, A. G. & Catts, H. W. (Eds.) (2012). *Language and Reading Disabilities*. 3rd ed. Boston: Pearson.
- Klicpera, C. & Gasteiger-Klicpera, B. (1993). *Lesen und Schreiben – Entwicklung und Schwierigkeiten*. Bern: Hans Huber.
- Klicpera, C., Schabmann, A. & Gasteiger-Klicpera, B. (2013). *Legasthenie – LRS: Modelle, Diagnose, Therapie und Förderung*. 4., aktualisierte Auflage. München: Reinhardt.
- Kohn, J., Wyschkon, A., Ballaschk, K., Ihle, W. & Esser, G. (2013). Verlauf von Umschriebenen Entwicklungsstörungen: Eine 30-Monats-Follow-up-Studie. *Lernen und Lernstörungen*, 2 (1), 77–89. <https://doi.org/10.1024/2235-0977/a000032>
- Kohn, J., Wyschkon, A. & Esser, G. (2013). Psychische Auffälligkeiten bei Umschriebenen Entwicklungsstörungen: Gibt es Unterschiede zwischen Lese-Rechtschreib- und Rechenstörungen? *Lernen und Lernstörungen*, 2 (1), 7–20. <https://doi.org/10.1024/2235-0977/a000027>
- Kowalski, K. & Voss, A. (2009). Die IGLU-Ergänzungsstudie 2006 zur Rechtschreibkompetenz von Viertklässlern. In R. Valtin & B. Hofmann (Hrsg.), *Kompetenzmodelle der Orthographie. Empirische Befunde und förderdiagnostische Möglichkeiten*, 26–38. Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- Landert, K. (2007). *Hochdeutsch im Kindergarten? Eine empirische Studie zum frühen Hochdeutscherwerb in der Deutschschweiz*. Bern/Landert: Peter Lang.
- Langer, W. (2004). *Mehrebenenanalyse. Eine Einführung für Forschung und Praxis*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lenhard, A. & Lenhard, W. (2017). Diagnoseverfahren zur Erfassung schriftsprachlicher Leistungen. In M. Philipp (Hrsg.), *Handbuch Schriftspracherwerb und weiterführendes Lesen*, 174–198. Weinheim: Beltz.
- Lindauer, T. & Schmellentin, C. (2017). Rechtschreibung: System, Kompetenzen und Erwerb. In M. Philipp (Hrsg.), *Handbuch Schriftspracherwerb und weiterführendes Lesen*, 20–35. Weinheim: Beltz.
- Lindauer, T. & Senn, W. (Hrsg.) (2016–2018). *Die Sprachstarken 2–9. Sprachlehrmittel für die Schweizer Volksschulen*. Zug: Klett & Balmer.
- Martschinke, S. & Kammermeyer, G. (2003). Jedes Kind ist anders. Jede Klasse ist anders. *Zeitschrift für Erziehungswissenschaft*, 6 (2), 257–275.
- Marx, A. (2017). Kinder und Jugendliche mit Zuwanderungshintergrund. In M. Philipp (Hrsg.), *Handbuch Schriftspracherwerb und weiterführendes Lesen*, 332–346. Weinheim: Beltz.
- May, P. (2009). Orthographie. In W. Bos, M. Bensen & C. Gröhlich (Hrsg.), *KESS 7. Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7*, 68–79. Münster: Waxmann.
- May, P. (2013). *HSP 1–10. Hamburger Schreibprobe zur Erfassung der grundlegenden Rechtschreibkompetenzen. Manual/Handbuch: Diagnose orthografischer Kompetenz*. Stuttgart: Klett.
- May, P., Malitzky, V. & Vieluf, U. (2001). Rechtschreibtests im Vergleich. Wie stellt man deren Güte fest und wie besser nicht? *Psychologie in Erziehung und Unterricht*, 48 (2), 146–152.
- Meyer, S. (2000). *Die Hamburger Schreibprobe im Kanton Solothurn. Bericht über eine Normierung*. Solothurn: Kantonale Drucksachenverwaltung.
- Moll, K. & Landerl, K. (2014). *SLRT II. Lese- und Rechtschreibtest*. 2., korrigierte und aktualisierte Auflage. Bern: Huber.
- Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass*, 2 (2), 842–860. <https://doi.org/10.1111/j.1751-9004.2007.00059.x>
- Nezlek, J., Schröder-Abé, M. & Schütz, A. (2006). Mehrebenenanalysen in der psychologischen Forschung. *Psychologische Rundschau*, 57 (4),

- 213–223. <https://doi.org/10.1026/0033-3042.57.4.213>
- Philipp, M. (2017). Geschlechterspezifika im Lesen und Schreiben. In M. Philipp (Hrsg.), *Handbuch Schriftspracherwerb und weiterführendes Lesen*, 318–331. Weinheim: Beltz.
- Polytechnic University Hong Kong. *Effect size calculators*. Abgerufen am 03. 10. 2015 von <http://www.polyu.edu.hk/mm/efficientsizefaqs/calculator/calculator.html>
- Psychometrica (2016). *Berechnung von Effektstärken*. Abgerufen am 17. 10. 2017 von <https://www.psychometrica.de/effektstaerke.html>
- Schmid, S. (2014, 10. November). Von Therapiewahn kann keine Rede sein. *Tages Anzeiger (online)*. Abgerufen am 13.9.2017 von <https://www.tagesanzeiger.ch/zuerich/stadt/Von-Therapiewahn-kann-keine-Rede-sein/story/22278534>
- Schmidlin, R. (2003). Zum Orthographieerwerb bei Deutschschweizer Kindern. In A. Häcki-Buhofer (Hrsg.), *Spracherwerb und Lebensalter*, 209–225. Tübingen & Basel: Francke.
- Schulte Körne, G. (2010). Diagnostik und Therapie der Lese-Rechtschreib-Störung. *Deutsches Ärzteblatt International*, 107(41), 718–727. <https://doi.org/10.3238/larztebl.2010.0718>
- SISA (o.J.). *Simple Interactive Statistical Analysis*. Abgerufen am 11. 10. 2015 von <http://www.quantitativeskills.com/sisa/>
- Sturm, A. (2017). Förderung hierarchieniedriger Schreibprozesse. In M. Philipp (Hrsg.), *Handbuch Schriftspracherwerb und weiterführendes Lesen*, 266–284. Weinheim: Beltz.
- Tacke, G., Völker, R. & Lohmüller, R. (2001). Die Hamburger Schreibprobe: Probleme mit einem neuen Rechtschreibtest. *Psychologie in Erziehung und Unterricht*, 48(2), 135–145.
- Treutlein, A., Roos, J. & Schöler, H. (2012). Lese- und Rechtschreibleistungen und ihre Abhängigkeit von individuellen Deutschkenntnissen, sprachlicher Zusammensetzung der Klasse und Klassengröße. *Psychologie in Erziehung und Unterricht*, 59(2), 122–131. <https://doi.org/10.2378/peu2012.art07d>
- Wimmerberger, A., Bächler, A. & Lüdin, C. (2017, 26. Oktober). Föderalismus und heilpädagogische Leistungen. *Neue Zürcher Zeitung (online)*. Abgerufen am 13.11.2017 unter <https://www.nzz.ch/meinung/foederalismus-und-heilpaedagogische-leistungen-ld.1324135>

Anschrift der Autor/innen

Prof. Dr. phil. Erich Hartmann
Dr. phil. Julia Winkes
Dr. phil. Felix Studer
 Universität Freiburg/Schweiz
 Departement für Sonderpädagogik
 Petrus-Kanisius-Gasse 21
 CH-1700 Freiburg
 E-Mail: erich.hartmann@unifr.ch