



Bildungssprachliche Kompetenzen (BiSpra): Anforderungen, Sprachverarbeitung und Diagnostik (Verbundvorhaben)

Forschungsvorhaben im Rahmen des Programms des Bundesministeriums für Bildung und Forschung zur Förderung der empirischen Bildungsforschung im Bereich Sprachdiagnostik/Sprachförderung FiSS (2. Förderphase)

Förderkennzeichen: 01GJ1209A/B

Schlussbericht 2016¹

Projektleitung²:

Prof. Dr. Sabine Weinert
(Otto-Friedrich-Universität Bamberg)
&

Prof. Dr. Petra Stanat
(Humboldt-Universität zu Berlin in Kooperation mit dem Institut zur Qualitätsentwicklung im Bildungswesen)

mit zeitweiliger Unterstützung durch
Prof. Dr. Angelika Redder
(Universität Hamburg)

Autorinnen des vorliegenden Schlussberichts:

Sabine Weinert, Petra Stanat, Birgit Heppt, Elisabeth Schuth³

¹ Der Aufbau des vorliegenden Berichts orientiert sich an den Vorgaben für Schlussberichte des BMBF

² Projektmitarbeiterinnen in der 2. Förderphase: Laura Babeliowsky, MSc. Psych.; Dr. Birgit Heppt; Dr. Judith Köhne; Elisabeth Schuth, MSc. Psych.

³ Wir danken Laura Babeliowsky für ihre Unterstützung.

Inhaltsverzeichnis

I Kurze Darstellung	3
1 Projektziele / Aufgabenstellung	4
2 Planung und Ablauf des Vorhabens sowie Voraussetzungen, unter denen das Vorhaben durchgeführt wurde	4
2.1 Entwicklung und Weiterentwicklung von Aufgaben / Instrumenten zur Erfassung bildungssprachlicher Kompetenzen und Pilotierung.....	4
2.2 Rekrutierung von Schulen und Horten für die Teilnahme an den Erhebungen und Genehmigungsverfahren	11
2.3 Zentrale Erhebungen während der Projektlaufzeit von BiSpra II	12
2.4 Auswahl und Schulung von Testleiterinnen und Testleitern	15
2.5 Durchführung der Validierungsstudie und Erhebungsdesign	16
2.6 Kodierung, Eingabe, Aufbereitung und Analyse der Daten.....	19
2.7 Kommunikation der Projektergebnisse sowie Manuskripterstellung.....	19
3 Wissenschaftlicher Stand, an den angeknüpft wurde.....	20
3.1 Angabe bekannter Konstruktionen, Verfahren und Schutzrechte, die für die Durchführung des Vorhabens benutzt wurden.....	20
3.2 Angabe der verwendeten Fachliteratur sowie der benutzten Informations- und Dokumentationsdienste	20
4 Zusammenarbeit mit anderen Stellen.....	22
II Eingehende Darstellung.....	23
1 Ergebnisse	23
1.1 Ergebnisse der Pilotierungsstudien	23
1.2 Ergebnisse der Validierungsstudie	26
1.3 Weiterführende Analysen im Rahmen von BiSpra.....	32
2 Voraussichtlicher Nutzen, insbesondere der Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans	34
3 Darstellung der wichtigsten Positionen des zahlenmäßigen Nachweises und der Notwendigkeit und Angemessenheit der geleisteten Arbeiten	35
4 Während der Durchführung des Vorhabens dem ZE bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen.....	35
5 Erfolgte und geplante Veröffentlichungen der Ergebnisse	36
6 Im Bericht zitierte Literatur	41

I Kurze Darstellung

Großangelegte Schulleistungsstudien wie PISA, IGLU und TIMSS weisen wiederholt auf soziale und zuwanderungsbezogene Disparitäten im Bildungserfolg hin (z. B. OECD, 2006, 2010; Schwippert, Wendt & Tarelli, 2012). Es wird angenommen, dass hierbei eingeschränkte Kenntnisse der Unterrichtssprache, der sogenannten Bildungssprache, eine Rolle spielen. Trotz der Bedeutung, die bildungssprachlichen Kompetenzen zugeschrieben wird, und verschiedenen vorliegenden Konzeptualisierungen von Bildungssprache fehlt bislang eine klare Beschreibung derjenigen Aspekte von Sprache, die unter dem Begriff Bildungssprache zusammengefasst werden. Zugleich existieren – insbesondere im deutschen Sprachraum – noch keine validen Instrumente zur Erfassung von Kompetenzen im Umgang mit diesen sprachlichen Anforderungen. Vor diesem Hintergrund verfolgte das Projekt „Bildungssprachliche Kompetenzen: Anforderungen, Sprachverarbeitung und Diagnostik“ (BiSpra) vor allem zwei Ziele: (1) Zum einen sollten bildungssprachliche Merkmale, von denen vermutet wird, dass sie eine zentrale Rolle im Bildungserfolg spielen und spezielle Hürden für Kinder mit unterschiedlichem familiären Hintergrund darstellen, identifiziert werden. Hierauf aufbauend sollte (2) zum anderen ein linguistisch und (sprach-)entwicklungspsychologisch fundiertes Instrument zur Erfassung bildungsbezogener sprachlicher Kompetenzen von Schülerinnen und Schülern in der Primarstufe entwickelt und validiert sowie die Bedeutung entsprechender Fähigkeiten für den Schulerfolg empirisch geprüft werden. Speziell erfasst werden sollten Kompetenzen des funktional-integrativen Hörverstehens bildungssprachlich anspruchsvoller Texte sowie sprachkomponenten-bezogene bildungssprachliche Fähigkeiten (Wortschatz, Verständnis von Satzverbindungen). In einer Messwiederholungsstudie mit zwei Alterskohorten wurden die bildungssprach-bezogenen Aufgaben für die Jahrgangsstufen 2 bis 4 validiert. Über die Validierung hinaus sollten zugleich Erkenntnisse über die Entwicklung und prognostische Bedeutung von bildungssprachlichen Kompetenzen gewonnen werden. In der ersten Projektphase (BiSpra I; siehe Abschlussbericht BiSpra I; Weinert, Stanat, Redder, Dragon, Heppt & Uessler, 2014) stand das erstgenannte Ziel im Fokus; in der zweiten Projektphase lag der Schwerpunkt auf dem zuletzt genannten Ziel. Das Projekt war interdisziplinär angelegt. BiSpra II bestand aus zwei Teilprojekten – einem psychologischen Teilprojekt (Teilprojekt Bamberg; Leitung: Prof. Dr. Sabine Weinert) und einem erziehungswissenschaftlichen Teilprojekt (Teilprojekt Berlin; Leitung: Prof. Dr. Petra Stanat). Um eine linguistische Fundierung der Messung von bildungssprachlichem Wortschatzwissen sicherzustellen, kooperierte insbesondere das Bamberger Teilprojekt zudem für einen Zeitraum von sechs Monaten mit der Arbeitsgruppe von Prof. Dr. Angelika Redder (Universität Hamburg). Die Projektleitungen und Mitarbeiterinnen der Teilprojekte standen in stetem Austausch und wechselseitiger Beratung und arbeiteten bei allen Schritten des empirischen Vorgehens (u. a. Materialentwicklung, Erprobung der Aufgaben, Versuchsplanung, Umsetzung der Studien) sehr eng zusammen. Im Folgenden werden zunächst die Zielsetzungen des Projekts und die organisatorischen Rahmenbedingungen beschrieben sowie insbesondere ein Überblick über die entwickelten Aufgaben/Aufgabenformate gegeben. Hieran schließt sich eine Darstellung wichtiger Ergebnisse an.

1 Projektziele / Aufgabenstellung

Ziel der beiden Teilprojekte war es, im interdisziplinären Verbund ein diagnostisches Instrument zu entwickeln, mit dem bedeutsame bildungssprachliche Kompetenzen im Grundschulalter reliabel und valide erfasst werden können. Dieses Instrument sollte so konzipiert sein, dass die bildungssprachlichen Fähigkeiten von der zweiten bis zur vierten Jahrgangsstufe eingeschätzt werden können. Dies soll es mittelfristig ermöglichen, den Entwicklungsverlauf in schulbezogenen sprachlichen Kompetenzen zu beobachten und gegebenenfalls gezielte Förderungen einzuleiten. Zu diesem Zweck wurden aufbauend auf den Entwicklungsarbeiten und empirischen Ergebnissen der ersten Projektphase zusätzliche Aufgaben zur Erfassung von bildungssprachlichem Hörverstehen, allgemein bildungssprachlichem Wortschatzwissen und dem Verständnis von Satzverbindungen (Konnektoren) konzipiert. Anhand ergänzender Testverfahren wurden Analysen zur Kriteriumsvalidität der entwickelten Aufgaben durchgeführt. Die längsschnittliche Anlage der Studie mit zwei Messzeitpunkten erlaubte zudem, erste Erkenntnisse über die (prognostische) Bedeutung dieser Kompetenzen für den Schulerfolg zu gewinnen, die im anschließenden Projekt „BiSpra-Aufgaben“ noch vertieft werden sollen.

2 Planung und Ablauf des Vorhabens sowie Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Aufbauend auf den Entwicklungsarbeiten und Ergebnissen der ersten Projektphase wurde die zweite Phase im Zeitraum von März 2013 bis Ende Juli 2016 durchgeführt. Verschiebungen gegenüber der ursprünglich beantragten Laufzeit ergaben sich (a) durch eine Verzögerung bei der Bewilligung (Febr. 2013) sowie (b) aufgrund eines personellen Wechsels innerhalb des Bamberger Teilprojekts, durch den auch eine kostenneutrale Verlängerung der Projektlaufzeit notwendig wurde.

2.1 Entwicklung und Weiterentwicklung von Aufgaben / Instrumenten zur Erfassung bildungssprachlicher Kompetenzen und Pilotierung

2.1.1 Textbezogenes Hörverstehen

Um zu ermitteln, inwieweit Kinder deutscher und Kinder nicht-deutscher Familiensprache in der Lage sind, Informationen aus (bildungs-)sprachlich unterschiedlich anspruchsvollen Texten zu verstehen und Schlussfolgerungen daraus zu ziehen, wurden unter Federführung des Berliner Teilprojekts kurze Hörverstehenstexte und dazugehörige Items – aufbauend auf den Entwicklungsarbeiten der ersten Projektphase – weiterentwickelt. Inhaltlich erzählen die Texte eine fiktive Geschichte von einem unbekanntem Planeten; mögliche Vorwissenseffekte auf die Verstehensleistung und Aufgabenbearbeitung sollten dabei möglichst gering gehalten werden. Die Texte enthalten daher zudem unbekannte Pseudowörter, deren Bedeutung aus dem Kontext erschlossen werden kann. Einige der Texte und Items, die in der ersten Förderphase entwickelt und eingesetzt worden waren, wurden beibehalten und in der Validierungsstudie zur Erfassung bildungssprachlichen Hörverstehens (Textebene) erneut eingesetzt. Zusätzlich wurden neue Texte und Items konstruiert, deren Schwierigkeitsgrad auch für ältere Grundschul Kinder (vierte Jahrgangsstufe) geeignet ist. Im Unterschied zur ersten Förderphase, in der verschiedene lexikalische und grammatische Anforderungen systematisch variiert und in ihren Effekten auf die Verstehensleistungen untersucht wurden, wurden in der

zweiten Projektphase gezielt anspruchsvolle (bildungs-)sprachlich geprägte Texte entwickelt, um die Leistungen auch der Viertklässler differenziert mittels des weiterentwickelten Instruments abbilden zu können.

Die Befunde der ersten Förderphase haben gezeigt, dass sich – bezogen auf die dort entwickelten Hörverstehensaufgaben (Textebene) – insbesondere der bildungssprachlich geprägte Wortschatz, nicht aber die bildungssprachlich geprägte Grammatik auf die Hörverstehensleistungen der Schülerinnen und Schüler ausgewirkt hat (Heppt, Stanat, Dragon, Berendes & Weinert, 2014). Dass sich speziell der bildungssprachlich geprägte Wortschatz und weniger die Komplexität der grammatischen Konstruktionen als besondere Hürde herausgestellt hat, könnte möglicherweise damit zusammenhängen, dass in dieser Studie einige Items auch ohne ein vollständiges Verständnis der jeweiligen grammatischen Strukturen korrekt beantwortet werden konnten. So war etwa eine gezielte Überprüfung der Bedeutung des Verständnisses von Satzverbindungen mit Konnektoren aufgrund des experimentellen Designs nicht möglich, da sich die entsprechenden Satzzusammenhänge in den lexikalisch und grammatisch einfachen Texten kaum angemessen ausdrücken ließen. Bei der Entwicklung zusätzlicher Hörverstehensaufgaben lag daher ein besonderes Augenmerk darauf, auch das Verständnis grammatischer Strukturen, die als charakteristisch für Bildungssprache gelten, gezielt zu erfassen. Komplexe Nebensatzstrukturen, Konnektoren und lange Nominal- und Präpositionalphrasen wurden hierzu vermehrt in die Hörverstehenstexte aufgenommen und zudem gezielt Items entwickelt, die explizit auf das Verständnis dieser syntaktischen Merkmale abzielen (vgl. Abbildung 1a, b).

Zwar ist Sambelos Verunsicherung zunächst deutlich zu erkennen, jedoch klärt sich seine Miene umgehend auf, als er Philipp, Mia und Emma entdeckt. Um seiner freudigen Überraschung Ausdruck zu verleihen, führt er seine Hände zu den Schultern und bewegt die angewinkelten Arme im Wechsel auf und ab, während er gleichzeitig schnalzende Zungenlaute erzeugt. Wie Sambelos anschließenden Erläuterungen zu entnehmen ist, habe er dieses Verhalten als Ausdruck seiner Malkigkeit gezeigt, womit auf Kumubila die Freude über ein unerwartetes Wiedersehen bezeichnet wird. Der Begriff der Malkigkeit kann sich sowohl auf Kinder und Jugendliche als auch auf Erwachsene beziehen, sofern diese keine eigenen Kinder haben.

Bildungssprachliche Wörter	Nominalisierungen	Lange Nominal- bzw. Präpositionalphrasen
Konnektoren	Passiv/unpersönliche Formulierungen	Eingeschobene Relativsätze
Funktionsverbgefüge	Partizipien	

Abbildung 1a: Illustration bildungssprachlicher Merkmale anhand eines der Hörverstehenstexte (aus Schuth et al., 2015, S. 99).

Beispielitem 1:	
Können sich Sambelos Eltern malkig fühlen?	ja <input type="checkbox"/> nein <input type="checkbox"/>
Beispielitem 2:	
Können sich nur Kinder malkig fühlen?	ja <input type="checkbox"/> nein <input type="checkbox"/>

Abbildung 1b: Illustration dazugehöriger Items (Hörverstehenstext s. Abbildung 1a; aus Schuth et al., 2015, S. 99).

Um Beispielitem 1 korrekt beantworten zu können, müssen die Kinder die Bedeutung des konditionalen Konnektors *sofern* kennen. Zur Beantwortung von Beispielitem 2 muss die Bedeutung des zweiteiligen Konnektors *sowohl... als auch* verstanden werden.

Über die Aufgaben zum textbezogenen bildungssprachlichen Hörverständnis hinaus wurden unter der Federführung des Bamberger Teilprojekts zwei sprachkomponenten-bezogene Skalen entwickelt, die speziell der Erfassung des Verständnisses von Satzverbindungen (Konnektoren) und von bildungssprachlichem Wortschatzwissen dienen.

2.1.2 Aufgaben zum Verständnis von Satzverbindungen mit Konnektoren

Als besonders wichtig für das Verständnis komplexer sprachlicher Zusammenhänge, wie sie in der Schule über Bildungssprache vermittelt werden, gilt das Verständnis von „Konnektoren“. Konnektoren, wie z. B. „trotzdem“, „dadurch“ oder „bevor“, sind ein häufig genutztes Mittel zur Bildung komplexer Sätze bzw. zur Vermittlung komplexer Sinnzusammenhänge und gelten daher als ein besonders charakteristisches Merkmal von Bildungssprache (z. B. Bailey & Butler, 2003).

So unstrittig die Bedeutung des Verständnisses von Konnektoren in Zusammenhang mit bildungssprachlichen Anforderungen ist, so schwierig gestaltet sich seine Erfassung. Wie in Schuth, Heppt, Köhne, Weinert und Stanat (2015) dargestellt⁴, konnte bei der Auswahl eines geeigneten Aufgabenformats in der zweiten Förderphase von BiSpra auf Erfahrungen aus der ersten Projektlaufzeit zurückgegriffen werden. Die Ergebnisse aus mehreren Pilotierungsstudien zu unterschiedlichen Aufgabentypen sind ausführlich in Dragon, Berendes, Weinert, Heppt und Stanat (2015) beschrieben. Erprobt wurden zunächst Aufgaben, bei denen die Kinder die inhaltliche Plausibilität von Sätzen beurteilen sollten, die über unterschiedliche Satzverbindungen (Konnektoren) verbunden waren. Die Plausibilität beziehungsweise fehlende Plausibilität der Aussagen/Sätze ergab sich jeweils durch einen entweder passenden oder unpassenden Konnektor. Interessanterweise zeigte sich, dass die Kinder stärker auf die semantische Konsistenz der jeweiligen Inhalte der beiden Satzteile zu achten schienen als auf den Konnektor, den sie offenbar sogar ignorierten (vgl. hierzu auch Knoepke, Richter, Isberner, Naumann, Neeb & Weinert, 2016). Als ein geeigneteres Testformat zur Erfassung des Konnektorenverständnisses erwiesen sich in einer weiteren Studie dagegen Lückensatzaufgaben (vgl. auch Uessler, Runge & Redder, 2013). Da dieses Format nur für Kinder geeignet ist, die bereits über grundlegende Lesefähigkeiten verfügen, wurde in BiSpra II zusätzlich ein weiteres Aufgabenformat erprobt. Bei diesem sollte der Einfluss der Lesefähigkeit auf die Aufgabenlösung durch Illustrationen der Antwortmöglichkeiten vermindert

⁴ Die Ausführungen folgen in Teilen der Darstellung in Schuth et al. (2015).

werden. Dies sollte ein erster Schritt sein in Richtung eines Aufgabenformats, das einen Einsatz auch bei Schülerinnen und Schülern in der ersten Klasse ermöglicht. Durch farbige Illustrationen sollte dieses Format zudem für Kinder besonders ansprechend wirken und so potentiell das Interesse und die Aufmerksamkeit der Kinder erhöhen. Im folgenden Abschnitt wird zunächst das Lückensatzformat zur Erfassung des Verständnisses von Konnektoren beschrieben und im Anschluss das Bildformat.

Lückensatzformat zur Erfassung von Konnektorenverständnis. Die einzelnen Items/Aufgaben bestehen aus je zwei Sätzen oder aus einem Satz mit zwei aufeinanderfolgenden Teilsätzen (Rahmensätze). Jedes Item enthält genau eine Lücke. Aufgabe der Kinder ist es, aus jeweils vier Konnektoren den richtigen, in die Lücke passenden, auszuwählen und anzukreuzen. Die Stellung des Konnektors im Satz (Satzbeginn oder Satzmitte), die Wortart des Konnektors und die resultierende Wortstellung wurden bei der Itemkonstruktion systematisch variiert. Von den drei Distraktoren war je einer semantisch, nicht aber syntaktisch passend; zwei waren syntaktisch, nicht aber semantisch passend. Es wurden insgesamt 24 Lückensätze konzipiert, acht mit kausalen/konsekutiven, sechs mit temporalen, acht mit konzessiven und je einer mit einem konditionalen und einem modalen Konnektor/en (siehe Tabelle 1; für weitere Beispiele und insbesondere eine ausführliche Erläuterung der Auswahl der in die Aufgaben einbezogenen Konnektoren vgl. Schuth et al., 2015). Die Items wurden den Kindern sowohl auditiv dargeboten (gesprochen von demselben Sprecher wie die textbezogenen Hörverstehensaufgaben) als auch in den Aufgabenheften abgedruckt vorgelegt.

Tabelle 1

Beispielitems des Lückensatzformates (die richtige Antwortmöglichkeit ist kursiv markiert)

Konnektortyp (Position Konnektor, Wortart Konnektor, Wortstellung)	Itembeispiel	Antwortmöglichkeiten
konzessiv (Satzmedial, Adverb, Verbzweitstellung)	Heute kommt Elinas Lieblingsserie im Fernsehen, ___ liest sie lieber ein Buch.	<i>trotzdem</i> trotz/daher/also
temporal (Satzmedial, Präposition, Verbzweitstellung)	„Paul darf ___ seinem Sturz mit dem Fahrrad nicht mehr am Schulsport teilnehmen.“	<i>seit</i> seitdem/bei/vor

Anm.: Für weitere Beispiele siehe Schuth et al., 2015.

Bildgestütztes Format zur Erfassung von Konnektorenverständnis. Für das bildgestützte Aufgabenformat zur Erfassung von Konnektorenwissen wurden 25 Items konstruiert. Die Aufgabe der Kinder besteht bei diesem Format darin, zu einem unvollständigen Satz das semantisch korrekte Satzende auszuwählen und anzukreuzen. Die Satzenden bestehen aus einzelnen Wörtern (oder kurzen Phrasen) sowie passenden Abbildungen (siehe Abbildung 2). Jedes Item ist aus zwei Sätzen oder zwei Satzteilen gebildet, die durch einen Konnektor verbunden sind, dessen Verständnis essentiell für die Erfassung des Sinnzusammenhangs

ist. Bei der Konzipierung der Aufgaben wurde gezielt darauf geachtet, dass der Konnektor tatsächlich verstanden werden muss, um das jeweils semantisch sinnvolle Satzende auswählen zu können. Distraktoren wurden, wenn möglich, so gewählt, dass sie bei Ignorieren des Konnektors ein semantisch passendes Satzende ergeben würden (z. B. „Zebras“ in Abbildung 2).

Obwohl Eva im Zoo am liebsten die Zebras und Elefanten anschaut, geht sie dort heute mit ihrem Bruder zu den...

	<input type="checkbox"/> Zebras.
	<input type="checkbox"/> Großeltern.
	<input type="checkbox"/> Affen.

Abbildung 2: Beispielitem des bildgestützten Formats für das Verständnis von Satzverbindungen (aus Schuth et al., 2015, S. 102)

2.1.3 Aufgaben zur Erfassung bildungssprachlichen Wortschatzwissens

In Kooperation mit Angelika Redder wurden unter Federführung des Bamberger Teilprojekts sowohl eine bildungssprachliche Wortschatzliste für das Grundschulalter erstellt als auch Aufgaben zur Erfassung bildungssprachlichen Wortschatzwissens konzipiert. Eine ausführliche Darstellung der Entwicklungsarbeiten und ihrer Begründungszusammenhänge findet sich in Köhne, Kronenwerth, Redder, Schuth und Weinert (2015).⁵

Der bildungssprachliche Wortschatz gilt als ein zentrales Merkmal von Bildungssprache (z. B. Köhne et al., 2015; Coxhead, 2000). Dennoch lag zu Beginn der Arbeiten noch kein linguistisch begründetes und empirisch fundiertes Rational vor, um bildungssprachliche Ausdrücke systematisch identifizieren zu können. Als wesentliche Voraussetzung zur Entwicklung valider Aufgaben/Items wurden daher im Rahmen des Projekts kooperativ und unter Einbezug linguistischer Expertise der Arbeitsgruppe um Angelika Redder zunächst Quellen identifiziert, die einen Einblick erlauben, welche Wörter in der Schule überhaupt verwendet werden, und ein umfängliches Korpus aus verschiedenen authentischen mündlichen (z. B. Unterrichtsdiskurse) wie schriftlichen (z. B. Arbeitsblätter, Schulbücher) Sprachdaten erstellt

⁵ Die hier vorliegende Darstellung orientiert sich in Teilen an dem Beitrag von Köhne et al. (2015).

(s. ausführlich Köhne et al., 2015 zur Entwicklung des Hamburg-Bamberger Korpus). Abbildung 3 gibt einen Überblick über das umfangreiche Hamburg-Bamberger-Korpus.

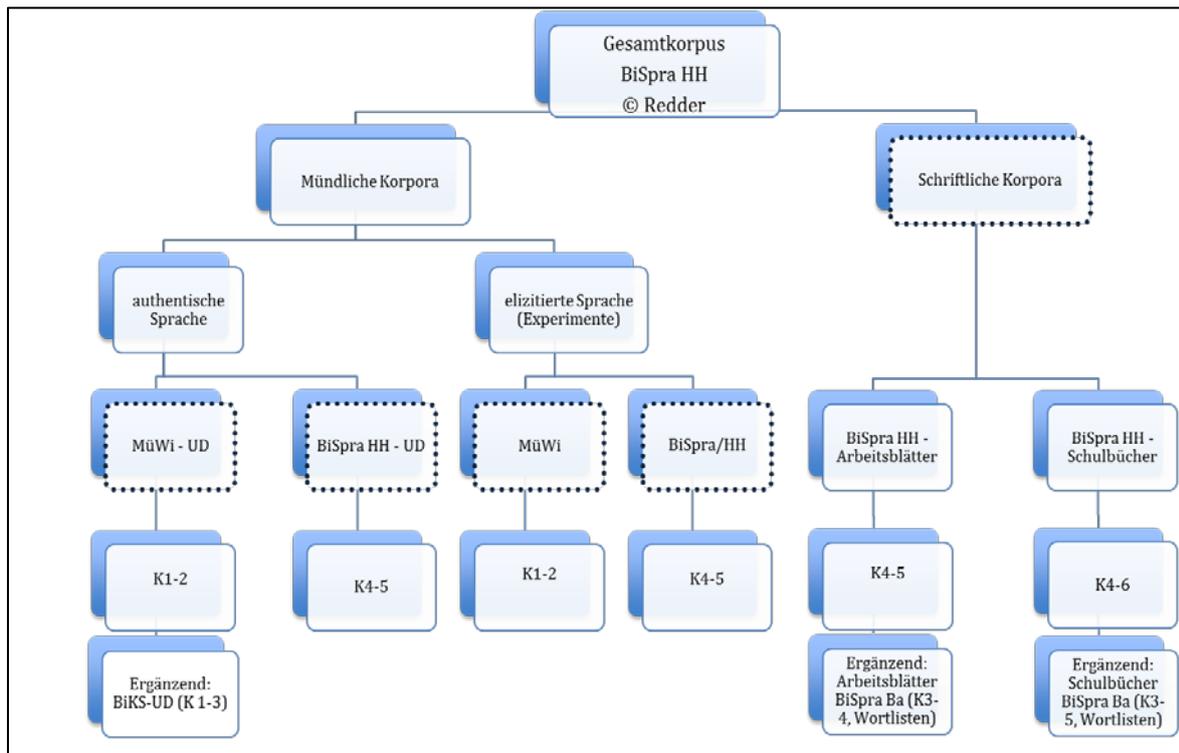


Abbildung 3: Übersicht über das Hamburg-Bamberger-Korpus von BiSpra II

Anmerkung: punktiertes Rahmen: Hauptlisten; K = Klasse/Jahrgangsstufe, UD = Unterrichtsdiskurs; MüWi, BiSpra HH, BiSpra Ba, BiKS: Drittmittelprojekte, in deren Rahmen die Datenbasis erhoben wurde (Abbildung entnommen aus Köhne et al., 2015, S. 76)

Die aus dem Hamburg-Bamberger-Korpus extrahierte Wortliste bietet eine reiche und authentische Quelle von „Schulsprache“: Sie beinhaltet sowohl mündliche als auch schriftliche Quellen und deckt Schüler- sowie Lehrersprache ab. Insgesamt umfasst sie knapp 700.000 Wörter (= token und knapp 30.000 types) aus mehr als 2.700 Transkripten.

Da es sich nicht bei allen Wörtern, die in der Schule verwendet werden, um bildungssprachliche Ausdrücke handelt, wurde im nächsten Schritt eine linguistisch und sprachpsychologisch begründete Auswahl der als bildungssprachlich einzustufenden Ausdrücke getroffen, wobei – mit Blick auf die Zielsetzung – nur solche einbezogen wurden, die als nicht fachspezifisch, sondern fachübergreifend bezeichnet werden können. Außerdem wurden nur Wörter berücksichtigt, die eine Minimalhäufigkeit von zwei Vorkommen im Hamburg-Bamberger-Korpus aufwiesen und mindestens eines der folgenden, formal gefassten Kriterien/Merkmale (oder Zugehörigkeiten) erfüllten (vgl. ausführlich Köhne et al., 2015):

1. sog. Instruktionvokabular, d. h. verbale oder nominale Bezeichnungen für sprachliches oder nicht-sprachliches Handeln, das in Aufgabenstellungen gefordert und insofern explizit relevant wird (*anlegen, erläutern, erklären, vergleichen, übersichtlich/Übersicht geben, Abbildung (erstellen)*)
2. Ausdrücke, die, lexikalisch isoliert betrachtet, ambig oder mehrdeutig (polysem) sind bzw. genauer: die eine konkrete, alltägliche, und eine (oder mehrere) abstraktere Be-

deutungen aufweisen (z. B. *abgeben, angeben, durchgehen, liegen, scheinen, abgeschlossen*)

3. Präfixverben, die sich aus einem Präfix wie *er-, ver-,* oder *be-* und einem alltäglich gebräuchlichen Verb wie *setzen, stellen* oder *legen* zusammensetzen
4. Partizipien/Adjektive/Adverbien, die komplexe Verknüpfungsvorgänge oder Modifizierungen erfordern (*bestimmt, folgend, beziehungsweise, quasi, einigermaßen*)
5. Trennbare Verben (Partikelverben)
6. Derivierte Nomina (*Anweisung, Bereich, Übergang*)

Die genannten Kriterien verbinden semantische und morphologische bzw. morphosyntaktische wie auch diskurspragmatische Aspekte. Auf dieser Basis wurde ein 118 Wörter umfassender, empirisch basierter bildungssprachlicher Ausdruckspool, die Hamburg-Bamberger-BiSpra-Liste, erstellt, die sich auf die unterrichtsbezogene Kommunikation in der Grundschule bezieht und als Grundlage der Aufgabenentwicklung zur Erfassung von bildungssprachlichem Wortwissen diente.

Bei der Konzeption der Items wurde darauf geachtet, dass die Aufgabenbearbeitung kein sehr breites oder spezielles Weltwissen erfordert und die Sätze (abgesehen von der Lücke im Lückensatzformat) lexikalisch und grammatisch altersangepasst formuliert sind. Die auszuwählenden Wörter sollen diskurspragmatisch zu der Struktur des Satzes und den anderen Wörtern im Satz passen. Die Einzelitems wurden zudem so gestaltet, dass sich eine lose Geschichte über den Tagesablauf eines Mädchens mit Schwerpunkt auf Erlebnissen in der Schule ergab.

Analog zu den beiden Aufgabenformaten zum Verständnis von Satzverbindungen (Konnektoren) wurden, wie im Folgenden beschrieben, auch für die Erfassung von bildungssprachlichem Wortschatzwissen Aufgaben sowohl im Lückensatzformat als auch im bildgestützten Format entwickelt und erprobt.

Lückensatzformat. Konzipiert wurden 29 Items aus je einem Satz mit einer Lücke, für die die Kinder das passende Wort aus drei Wahlmöglichkeiten auswählen sollen. Alle Wahlmöglichkeiten sind dem oben beschriebenen bildungssprachlichen Ausdruckspool (Hamburg-Bamberger-BiSpra-Liste) entnommen und ergeben einen grammatikalisch sinnvollen Satz.

Annas Freund Adam ist heute nicht in der Schule. Die Lehrerin _____, dass er krank ist.

- zeigt an
- nimmt an
- legt an

Abbildung 4: Beispielitem zum bildungssprachlichen Wortschatz im Lückensatzformat (aus Köhne et al., 2015, S. 80)

Bildgestütztes Format. Da Lückentexte, wie bereits angesprochen, grundlegende Lesefähigkeiten erfordern, wurde zudem ein zweites Aufgabenformat mit bildlicher Unterstützung erprobt: Dem Kind wird eine Frage über die Bedeutung eines Satzes gestellt, wobei es zwi-

schen zwei bildgestützten Antworten wählen kann. Dabei enthält jeder der 34 konzipierten Items/Sätze genau einen bildungssprachlichen Ausdruck (vgl. Abbildung 5).

Der Lehrer fragt Anne, was der Auslöser für den Streit mit Steffen war. Was möchte er wissen?

 Warum sie sich gestritten haben.

 Wie sie sich wieder vertragen haben.

Abbildung 5: Beispielitem zum bildungssprachlichen Wortschatz im bildgestützten Format (aus Köhne et al., 2015, S. 80)

2.2 Rekrutierung von Schulen und Horten für die Teilnahme an den Erhebungen und Genehmigungsverfahren

Rekrutierung. Die Rekrutierung der Institutionen und Teilnehmer/-innen für die Pilotierungsstudien erfolgte durch die Projektmitarbeiterinnen. Die Pilotierungsstudien fanden zum einen in 24 Kinderhorten in Nürnberg statt (Bamberger Teilprojekt). Dabei wurden gezielt verschiedene Stadtteile berücksichtigt, die sich hinsichtlich ihrer Anteile an Familien mit Migrationshintergrund deutlich voneinander unterscheiden, um die gewünschte Heterogenität der Stichprobe in Bezug auf den familiären Migrationsstatus und den sozio-ökonomischen Hintergrund zu gewährleisten. Zum anderen wurden vier Berliner Grundschulen einbezogen (Berliner Teilprojekt). Basierend auf Informationen der Senatsverwaltung für Bildung, Wissenschaft und Forschung waren hierzu zunächst 15 Schulen mit einer sprachlich wie sozio-ökonomisch heterogenen Schülerschaft kontaktiert worden.

Die Rekrutierung der Schulen für die Validierungsstudie erfolgte – unter der Federführung des Berliner Teilprojekts – in enger Absprache mit dem Datenerhebungsinstitut IEA DPC Hamburg. Die Stichprobe umfasste zunächst 196 Schulen aus vier Bundesländern – Hamburg, Brandenburg, Nordrhein-Westfalen und Baden-Württemberg. Um eine möglichst heterogene Schülerschaft in die Stichprobe einbeziehen zu können, wurden bei der Zusammenstellung der Schulstichprobe Informationen zum zugewanderungsbezogenen und, falls vorhanden, zum sozio-ökonomischen Hintergrund der Schülerinnen und Schüler berücksichtigt. Die ausgewählten Schulen wurden anschließend von den Projektmitarbeiterinnen postalisch und telefonisch kontaktiert und gebeten, bei der Studie mitzuwirken. Zum ersten Messzeitpunkt konnten 38 Schulen für die Teilnahme an der Studie gewonnen werden, von denen 36 auch am zweiten Messzeitpunkt teilnahmen. In jeder Schule wurde je eine zweite und eine dritte Klasse zufällig für die Teilnahme an der Studie gezogen.

Genehmigungsverfahren. Zur Durchführung der Pilotierungsstudie in Horten (Nürnberg) wurde ein informiertes Einverständnis der Hortleitungen und der Eltern eingeholt. Für die Berliner Grundschulen wurde eine Genehmigung der Erhebungen nach § 65 des geltenden Schulgesetzes durch die Berliner Senatsverwaltung für Bildung, Jugend und Wissenschaft eingeholt. Letztere setzte das schriftliche Einverständnis der Eltern in die Studienteilnahme ihrer Kinder voraus.

Das Genehmigungsverfahren für die Teilnahme an der Validierungsstudie wurde unter Federführung des Bamberger Teilprojekts in enger Absprache mit dem Datenerhebungsinstitut IEA DPC Hamburg durchlaufen. Für Brandenburg, Baden-Württemberg und Hamburg wurde hierzu die fachliche und datenschutzrechtliche Bewilligung durch die jeweiligen Kultusministerien eingeholt und erteilt; in Nordrhein-Westfalen wurde eine fachliche Genehmigung durch das Ministerium für Schule und Weiterbildung erteilt. In allen vier Bundesländern war die schriftliche Einwilligung der Eltern in die Studienteilnahme ihrer Kinder Voraussetzung für die Teilnahme am Projekt.

2.3 Zentrale Erhebungen während der Projektlaufzeit von BiSpra II

Durchgeführt wurden zwei Pilotierungsstudien sowie eine längsschnittliche Hauptstudie (Validierungsstudie). Der Fokus der Pilotierungen lag zum einen auf der empirischen Prüfung der neu entwickelten textbezogenen Hörverstehensaufgaben sowie zum anderen auf der Erprobung der Aufgaben zur Erfassung von bildungssprachlich geprägtem Wortschatzwissen und dem Verständnis von Satzverbindungen. Auf der Basis der Pilotierungsergebnisse sollten zum einen geeignete Aufgaben ausgewählt und zum anderen – bezogen auf die beiden Sprachkomponenten (Wortschatz, Konnektoren) – die beiden beschriebenen Aufgabenformate getestet werden. Über das jeweilige bildgestützte Format wurde zugleich geprüft, ob die konzipierten Aufgaben prinzipiell auch bei jüngeren Kindern, die noch nicht über grundlegende Lesefähigkeiten verfügen, einsetzbar sind. Wie im Folgenden dargestellt wird, erwiesen sich diese Aufgaben jedoch nicht als sehr erfolgversprechend. Die jeweiligen Erhebungszeiträume der Studien können Abbildung 6 entnommen werden.

Die **Pilotierungserhebungen** fanden im Herbst 2013 statt. Insgesamt nahmen 624 Schülerinnen und Schüler der Jahrgangsstufen 2, 3 und 4 an Grundschulen in Berlin sowie Kinderhorten in Nürnberg teil (vgl. Abschnitt 2.2). Tabelle 2 zeigt, dass – der Rekrutierungsstrategie entsprechend – Kinder mit nicht-monolingual deutschem Sprachhintergrund im Vergleich zum bundesdeutschen Durchschnitt überrepräsentiert waren (siehe Tabelle 2). Neben den Testaufgaben für die Schülerinnen und Schüler kam ein Elternfragebogen zum Einsatz, um u. a. wichtige Informationen zur Spracherwerbssituation der Kinder zu erfassen.

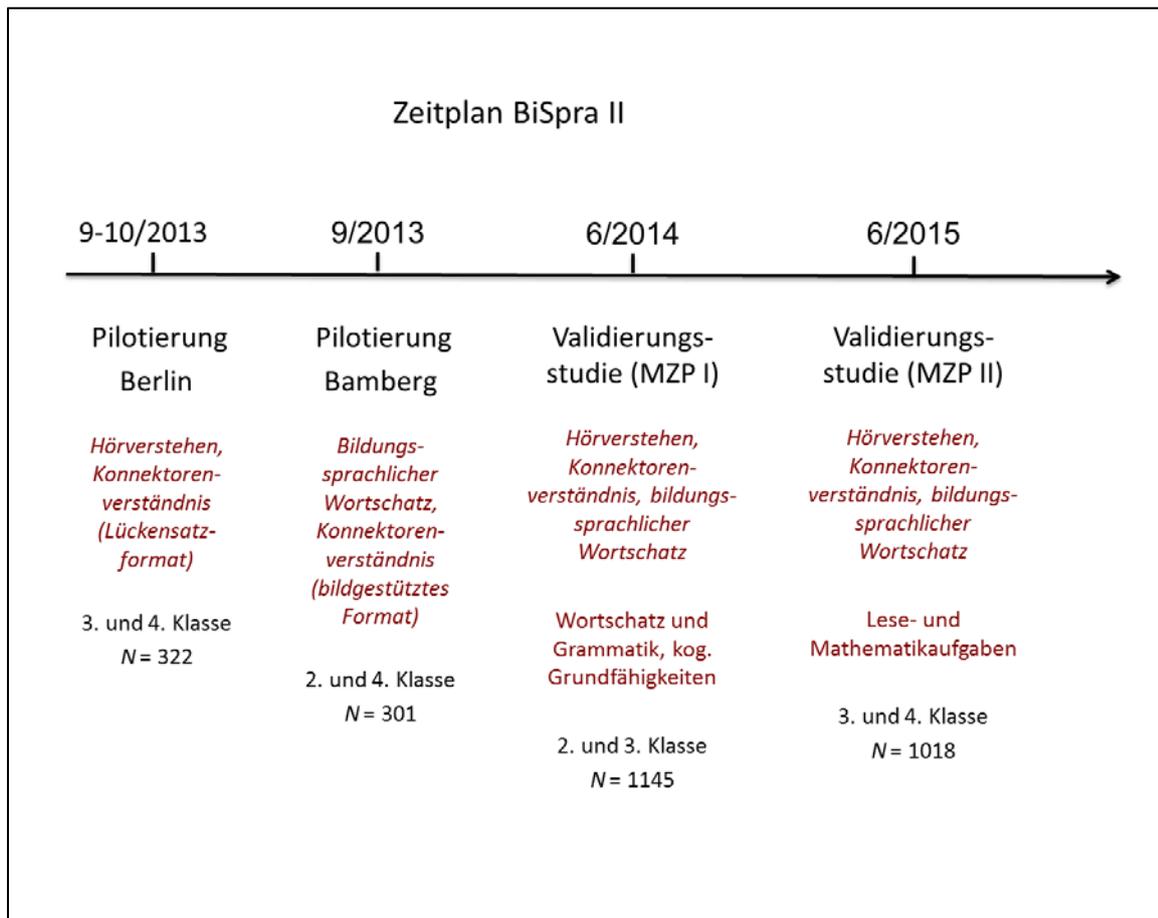


Abbildung 6: Übersicht über die in BiSpra II durchgeführten Erhebungen (inklusive Stichprobenumfängen)

Die **erste Erhebung der Messwiederholungsstudie (Validierungsstudie)** fand im Juni 2014 statt. In der ersten Erhebungsphase wurden sowohl die selbst entwickelten Hörverstehens-, Konnektoren- sowie Wortschatzaufgaben als auch bereits etablierte Aufgaben zur Erfassung des allgemeinen Wortschatzes, der Grammatik, der kognitiven Grundfähigkeiten und des phonologischen Arbeitsgedächtnisses eingesetzt. Insgesamt nahmen 1145 Zweit- und Drittklässler an den Erhebungen teil (siehe Abbildung 6 und Tabelle 2)⁶. Zudem wurden die Eltern dieser Schülerinnen und Schüler schriftlich u. a. zur Spracherwerbssituation der Kinder und zum sozio-ökonomischen und bildungsbezogenen Hintergrund der Familie befragt. Den Lehrkräften wurde ebenfalls ein Fragebogen ausgehändigt, mit der Bitte, Fragen zu ihrer Lehrtätigkeit, ihrer Unterrichtsgestaltung sowie der Zusammensetzung ihrer Schülerschaft zu beantworten.

Die Erhebungen des **zweiten Messzeitpunkts** der Validierungsstudie fanden im Juni 2015 statt. Neben den im Projekt entwickelten Aufgaben zum bildungssprachlichen Hörverstehen, zum Verständnis von Satzverbindungen mit Konnektoren und zum bildungssprachlichen

⁶ Durch einen Erhebungsfehler (Vertauschung des Testmaterials der beiden Klassenstufen in einer Schule) können in die Analysen der Daten des ersten Messzeitpunkts nur 37 der 38 Schulen und 1107 der 1145 Kinder einbezogen werden. Derselbe Erhebungsfehler trat auch zum zweiten Messzeitpunkt auf, betraf dort aufgrund des Testheftdesigns jedoch nur die Hörverstehensaufgaben.

Wortschatz wurden zum zweiten Messzeitpunkt auch Aufgaben zum Leseverständnis und zu mathematischen Kompetenzen eingesetzt, um zentrale schulisch vermittelte Kompetenzen zu erfassen. Am zweiten Messzeitpunkt nahmen 1018 Schülerinnen und Schüler teil, von denen 993 bereits zum ersten Messzeitpunkt teilgenommen hatten (siehe Tabelle 2).

Eltern, die den Elternfragebogen im Vorjahr noch nicht ausgefüllt hatten, bekamen diesen erneut vorgelegt. In Klassen, in denen es zwischen beiden Messzeitpunkten zu einem Lehrerwechsel gekommen war, wurden auch die neuen Lehrkräfte gebeten, den Lehrerfragebogen auszufüllen. Tabelle 3 gibt einen Überblick über die Konstrukte, die in den einzelnen Erhebungen erhoben wurden, und die eingesetzten Testverfahren.

Tabelle 2

Stichprobengrößen sowie Geschlechter-, Klassenstufen- und Familiensprachverteilung für die Erhebungen

Erhebung	N (♂; ♀)	N je Klassenstufe	N je Familiensprache (1 = monolingual deutsch, 2 = nicht-monolingual deutsch) ⁷
Pilotierung	624 (293; 331)	2. Kl.: 177	1 = 243
		3. Kl.: 154	2 = 381
		4. Kl.: 293	
Validierungsstudie Messzeitpunkt I	1145 (571; 574)	2. Kl.: 546	1 = 490 ⁸
		3. Kl.: 599	2 = 652
Validierungsstudie Messzeitpunkt II	1018 (516; 502)	3. Kl.: 478	1 = 412 ⁹
		4. Kl.: 540	2 = 577

Anmerkungen. N = Anzahl; Kl = Klassenstufe

⁷ Die Gruppe der Schülerinnen und Schüler (SuS) mit nicht-monolingual deutscher Familiensprache umfasst dabei Kinder mit teilweise sehr unterschiedlichen Spracherwerbsbiographien. Bilinguale Schülerinnen und Schüler, die Deutsch und eine weitere Muttersprache simultan erwerben, zählen ebenso dazu wie SuS, die Deutsch erst zu einem späteren Zeitpunkt und folglich als Zweitsprache erwerben.

Im Fragebogen wurde der Sprachhintergrund zu den beiden Messzeitpunkten (MZP) jeweils unterschiedlich erfasst. Zu MZP 1 wurde nach den zu Hause gesprochenen Sprachen gefragt. Hierauf basiert die Zuordnung in Tabelle 2. Zu MZP 2 wurde nach der Häufigkeit gefragt, mit der die Schülerinnen und Schüler (SuS) zu Hause mit verschiedenen Personen Deutsch sprechen. Diese Frage wurde bei der Kategorienbildung in Tabelle 2 nicht berücksichtigt.

⁸ Von 3 Kindern liegen keine Informationen zu ihrem Sprachhintergrund vor, weshalb sich die angegebenen Werte nicht auf 1145 addieren.

⁹ Von 29 Kindern liegen keine Informationen zu ihrem Sprachhintergrund vor, weshalb sich die angegebenen Werte nicht auf das N von 1018 addieren.

Tabelle 3

Testverfahren, die zur Erfassung der einzelnen Konstrukte in den Erhebungen eingesetzt wurden

Erhebung	erfasste Konstrukte	Testverfahren/Untertest	
Pilotierung	Hörverstehen (Text)	selbst entwickelte Aufgaben	
	Satzverbindungen mit Konnektoren	selbst entwickelte Aufgaben	
	bildungssprachl. Wortschatz	selbst entwickelte Aufgaben	
Validierungsstudie Messzeitpunkt I	Hörverstehen (Text)	selbst entwickelte Aufgaben	
	Satzverbindungen mit Konnektoren	selbst entwickelte Aufgaben	
	bildungssprachlicher Wortschatz	selbst entwickelte Aufgaben	
	Grammatik (rezeptiv)	34 Items (Kurzversion TROG-D; Fox, 2007)	
	Allg. Wortschatz (rezeptiv)	60 Items aus dem PPVT (PPVT III; Bulheller & Häcker, 2003; PPVT-R unveröffentlichte Forschungsversion; Roßbach, Tietze & Weinert, 2005)	
	kognitive Grundfähigkeiten	12 Matrizenaufgaben (Subtest, CFT 1; Cattell, Weiß & Osterland, 1997)	
	phonologisches Arbeitsgedächtnis	Zahlen nachsprechen (Subtest, K-ABC; Melchers & Preuß, 2009)	
Validierungsstudie Messzeitpunkt II	Hörverstehen (Text)	selbst entwickelte Aufgaben	
	Satzverbindungen mit Konnektoren	selbst entwickelte Aufgaben	
	bildungssprachlicher Wortschatz	selbst entwickelte Aufgaben	
	Leseverständnis	15 Items (Subtest, ELFE; Lenhard & Schneider, 2006)	
	mathematische Basiskompetenzen	40 Items (Subtest, HRT 1-4; Haffner, Baro, Parzer & Resch, 2005)	
	mathematische Kompetenzen		5 Aufgaben aus dem Aufgabenpool des IQB zur Operationalisierung der Bildungsstandards (Granzer, Walther, Winkelmann, Robitzsch & Köller, 2008)

2.4 Auswahl und Schulung von Testleiterinnen und Testleitern

Pilotierungsstudien. Als Testleiterinnen und Testleiter wurden Studierende inhaltlich einschlägiger Studiengänge (Psychologie, Erziehungswissenschaft, Lehramt) mit relevanter praktischer Erfahrung (z. B. in der Durchführung von Erhebungen in Grundschulen) der Otto-Friedrich Universität Bamberg, der Freien Universität Berlin und der Humboldt-Universität zu Berlin rekrutiert (über Aushänge, Anzeigen auf der Homepage der Universitäten, Email-Verteiler).

Um eine standardisierte Testdurchführung sicherzustellen, wurden in beiden Teilprojekten umfassende Schulungen durchgeführt, die für alle Testleiterinnen und Testleiter verpflichtend waren. Im Rahmen der Schulungen wurden das Projekt, dessen Ziele sowie insbesondere die zum Einsatz kommenden Verfahren und deren Durchführung ausführlich dargestellt. Anschließend wurde die Testdurchführung (inklusive der Bedienung der technischen Geräte zum Abspielen der Texte und Items) im Rollenspiel eingeübt und der flexible Umgang mit Problemsituationen besprochen. Die Schulung umfasste ca. 4 Stunden.

Die Validierungsstudie wurde durch ein Erhebungsinstitut (IEA DPC Hamburg) durchgeführt. Auch hier wurden die Erhebungen von trainierten Studierenden einschlägiger Fächer (z. B. Psychologie, Erziehungswissenschaften, Lehramt), die überwiegend schon über Erfahrung in der Durchführung standardisierter Leistungserhebungen verfügten, durchgeführt. Auch hier durchliefen alle Testleiterinnen und Testleiter eine umfangreiche Schulung (wie oben beschrieben), die von den Projektmitarbeiterinnen gemeinsam mit den verantwortlichen Mitarbeiterinnen und Mitarbeitern des IEA DPC durchgeführt wurde. Zu jedem Messzeitpunkt wurden zwei Schulungstermine in Hamburg und Frankfurt angeboten.

2.5 Durchführung der Validierungsstudie und Erhebungsdesign

Die Validierungsstudie wurde von beiden Teilprojekten in enger Abstimmung geplant und durchgeführt.

Überblick. Durchgeführt wurde eine Messwiederholungsstudie mit zwei Kohorten (Kohortensequenzdesign). Zum ersten Messzeitpunkt wurden Kinder der 2. und 3. Klassenstufe einbezogen; der zweite Messzeitpunkt fand ein Jahr später statt (3. / 4. Klassenstufe). Während zum ersten Messzeitpunkt neben den im Projekt entwickelten Skalen zum bildungssprachlichen textbezogenen Hörverstehen, zum Verständnis von Satzverbindungen (Konnektoren) sowie zum bildungssprachlichen Wortschatz zusätzliche Validierungsinstrumente (Tests zum allgemeinen rezeptiven Wortschatz, Grammatik, kognitive Grundfähigkeiten, phonologisches Arbeitsgedächtnis) eingesetzt wurden, wurden zum zweiten Messzeitpunkt zusätzlich „outcome“-Indikatoren (Leseverständnis, mathematische Kompetenzen) eingesetzt. Die BiSpra-Skalen wurden mit klassenstufen-angepasster Schwierigkeit vorgegeben (s. unten) und lagen aufgrund der begrenzten Erhebungszeiten als Kurz- und Langversion vor.

Durchführung. Die Erhebungen wurden in den Klassenräumen der teilnehmenden Schulen durchgeführt und fanden im Klassenverband statt. Eine Ausnahme bildete lediglich das Zahlen nachsprechen, das im Einzelsetting durchgeführt wurde.

Um eine möglichst standardisierte Testdurchführung zu gewährleisten, wurden neben den im Projekt entwickelten Aufgaben zum textbezogenen Hörverstehen, zum Verständnis von Satzverbindungen mit Konnektoren und zum bildungssprachlichen Wortschatz auch die Items der Validierungsvariablen (allgemeiner Wortschatz- und Grammatiktest) vor der Erhebung durch einen professionellen Sprecher in einem Tonstudio eingesprochen und den Kindern während der Testung von CD vorgespielt. Zu beiden Messzeitpunkten nahmen die Schülerinnen und Schüler an je einer Testsitzung teil, die inklusive Pause ca. 90 Minuten dauerte.

Testheftdesign und Durchführung zum ersten Messzeitpunkt. Eingesetzt wurden sieben verschiedene Testhefte (siehe Tabelle 4). Dies war notwendig, (a) um die Erhebungszeit be-

grenzt zu halten und (b) um die Aufgabenschwierigkeit der jeweiligen Klassenstufe anzupassen (Hörverstehensaufgaben (Testheft 1A & 1B); Aufgaben zum bildungssprachlichen Wortschatz (Testheft 2A & 2B)). Zur Erfassung des Verständnisses von Satzverbindungen mit Konnektoren (Testheft 3) wurde in beiden Klassenstufen (auf Basis der Pilotierungsergebnisse) dasselbe Aufgabenset eingesetzt. Je Skala wurden jeweils eine kurze und eine lange Version gebildet.

Die Testhefte unterschieden sich zudem darin, ob sie zusätzlich zu einer der BiSpra-Skalen (Langversion) weitere Testverfahren enthielten, die der Konstruktvalidierung dienen sollten (Testhefte 1-3), oder ob sie ausschließlich die im Projekt entwickelten Skalen (Kurzversionen) enthielten (Testheft 4).

Die Testhefte 1 bis 3 enthielten somit jeweils einen der drei BiSpra-Tests in einer Langversion sowie Aufgaben zum allgemeinen Wortschatz, zum Grammatikverständnis und zu den kognitiven Grundfähigkeiten (siehe Tabelle 4). Testheft 4 enthielt die projekteigenen Aufgaben in gekürzter Fassung, um alle drei bildungssprachbezogenen Skalen gemeinsam erfassen zu können, ohne die zur Verfügung stehende Testzeit zu überschreiten.

Tabelle 4

Testheftdesign zum ersten Messzeitpunkt

MZP 1	TH1A 2. Klasse	TH1B 3. Klasse	TH2A 2. Klasse	TH2B 3. Klasse	TH3 2. & 3. Klasse	TH4A 2. Klasse	TH4B 3. Klasse
Schülerfragebogen							
BiSpra Hörverstehen 2. Klasse							
BiSpra Hörverstehen 3. Klasse							
BiSpra Wortschatz 2. Klasse							
BiSpra Wortschatz 3. Klasse							
BiSpra Konnektoren							
TROG-D							
PPVT							
CFT							
Zahlen nachsprechen							
N	141	156	90	106	344	153	155

Die Testhefte wurden den Klassen zufällig zugewiesen, wobei alle Schülerinnen und Schüler einer Klasse dasselbe Testheft erhielten. Insgesamt sollte darauf geachtet werden, dass alle Testhefte annähernd gleich häufig zum Einsatz kamen.

Zuerst führte der Testleiter / die Testleiterin die Kinder durch den Schülerfragebogen. Danach wurden den Kindern je nach Testheft die BiSpra-Aufgaben zum textbezogenen Hörverstehen, zu den Satzverbindungen mit Konnektoren oder zum bildungssprachlichen Wortschatz auditiv präsentiert. In den Testgruppen mit den Testheften 1 bis 3 folgten anschließend die Aufgaben zum allgemeinen Grammatikverständnis, zum allgemeinen Wortschatz und zu den kognitiven Grundfähigkeiten, die die Kinder alleine und ohne Unterstützung des Testleiters oder der Testleiterin bearbeiteten. Im Anschluss an die Testsitzung im Klassenverband bearbeitete eine kleine Stichprobe von 32 Kindern Aufgaben zum phonologischen Arbeitsgedächtnis im Einzelsetting. Diese Kinder wurden zufällig ausgewählt.

Die Elternfragebögen wurden den Kindern mit Umschlägen mit nach Hause gegeben, so dass die Eltern diese ausfüllen und an die Lehrkräfte zurückgeben konnten.

Das Testheftdesign und die Erhebungsdurchführung zum zweiten Messzeitpunkt erfolgten analog zum ersten Messzeitpunkt. Da sich die Aufgaben zum bildungssprachlichen Wortschatz in der Pilotierungsstudie für die Jahrgangsstufen 3 und 4 als gleichermaßen geeignet erwiesen haben, wurde für beide Jahrgangsstufen dasselbe Aufgabenset eingesetzt (siehe Tabelle 5).

Tabelle 5

Testheftdesign zum zweiten Messzeitpunkt

MZP 2	TH1A 3. Klasse	TH1B 4. Klasse	TH2 3. & 4. Klasse	TH3 3. & 4. Klasse	TH4A 3. Klasse	TH4B 4. Klasse
Schülerfragebogen						
BiSpra Hörverstehen 3. Klasse						
BiSpra Hörverstehen 4. Klasse						
BiSpra Wortschatz						
BiSpra Konnektoren						
ELFE						
HRT						
Mathematikaufgaben zu den Bildungsstandards (Aufgabenpool des IQB)						
N	111	126	184	312	132	153

Zum zweiten Messzeitpunkt kamen daher nur noch sechs verschiedene Testhefte zum Einsatz (siehe Tabelle 5). Um die prädiktive Validität bildungssprachlicher Fähigkeiten für schulische Leistungsmaße überprüfen zu können, wurden neben den textbezogenen Hörverstehens-, den Konnektoren- und den bildungssprachlich geprägten Wortschatzitems Aufgaben zum Leseverständnis und zu mathematischen Kompetenzen eingesetzt (Testhefte 1 bis 3). Testheft 4 enthielt gekürzte Fassungen der im Projekt entwickelten Skalen zur Erfassung bildungssprachlicher Fähigkeiten.

Bei der Zuordnung der Testhefte zu den Klassen wurde darauf geachtet, dass alle Klassen diejenigen Testheftversionen erhielten, die sie auch beim ersten Messzeitpunkt bearbeitet hatten. Den Schülerinnen und Schülern, die beim ersten Messzeitpunkt Kurzversionen aller drei BiSpra-Skalen bearbeitet hatten, wurden also auch beim zweiten Messzeitpunkt wieder die Testheftversion vorgelegt, die Kurzversionen aller drei BiSpra-Tests enthielt, um Entwicklungsverläufe über die Zeit abbilden zu können.

Denjenigen Schülerinnen und Schülern, bei denen nach dem ersten Messzeitpunkt noch kein ausgefüllter Elternfragebogen vorlag, wurde am Ende der Testsitzung erneut ein Elternfragebogen mitgegeben.

2.6 Kodierung, Eingabe, Aufbereitung und Analyse der Daten

Die Eingabe der Daten der Pilotstudien erfolgte unter Verwendung des Dateneingabeprogramms WEingabe nach Erstellung einer Eingabemaske.¹⁰ Die Daten der Validierungsstudie wurden im Anschluss an den jeweiligen Messzeitpunkt vom IEA DPC Hamburg eingescannt und anschließend mit Hilfe von Codebüchern (Variablen inkl. Variablen- und Wertelabels) die Datensätze erstellt.

Die Aufbereitung der Daten aller Erhebungen (u. a. Umkodierung der Rohwerte, Mergen der Datensätze, Variablenerstellung etc.) erfolgte in SPSS. Die Datenanalyse erfolgte unter Verwendung der Programme SPSS, ConQuest und Mplus.

2.7 Kommunikation der Projektergebnisse sowie Manuskripterstellung

Die Projektergebnisse wurden auf Fachkonferenzen im In- und Ausland (Kongress der Gesellschaft für Empirische Bildungsforschung (GEBF), Annual Meeting of the American Educational Research Association (AERA), European Congress of Psychology (ECP), Association for Research on Learning and Instruction (EARLI), Conference of Junior Researchers of EARLI (JURE), Jahrestagung der Kommission Grundschulforschung und Pädagogik der Primarstufe der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE), Fachgruppentagung Pädagogische Psychologie der DGPs (PAEPS)) in Form von Vorträgen oder Postern präsentiert (siehe Kap. 4). Zudem nahmen die Projektmitarbeiterinnen an den Veranstaltungen (Herbstschulen, Frühjahrsworkshops) der Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS), in deren Rahmen BiSpra gefördert wurde, teil und präsentierten hier regelmäßig die neuesten Projektergebnisse und stellten diese zur Diskussion. Darüber hinaus beteiligten sich die Projektmitarbeiterinnen an diversen Fortbildungsveranstaltungen in der Bildungspraxis, wo sie die im Projekt entwickelten Aufgaben sowie erste Projektergeb-

¹⁰ Diese Arbeiten wurden von geschulten studentischen Mitarbeiterinnen und Mitarbeitern in Berlin und Bamberg unterstützt. Wir danken diesen für ihr Engagement.

nisse vorstellten und mit den Teilnehmerinnen und Teilnehmern Möglichkeiten der Erfassung und Förderung bildungssprachlicher Fähigkeiten diskutierten (Workshop auf dem Frankfurter Fachforum „Sprache und Naturwissenschaften“, Vortrag im Rahmen des Qualifizierungsprogramms „Alltagsintegrierte Sprachbildung und Sprachförderung für (Kinder-)Tageseinrichtungen im Landkreis Goslar“).

Weiterhin entstanden in Zusammenarbeit beider Teilprojekte sowie gemeinsam mit der Arbeitsgruppe um Angelika Redder zwei Buchkapitel für den neuesten Sammelband der FiSS. Außerdem wurden mehrere Zeitschriftenartikel mit Projektergebnissen veröffentlicht (Zeitschrift für Pädagogische Psychologie, Zeitschrift für Erziehungswissenschaft). Weitere Zeitschriftenartikel befinden sich derzeit in Vorbereitung bzw. in Planung.

3 Wissenschaftlicher Stand, an den angeknüpft wurde

3.1 Angabe bekannter Konstruktionen, Verfahren und Schutzrechte, die für die Durchführung des Vorhabens benutzt wurden

Um Aussagen über die Kriteriums- sowie Konstruktvalidität treffen zu können, wurden in der Validierungsstudie neben den Eigenentwicklungen zum funktional-integrativen Hörverstehen und zu den sprachkomponenten-bezogenen bildungssprachlichen Fähigkeiten (Verständnis von Satzverbindungen mit Konnektoren und bildungssprachlicher Wortschatz) auch bestehende Verfahren zur Erfassung des grundlegenden Grammatikverständnisses (TROG-D; Fox, 2007), des Wortschatzes (PPVT; Version III adaptiert von Bulheller & Häcker, 2003; PPVT-R von Dunn & Dunn, 1981; unveröffentlichte deutsche Forschungsversion von Roßbach, Tietze & Weinert, 2005), des phonologischen Arbeitsgedächtnisses (Zahlen nachsprechen aus K-ABC; Melchers & Preuß, 2009) und der kognitiven Grundfähigkeiten (Subskala des CFT 1; Cattell, Weiß & Osterland, 1997) eingesetzt. Als schulische Leistungsindikatoren dienten das Leseverständnis (ELFE; Lenhard & Schneider, 2006) und mathematische Kompetenzen. Grundlegende arithmetische Fähigkeiten wurden über den HRT 1-4 (Subtest Ergänzungen Haffner, Baro, Parzer & Resch, 2005) erfasst. Zur Bestimmung komplexerer mathematischer Problemlösefähigkeiten dienten Aufgaben aus dem Itempool des IQB, mit denen das Erreichen der Bildungsstandards überprüft werden kann (Granzer, Walther, Winkelmann, Robitzsch & Köller, 2008).

3.2 Angabe der verwendeten Fachliteratur sowie der benutzten Informations- und Dokumentationsdienste

Als Grundlage zur Definition und Charakterisierung von Bildungssprache wurde im BiSpraprojekt auf die einschlägige nationale und internationale Fachliteratur zu Bildungssprache bzw. *academic language* zurückgegriffen. Aus dem angloamerikanischen Sprachraum zählen hierzu neben theoretischen Überlegungen von Cummins (z. B. 1979, 2000) zur Unterscheidung zwischen *Basic Interpersonal Communication Skills* (BICS) und *Cognitive Academic Language Proficiency* (CALP) und Forschungsarbeiten aus dem Bereich der funktionalen Linguistik zum Begriff des sprachlichen Registers (z. B. Halliday & Hasan, 1989; Schleppegrell, 2001, 2004) insbesondere Ansätze zur evidenzbasierten Operationalisierung von Bildungssprache, die von der Arbeitsgruppe um Alison Bailey am Center for Research on Eval-

uation, Standards, and Student Testing erarbeitet wurden (Bailey, Butler, Stevens & Lord, 2007; Bailey, Stevens, Butler & Huang, 2005; Butler, Bailey, Stevens & Huang, 2004). Im deutschen Sprachraum sind insbesondere die Arbeiten von Ingrid Gogolin relevant (z. B. Gogolin, 2006, 2009; Gogolin et al., 2011). Der von ihr vorgeschlagene Begriff der *Bildungssprache* als Äquivalent zu *academic language* hat sich mittlerweile fest etabliert und bezeichnet im weitesten Sinne diejenige Sprache, die in der Schule verwendet und deren Verständnis vorausgesetzt wird, um Wissen zu erwerben und weiterzugeben (vgl. z. B. Anstrom, DiCerbo, Butler, Katz, Millet & Rivera, 2010; Chamot & O'Malley, 1994).

Mit Blick auf die Aufgabenentwicklung wurden zum einen Befunde und Materialien herangezogen, die (1) als Grundlage für eine möglichst valide Operationalisierung lexikalischer und grammatischer Merkmale von Bildungssprache dienen konnten. Zum anderen wurden (2) auch Forschungsarbeiten berücksichtigt, in denen bereits Vorarbeiten zur Erfassung bildungssprachlicher Fähigkeiten geleistet worden sind.

Der erst genannte Punkt umfasst nicht nur Literatur zu allgemeinen Charakterisierungen der Bildungssprache im Deutschen (z. B. Eckhardt, 2008; Gogolin & Lange, 2011) und im Englischen (z. B. Bailey et al., 2007; Schleppegrell, 2004); darüber hinaus wurden auch Quellen herangezogen, die eine möglichst empirisch fundierte Auswahl schulrelevanter sprachlicher Merkmale ermöglichen sollten. Hierzu zählen etwa Testaufgaben zur Operationalisierung der in Deutschland länderübergreifend geltenden Bildungsstandards (vgl. Kultusministerkonferenz, 2005; Stanat, Pant, Böhme & Richter, 2012) sowie die Korpora DLex (Heister et al., 2011) und ChildLex3 (Schroeder, Würzner, Heister, Geyken & Kliegl, 2015). Während das Korpus DLex hauptsächlich auf schriftlichen Quellen basiert, die sich an Erwachsene richten (Zeitungsartikel, Belletristik, Wissenschaft, Gebrauchstexte), umfasst ChildLex vorrangig nicht schulische schriftliche Texte für 6- bis 12-jährige Kinder. Bei der Entwicklung von Aufgaben zum Verständnis von Satzverbindungen mit Konnektoren wurden in erster Linie solche Konnektoren ausgewählt, die in DLex häufiger auftraten als in ChildLex und für die daher vermutet werden kann, dass sie vor allem in einer anspruchsvollen formellen Sprache von Bedeutung sind (für eine ausführliche Beschreibung der Aufgabenentwicklung siehe Schuth et al., 2015). Um eine empirisch fundierte Auswahl allgemein bildungssprachlicher Begriffe zu gewährleisten, wurde ferner auf das oben beschriebene Hamburg-Bamberger-Korpus von BiSpra II (Köhne et al., 2015) und damit auf Daten/Befunde aus verschiedenen Forschungsprojekten, wie z. B. der Bamberger Forschergruppe „Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen“ (entwicklungspsychologisches Teilprojekt), dem FiSS-Projekt MüWi (Leitung: Angelika Redder) sowie BiSpra I rekuriert.

Überdies wurden bei der Aufgabenentwicklung in BiSpra auch Studien zur Operationalisierung von Bildungssprache berücksichtigt. In diesem Zusammenhang sind insbesondere die Arbeiten von Eckhardt (2008) zu nennen. Die Autorin operationalisierte erstmals im deutschsprachigen Raum verschiedene Aspekte von Bildungssprache und untersuchte deren Effekte auf das Lese- und Hörverstehen von Kindern mit deutscher und Kindern mit nicht-deutscher Herkunftssprache. In Anlehnung an das von ihr gewählte Vorgehen sowie an das Vorgehen in der BiKS-3-10-Studie (entwicklungspsychologisches Teilprojekt) wurden in BiSpra I textbezogene Hörverstehensaufgaben entwickelt, deren Stimulustexte sich systematisch hinsichtlich verschiedener lexikalischer und grammatischer Merkmale der Bildungssprache unterscheiden. Nicht zuletzt baute die Aufgabenentwicklung in BiSpra II auch auf den Erkennt-

nissen aus BiSpra I auf, die bereits an verschiedenen Stellen publiziert wurden (Dragon et al., 2015; Heppt, Dragon, Berendes, Stanat, & Weinert, 2012; Heppt et al., 2014).

4 Zusammenarbeit mit anderen Stellen

Aktivitäten im Rahmen der Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS). Das BiSpra-Verbundprojekt war konzipiert und wurde durchgeführt in enger Zusammenarbeit zwischen verschiedenen Disziplinen (Psychologie, Erziehungswissenschaft, Linguistik) und den beiden Teilprojekten in Bamberg und Berlin. Die enge interdisziplinäre Zusammenarbeit wird in allen Schritten der Projektentwicklung, der Aufgabenkonstruktion und –überprüfung sowie der Durchführung und Auswertung der Validierungsstudie deutlich. Ein kooperativer und stetiger Austausch zwischen den Mitarbeiterinnen beider Teilprojekte sowie verschiedene Arbeitstreffen und Skype-Konferenzen der Teilprojekte auf Leitungs- und Mitarbeitererebene haben hierzu beigetragen.

Darüber hinaus erwies sich die Zusammenarbeit mit anderen Projekten der Forschungsinitiative als hilfreich und anregend. Dies gilt nicht zuletzt für die verschiedenen Veranstaltungen der Forschungsinitiative Sprachdiagnostik und Sprachförderung (Herbstschulen, Frühjahrsworkshops), in denen der jeweilige Stand der Projektarbeit durch Kurzvorträge vorgestellt und diskutiert wurde, und schlägt sich auch in gemeinsamen Publikationen (Knoepke et al., 2016) sowie in Veröffentlichungen der Forschungsinitiative (FiSS Sammelband) nieder, in dem das BiSpra II-Projekt mit zwei Beiträgen vertreten ist.

Die Arbeit im Verbund ermöglichte nicht nur eine äußerst enge Kooperation zwischen dem Bamberger und dem Berliner Teilprojekt im Hinblick auf Aufgabenentwicklung, Datenerhebung und Datenanalyse, sondern auch den wechselseitigen Austausch zwischen den psychologisch-erziehungswissenschaftlichen Teilprojekten und der Arbeitsgruppe von Angelika Redder.

II Eingehende Darstellung

1 Ergebnisse¹¹

1.1 Ergebnisse der Pilotierungsstudien¹²

Die Items der drei Skalen (textbezogenes Hörverstehen, Satzverbindungen mit Konnektoren, bildungssprachlicher Wortschatz) wurden jeweils separat in einem mehrschrittigen Verfahren mittels Rasch-Analysen (Embretson & Reise, 2000) unter Verwendung des Programms ConQuest (Wu, Adams, Wilson & Haldane, 2007) skaliert. Zur Itemauswahl wurden nach jedem Analyseschritt verschiedene itemstatistische Kennwerte herangezogen. Erstens dienten Maße des Modellfits als Kriterium (in Anlehnung an PISA (OECD, 2005): standardisierter *Mean Square (MNSQ)* zwischen 0.75 und 1.25 sowie daraus abgeleitete *t*-Wert kleiner als 1.96, nicht signifikant). Zweitens wurden die punkt-biserialen Korrelationen (r_{pb}) zwischen Item und Gesamtskala (vergleichbar mit der klassischen Trennschärfe) als Indikatoren für die Diskriminationsfähigkeit der Items herangezogen ($r_{pb} > .25$; vgl. OECD, 2005). Drittens wurde *EAP/PV-Reliabilität (expected a posteriori/plausible value)* als Maß für die interne Konsistenz des Gesamtitemsatzes einbezogen (Werte nahe 1 sprechen für eine hohe interne Konsistenz).

1.1.1 Itemauswahl für die textbezogenen Hörverstehensaufgaben

Anhand der beschriebenen Auswahlkriterien haben sich 13 der 16 textbezogenen Hörverstehensaufgaben mit insgesamt 78 Items bewährt. Diese wiesen sowohl bezogen auf die Gesamtstichprobe als auch bei getrennter Betrachtung der beiden Klassenstufen (Klasse 3; Klasse 4) und der beiden Sprachgruppen (monolingual deutschsprachig und nicht-monolingual deutschsprachig) einen sehr guten Itemfit ($0.93 \leq M_{MNSQ} \leq 1.02$, $0.11 \leq SD_{MNSQ} \leq 0.16$), hohe interne Konsistenzen (zwischen .86 und .89 für die verschiedenen Gruppen) und gute Diskriminierungswerte auf. Letztere variierten in den verschiedenen Sprachgruppen und Klassenstufen zwischen $.03 \leq r_{pb} \leq .79$, wobei aber nur zwischen 5 und 15% der Items eine Trennschärfe von unter $r_{pb} = .25$ aufwiesen. Die mittlere Itemschwierigkeit lag für die Gesamtstichprobe bei $M = .73$ ($SD = .10$) und unterschied sich zwischen Schülerinnen und Schülern der Klassenstufen 3 und 4 sowie zwischen monolingual deutschsprachigen Kindern (md) und nicht-monolingual deutschsprachigen Kindern (nmd) in erwartbarer Richtung ($M_{K3} = .67$, $SD_{K3} = .11$; $M_{K4} = .78$, $SD_{K4} = .11$; $M_{md} = .83$, $SD_{md} = .10$; $M_{nmd} = .68$, $SD_{nmd} = .11$). Ausgehend von dem resultierenden Pool von 13 Texten und 78 Items wurden drei Aufgabensets für die Klassenstufen 2 bis 4 zusammengestellt. Diese wurden so gebildet, dass sie einerseits einen substanziellen Anteil identischer Aufgaben (Ankeritems) enthalten, um Entwicklungsverläufe abbilden zu können; andererseits wurden zu hohe Überlappungen vermieden,

¹¹ Der Schwerpunkt des Berliner Teilprojekts lag auf der Auswertung der Aufgaben zum textbezogenen Hörverstehen; der Schwerpunkt des Bamberger Teilprojekts auf den sprachkomponenten-bezogenen Aufgaben.

¹² Die Kapitel 1.1 und 1.2 des Ergebnisteils folgen in wesentlichen Teilen dem Beitrag von Schuth, E., Heppt, B., Köhne, J., Weinert, S. & Stanat, P. (2015). Die Erfassung schulisch relevanter Sprachkompetenzen bei Grundschulkindern – Entwicklung eines Testinstruments. In A. Redder, J. Naumann & R. Tracy (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS) – Ergebnisse* (S. 73-92). Münster: Waxmann.

um Erinnerungseffekte gering zu halten und ein auf verschiedene Klassenstufen angepasstes Schwierigkeitsniveau zu erreichen (entsprechend waren die durchschnittlichen Lösungswahrscheinlichkeiten für die zweite ($M = .70$, $SD = .16$) und dritte Klasse ($M = .70$, $SD = .15$) vergleichbar).

1.1.2 Itemauswahl für die Aufgaben zum Verständnis von Satzverbindungen mit Konnektoren

Bezüglich der Satzverbindungsaufgaben wurden, wie beschrieben, ein bildgestütztes und ein Lückensatzformat erprobt. Es zeigte sich, dass für die Lückensätze sowohl die Fit-Parameter und die Diskriminationsfähigkeiten (siehe Tabelle 6) als auch die internen Konsistenzen deutlich besser ausfielen als für das bildgestützte Format (EAP/PV_{bildgestütztesFormat} zwischen .51 und .67, EAP/PV_{Lückensatz} zwischen .73 und .76). Vor diesem Hintergrund wurde das Lückensatzformat für die Validierungsstudie ausgewählt. Im Speziellen mussten von den Lückensatz-Items nur zwei der 24 eingesetzten Items aufgrund ihrer itemstatistischen Kennwerte ausgeschlossen werden.

Tabelle 6

Range der Fit-Indizes, punkt-biserialen Korrelationen und internen Konsistenzen der Skalen zum Verständnis von Satzverbindungen (Konnektoren)

	Lückensätze	bildgestütztes Format
M_{MNSQ}^{13}	$0.99 \leq M_{MNSQ} \leq 1.00$	$0.99 \leq M_{MNSQ} \leq 1.00$
SD_{MNSQ}	$0.14 \leq SD_{MNSQ} \leq 0.17$	$.07 \leq SD_{MNSQ} \leq .09$
M_t	$-0.08 \leq M_t \leq 0.05$	$-0.03 \leq M_t \leq 0.04$
SD_t	$0.69 \leq SD_t \leq 1.35$	$0.56 \leq SD_t \leq 0.82$
$\min \leq r_{pb} \leq \max$ (über Gruppen ¹⁴ hinweg)	$.04 \leq r_{pb} \leq .75$	$.06 \leq r_{pb} \leq .72$
% Items mit r_{pb} unter .25 innerhalb der Gruppen (min-max)	0-8	8-16

Die ausgeschlossenen Items, die die beiden konzessiven Konnektoren *wenngleich* und *obgleich* beinhalteten, erwiesen sich sowohl für mehrsprachige als auch für monolingual deutschsprachige Kinder als äußerst schwierig: Die Lösungswahrscheinlichkeit lag in beiden Gruppen unterhalb der Ratewahrscheinlichkeit von 25% (4 Wahlmöglichkeiten). Auch die

¹³ $A \leq M/SD_x \leq B$ = Mittelwerte/Standardabweichungen von Maß X (MNSQ, t-Wert) für die fünf verschiedenen Gruppen liegen zwischen Werten A und B, jeweils gemittelt über die separaten Skalierungen. Gruppen: Gesamtstichprobe, monolingual deutschsprachige vs. nicht-monolingual deutschsprachige Kinder, Klassenstufen 2 vs. 4 (Satzverbindungen im bildgestützten Format) bzw. Klassenstufen 3 vs. 4 (Satzverbindungen im Lückensatzformat).

¹⁴ Gesamtstichprobe, monolingual deutschsprachige vs. nicht-monolingual deutschsprachige Kinder, Klassenstufen 2 vs. 4 (Satzverbindungen im bildgestützten Format) bzw. die Klassenstufen 3 vs. 4 (Satzverbindungen im Lückensatzformat).

Diskriminationsparameter sind im Vergleich zu den anderen Satzverbindungsitems niedrig (*wenngleich* in Klasse K3 $r_{pb.} = .34$ und $r_{pb.} = .28$ in Klasse K4; *obgleich* $r_{pb.} = .20$ K3 und $r_{pb.} = .27$ K4). Auf Basis dieser Befunde wurden 22 Satzverbindungs-Lückensatzitems für den Einsatz in der zweiten und dritten Klasse in der Validierungsstudie ausgewählt. Die mittlere Itemschwierigkeiten dieser Auswahl beträgt für die Gesamtstichprobe $M = .55$ ($SD = .47$). Ein Vergleich der mittleren Schwierigkeiten zwischen den beiden Klassenstufen sowie zwischen den beiden Sprachgruppen zeigt Unterschiede in der erwarteten Richtung ($M_{K3} = .54$, $SD_{K3} = .16$; $M_{K4} = .70$, $SD_{K4} = .15$; $M_{md} = .74$, $SD_{md} = .15$; $M_{nmd} = .57$, $SD_{nmd} = .16$).

1.1.3 Itemauswahl für die Aufgaben zum bildungssprachlichen Wortschatz

Für die Aufgaben zum bildungssprachlichen Wortschatz wurden ebenfalls ein bildgestütztes Format und ein Lückensatzformat erprobt. Auch hier zeigten sich sowohl für die Gesamtstichprobe als auch getrennt für einzelne Subgruppen (2. und 4. Jahrgangsstufe; monolingual deutsche und nicht-monolingual deutsche Familiensprache) für das Lückensatzformat bessere interne Konsistenzen ($EAP/PV_{\text{bildgestütztesFormat}}$ zwischen $.51$ und $.75$, $EAP/PV_{\text{Lückensatz}}$ zwischen $.58$ und $.82$) und bessere itemstatistische Kennwerte als für das bildgestützte Format (s. Tabelle 7).

Die Diskriminationsparameter des bildgestützten Formats weisen auf eine vergleichsweise geringe Trennschärfe der Items hin. Bei einer Berücksichtigung all jener Diskriminationsparameter, die sich bei Betrachtung der Gesamtstichprobe und bei getrennten Analysen für die beiden Jahrgangsstufen und Sprachgruppen ergeben, lagen 32 - 50% der 34 Items unter einem Diskriminationswert von $r_{pb} = .25$. Das Lückensatzformat zeigte deutlich bessere Diskriminationsparameter: Hier lagen nur 10 - 34% der 29 Items in den verschiedenen Gruppen unter $r_{pb} = .25$ (s. Tabelle 7).

Bei einem Vergleich der mittleren Schwierigkeiten zeigten sich für das bildgestützte Format erwartungskonforme Unterschiede zwischen den Gruppen ($M_{\text{Gesamt}} = .70$, $SD_{\text{Gesamt}} = .42$; $M_{K2} = .64$, $SD_{K2} = .44$; $M_{K4} = .76$, $SD_{K4} = .38$; $M_{md} = .75$, $SD_{md} = .40$; $M_{nmd} = .66$, $SD_{nmd} = .43$). Jedoch erwies sich ein Teil der Items mit einer Schwierigkeit von $M > .85$ als sehr leicht. Dies galt teilweise sogar für die Gesamtgruppe, insbesondere aber für Kinder der 4. Jahrgangsstufe sowie für monolingual deutschsprachige Kinder und könnte auch die geringe Diskriminationsfähigkeit der Items erklären. Das bildgestützte Format wurde daher nicht in der Validierungsstudie eingesetzt.

Die Lückensatzaufgaben erzielten weitaus bessere Itemfitstatistiken. Die mittleren Itemschwierigkeiten lagen bei $M_{\text{Gesamt}} = .55$, $SD_{\text{Gesamt}} = .47$; $M_{K2} = .45$, $SD_{K2} = .47$; $M_{K4} = .68$, $SD_{K4} = .43$; $M_{md} = .61$, $SD_{md} = .45$; $M_{nmd} = .51$, $SD_{nmd} = .48$. Auch hier zeigt sich erwartungsgemäß, dass Kinder der 4. Klasse und die Gruppe der monolingual deutschsprachigen Kinder etwas besser in den Wortschatz-Lückensatzaufgaben abschnitten als Kinder der 2. Klasse und nicht monolingual deutschsprachige Kinder. Um die Schwierigkeit der Items der Klassenstufe anzupassen, wurden für die Validierungsstudie zwei Testversionen mit je 24 Items, von denen jeweils 19 überlappten (Ankeritems), erstellt: eine für die 2. Jahrgangsstufe und eine für die 3./4. Jahrgangsstufe.

Tabelle 7

Range der Fit-Indizes, punkt-biserialen Korrelationen und internen Konsistenzen der beiden Skalen zum bildungssprachlichen Wortschatz

	Lückensätze	bildgestütztes Format
M_{MNSQ} ¹⁵	$0.98 \leq M_{MNSQ} \leq 1.01$	$0.99 \leq M_{MNSQ} \leq 1.00$
SD_{MNSQ}	$0.05 \leq SD_{MNSQ} \leq 0.12$	$0.03 \leq SD_{MNSQ} \leq 0.10$
M_t	$-0.03 \leq M_t \leq 0.15$	$0.02 \leq M_t \leq 0.09$
SD_t	$0.78 \leq SD_t \leq 1.55$	$0.37 \leq SD_t \leq 1.02$
$\min \leq r_{pb} \leq \max$ (über Gruppen ¹⁶ hinweg)	$0.08 \leq r_{pb} \leq 0.65$	$0.00 \leq r_{pb} \leq 0.67$
% Items mit r_{pb} unter .25 innerhalb der Gruppen (min-max)	10-34	32-50

1.2 Ergebnisse der Validierungsstudie

1.2.1 Validitätsprüfung der Hörverstehensaufgaben, der Aufgaben zum Verständnis von Satzverbindungen mit Konnektoren sowie der Aufgaben zum bildungssprachlichen Wortschatz

Auf Grundlage der Daten des 1. Messzeitpunkts der Validierungsstudie überprüfen wir zunächst die internen Konsistenzen für die Skalen zur Erfassung des textbezogenen Hörverstehens, des Verständnisses von Satzverbindungen (Konnektoren) und des bildungssprachlichen Wortschatzes für die Klassenstufen 2 und 3 sowie für monolinguale und nicht-monolinguale Schülerinnen und Schüler. Anschließend wurden Leistungsunterschiede zwischen Klassenstufen einerseits und Kindern mit monolingualem oder nicht-monolingualem Sprachhintergrund andererseits, für die die Instrumente geeignet sein sollten, untersucht. Um Hinweise auf die konvergente und diskriminante Validität zu erhalten, wurden zudem auf Basis eines mehrdimensionalen Modells in ConQuest latente Korrelationen mit den Variablen textbezogenes Hörverstehen, Verständnis von Satzverbindungen, bildungssprachlicher Wortschatz, allgemeiner Wortschatz, Grammatikverständnis und kognitive Grundfähigkeiten geschätzt (für eine ausführliche Ergebnisdarstellung siehe Schuth et al., 2015).

Die internen Konsistenzen variierten für das textbezogene Hörverstehen und die Satzbindungsaufgaben für monolingual deutschsprachige Kinder zwischen .61 und .88 und liegen damit alle in einem akzeptablen Bereich (etwas geringer für nicht-monolingual deutschsprachige 2. Klässler). Demnach werden das globale Verständnis bildungssprachlicher Texte und das Verständnis von Satzverbindungen (Konnektoren) mit den vorliegenden Instrumenten insbesondere in der Klassenstufen 3 sowie generell bei Kindern mit monolingual deutscher

¹⁵ $A \leq M/SD_x \leq B$ = Mittelwerte/Standardabweichungen von Maß X (MNSQ, t-Wert) für die fünf verschiedenen Gruppen liegen zwischen Werten A und B, jeweils gemittelt über die separaten Skalierungen. Gruppen: Gesamtstichprobe, monolingual deutschsprachige vs. nicht-monolingual deutschsprachige Kinder, Klassenstufen 2 vs. 4.

¹⁶ Gesamtstichprobe, monolingual deutschsprachige vs. nicht-monolingual deutschsprachige Kinder, Klassenstufen 2 vs. 4.

Familiensprache hinreichend verlässlich erfasst. Die internen Konsistenzen der Wortschatzskala liegen in der 2. Klassenstufe bei .50 für die monolingual deutschsprachigen Kinder und bei .44 für mehrsprachige Kinder und sind damit für beide Sprachgruppen nicht zufriedenstellend. In der 3. Jahrgangsstufe hingegen erreichen die Reliabilitäten etwas bessere Werte ($EAP/PV_{md} = .69$, $EAP/PV_{nmd} = .61$), was darauf hindeutet, dass die Aufgaben mit zunehmender Klassenstufe eine reliablere Fähigkeitsschätzung ermöglichen.

Zur Überprüfung von Gruppenunterschieden wurden die drei Skalen zunächst auf Basis eines eindimensionalen Raschmodells getrennt skaliert. Dabei wurde die mittlere Personenfähigkeit für die Gesamtstichprobe in allen drei Tests auf 0 restringiert und die mittleren Personenfähigkeiten für die Substichproben jeweils als Abweichungen vom Mittelwert 0 der Gesamtstichprobe bestimmt (s. Tabelle 8).

Dabei zeigte sich erwartungsgemäß, dass die mittleren Fähigkeitsschätzer der monolingual deutschsprachigen Kinder bei allen drei Skalen über den Leistungen der Schülerinnen und Schüler mit nicht-monolingual deutscher Familiensprache liegen (textbezogenes Hörverstehen: $t = 5.20$, $df = 444.98$, $p < .01$, $d = 0.45$; Verständnis von Satzverbindungen: $t = 3.48$, $df = 587$, $p < .01$, $d = 0.29$; bildungssprachlicher Wortschatz $t = 6.93$, $df = 421$, $p < .01$, $d = 0.69$). Ebenso erwartungskonform sind die höheren Fähigkeitsschätzer für Schülerinnen und Schüler der Jahrgangsstufe 3 im Vergleich zu denjenigen für Schülerinnen und Schüler der Jahrgangsstufe 2 (textbezogenes Hörverstehen: $t = -8.23$, $df = 565$, $p < .01$, $d = 0.70$; Verständnis von Satzverbindungen: $t = -11.78$, $df = 589$, $p < .001$, $d = 0.98$; bildungssprachlicher Wortschatz $t = -7.37$, $df = 422$, $p < .01$, $d = 0.72$). Die Effektgrößen deuten zudem darauf hin, dass Unterschiede in den bildungssprachlichen Fähigkeiten beim textbezogenen Hörverstehen und beim Verständnis von Satzverbindungen (Konnektoren) offenbar in stärkerem Maße von der Klassenstufe abhängen als vom familiären Sprachhintergrund. Für den bildungssprachlichen Wortschatz zeigten sich für die Faktoren Klassenstufe und Sprachhintergrund vergleichbare Effekte.

Die latenten Korrelationen zwischen den drei Skalen zur Erfassung von bildungssprachlichen Kompetenzen und den zum selben Messzeitpunkt eingesetzten Tests zur Messung von allgemeinem Wortschatz, rezeptiver Grammatik sowie den Matrizenaufgaben als Indikator für die nonverbalen kognitiven Grundfähigkeiten sind in Tabelle 9 dargestellt. Erwartungsgemäß zeigt sich, dass die sprachbezogenen Tests untereinander deutlich höher korrelieren als mit den kognitiven Grundfähigkeiten. Dies kann als erster Hinweis auf das Vorliegen konvergenter und diskriminanter Validität interpretiert werden. Die numerisch höchste Korrelation besteht zwischen den beiden Wortschatzskalen. Da die BiSpra-Aufgaben zum Verständnis von Konnektoren nicht nur grammatische Kompetenzen messen, sondern sowohl grammatisches als auch Wortverständnis fordert, lassen sich sowohl mit dem rezeptiven allgemeinen Wortschatz als auch mit dem Grammatikverständnis substantielle Korrelationen erwarten, die durch die Daten bestätigt werden.

Tabelle 8

Deskriptive Statistiken für die geschätzten Personenfähigkeiten (WLE) für Hörverstehens-, Satzverbindungs- und Wortschatzaufgaben

Klasse	n	Sprachhintergrund (monolingual – nicht monolingual Deutsch)	Personenfähigkeit (WLE)	
			M	SD
Hörverstehen				
2	125	monolingual deutsch	-0.54	0.66
	151	nicht-monolingual deutsch	-1.11	1.24
	277⁴	Gesamt	-0.35	0.99
3	131	monolingual deutsch	0.55	0.94
	159	nicht-monolingual deutsch	0.11	0.81
	290	Gesamt	0.31	0.89
2 und 3				
567	Gesamt	-0.01	1.00	
Satzverbindungen				
2	125	monolingual deutsch	-0.35	1.18
	147	nicht-monolingual deutsch	-0.72	.88
	274⁴	Gesamt	-0.56	1.05
3	122	monolingual dt.	0.79	1.16
	195	nicht-monolingual dt.	0.31	1.07
	317	Gesamt	0.50	1.13
2 und 3				
591	Gesamt	0.01	1.21	
Wortschatz				
2	73	monolingual dt.	0.06	0.77
	124	nicht-monolingual dt.	-0.50	0.67
	197	Gesamt	-0.30	0.75
3	95	monolingual dt.	0.64	1.01
	132	nicht-monolingual dt.	0.08	0.77
	227	Gesamt	0.31	0.92
2 und 3				
424	Gesamt	0.03	0.90	

Tabelle 9

Latente Interkorrelationen der Maße zur Erfassung bildungssprachlicher Fähigkeiten und Korrelationen mit Validierungsvariablen (N = 1105)

		1	2	3	4	5
1	bildungssprachliches Hörverstehen	–				
2	Satzverbindungen	.74	–			
3	bildungssprachlicher Wortschatz	.83	.70	-		
4	Grammatik	.71	.77	.54	-	
5	allgemeiner Wortschatz	.83	.81	.89	.70	-
6	kognitive Grundfähigkeiten	.35	.57	.27	.51	.41

1.2.2 Erste Analysen zur prädiktiven Validität

Für jede der drei bildungssprachlichen Skalen wurde untersucht, inwieweit die Leistungen der Kinder zu Messzeitpunkt 1 (2. und 3. Klasse) prädiktiv für die Schulleistungen der Kinder ein Jahr später sind (3. Klasse und 4. Klasse). Als Indikatoren für die Schulleistungen wurden tatsächliche Schulnoten bzw. (sofern keine Noten vorhanden waren) Lehrereinschätzungen der Schülerleistungen auf der Notenskala sowie schulrelevante Kompetenzmaße (Leseverständnis [ELFE], arithmetische Fertigkeiten [HRT], mathematische Testaufgaben des IQB zu den Bildungsstandards der KMK [IQB-Aufgaben]; siehe Tabelle 3) herangezogen. Hierzu wurden für die beiden Messwiederholungsstudien getrennte hierarchische Regressionen für vier Noten (Mathematik, Sachunterricht, Lesen, Schreiben) und drei Kompetenzmaße (ELFE, HRT, IQB-Aufgaben) als abhängige Variablen berechnet.

Aufgrund des komplexen Erhebungsdesigns der Validierungsstudie muss ein quer- und längsschnittlich angelegtes Linkingverfahren durchgeführt werden, um die Stichproben und Kompetenzindikatoren der einzelnen Testheftversionen (auch längsschnittlich) zusammenzuführen und vergleichen zu können. Dieses Verfahren ist methodisch anspruchsvoll, aktuell noch in Arbeit und wird im Rahmen des Anschlussprojekts „BiSpra-Aufgaben“ (FKZ: 01 GJ 1704 A/B) weitergeführt werden. Die im Folgenden zusammenfassend berichteten Ergebnisse basieren derzeit noch auf vergleichsweise kleinen Substichproben.

Prädiktivität des textbezogenen bildungssprachlichen Hörverstehens. Um die Vorhersagekraft der Hörverstehensaufgaben für Schulnoten und schulrelevante Kompetenzen ein Jahr später zu untersuchen, wurden hierarchische Regressionen durchgeführt. Dabei wurden das Alter, das Geschlecht, der Sprachhintergrund (Familiensprache: nur Deutsch vs. auch eine andere Sprache als Deutsch) und der sozio-ökonomische Status (SES), erfasst über den HISEI der Eltern, als Kontrollvariablen in das Modell aufgenommen. Im Folgenden werden die gefundenen Koeffizienten für die Hörverstehensskala exemplarisch vor allem für den Längsschnitt von der zweiten zur dritten Klassenstufe berichtet.

Das integrative bildungssprachbezogene Hörverstehensmaß zeigte in der zweiten Klasse einen signifikanten Effekt auf die in der dritten Klasse erhobenen späteren Noten im Lesen ($\beta = -.40, p < .001$), Schreiben ($\beta = -.24, p < .05$), in Mathematik ($\beta = -.24, p < .05$) und im Sachunterricht ($\beta = -.19, p < .05$). Zudem sagten die Fähigkeiten der Kinder im bildungs-

sprachlichen Hörverstehen ihre Leistungen beim Lösen der mathematischen IQB-Aufgaben ein Jahr später vorher ($\beta = .41, p < .01$), nicht aber ihre arithmetischen Fertigkeiten. Für das getestete Leseverständnis erwiesen sich die Leistungen der Kinder im Hörverstehen sogar als einziger signifikanter Prädiktor im Modell ($\beta = -.52, p < .001$). Diese Befunde weisen darauf hin, dass die bildungssprachliche Hörverstehensskala Kompetenzen erfasst, die für die schulische Leistungsentwicklung relevant sind, was für die Validität des Verfahrens spricht. Auch die Ergebnisse für den Längsschnitt von der dritten zur vierten Klassenstufe ergeben entsprechende Hinweise (z. B. signifikante prädiktive Zusammenhänge mit allen erhobenen Noten: im Sachunterricht, $\beta = -.34, p < .001$; Rechnen, $\beta = -.18, p < .05$; Schreiben, $\beta = -.39, p < .001$; Lesen, $\beta = -.32, p < .001$; sowie der späteren Kompetenz der Kinder beim Lösen der mathematischen IQB-Aufgaben, $\beta = .27, p < .05$).

Prädiktivität des Verständnisses von Satzverbindungen mit Konnektoren. Die Vorhersagekraft der Aufgaben zum Verständnis von Satzverbindungen mit Konnektoren auf die Schulleistungen der Kinder ein Jahr später wurde ebenfalls mittels hierarchischer Regressionen und unter Kontrolle des Alters, des Geschlechts, des Sprachhintergrunds und des HISEI der Eltern geprüft. In einem zweiten Schritt wurden allgemeine grammatische Kompetenzen, erfasst über 34 Items des TROG-D (Fox, 2007), und in einem abschließenden dritten Schritt die BiSpra-Satzverbindungsaufgaben aufgenommen. Geprüft wird somit, inwieweit die BiSpra-Skala zusätzlich zum TROG-D, der nicht speziell auf die Erfassung bildungssprachlicher Kompetenzen ausgerichtet ist, Varianz in den erzielten Schulleistungen erklärt. Auch hier werden die Ergebnisse des dritten Schritts exemplarisch primär für den Längsschnitt von der zweiten zur dritten Klassenstufe berichtet.

In der zweiten Klassenstufe erweist sich das Verständnis von Satzverbindungen – selbst unter Kontrolle der allgemeinen grammatischen Kompetenzen – als signifikanter Prädiktor sowohl für die Noten im Lesen ($\beta = -.37, p < .01$) und Schreiben ($\beta = -.30, p < .05$) als auch für die Note im Sachunterricht ($\beta = -.20, p = .08$). Bezogen auf die Noten in Mathematik erwies sich in der zweiten Klasse nur das allgemeine Grammatikverständnis als signifikanter Prädiktor ($\beta = -.32, p < .05$). Betrachtet man statt Noten die erhobenen schulrelevanten Kompetenzmaße als abhängige Variablen, so zeigt sich ein signifikanter Effekt der Satzverbindungsaufgaben auf die Leistungen im Leseverständnis ($\beta = .45, p < .001$). Darüber hinaus trägt das Verständnis von Konnektoren auch zur Vorhersage der Kompetenz in den mathematischen IQB-Aufgaben bei ($\beta = .35, p < .01$). Die arithmetischen Fertigkeiten der Kinder konnten hingegen durch keine der untersuchten Variablen vorhergesagt werden. Im Unterschied dazu erweisen sich in der dritten Klasse sowohl die allgemeinen grammatischen Kompetenzen ($\beta = .22, p < .05$) als auch das bildungssprachliche Konnektorenverständnis ($\beta = .37, p < .001$) als signifikante Prädiktoren späterer arithmetischer Fertigkeiten, wobei der Effekt des bildungssprachlichen Maßes deutlich stärker ist. Zudem zeigten sich für die dritte Klasse Effekte des Konnektorenverständnisses auf alle Schulnoten ein Jahr später (Lesen, $\beta = -.38, p < .001$; Schreiben, $\beta = -.31, p < .01$; Mathematik, $\beta = -.37, p < .01$; Sachunterricht, $\beta = -.25, p < .05$) sowie auf die Leistungen im Leseverständnistest ($\beta = .45, p < .001$).

Prädiktivität des bildungssprachlichen Wortschatzwissens. Für die Aufgaben zur Erfassung des bildungssprachlichen Wortschatzes wurden analoge Berechnungen zur Überprüfung des prädiktiven Effekts auf Schulnoten und Testleistungen über ein Jahr hinweg durchgeführt. Dabei gingen neben den Kontrollvariablen im zweiten Schritt das allgemeine Wortschatzwissen, erfasst über 40 Items des Peabody Picture Vocabulary Tests (PPVT; adaptiert aus

den Versionen III von Bulheller & Häcker (2003) und einer deutschen Forschungsversion des PPVT-R (Dunn & Dunn, 1981) von Roßbach, Tietze & Weinert, 2005) und im dritten Schritt die Skala zum bildungssprachlichen Wortschatz aus BiSpra in das Modell ein.

Die Ergebnisse des dritten Schritts zeigen für die zweite Klassenstufe, dass die Leistung in der bildungssprachlichen Wortschatzskala die Noten im Lesen ($\beta = -.60, p < .01$) ein Jahr später signifikant vorhersagt. Vergleichbare Effekte für die Noten in Mathematik, im Sachunterricht und im Schreiben finden sich nicht. Allerdings zeigt sich bereits für den allgemeinen Wortschatzindikator (der zuerst in das Modell aufgenommen wurde) ein signifikanter Effekt auf die späteren Noten in Mathematik ($\beta = -.51, p < .05$) und im Sachunterricht ($\beta = -.58, p < .01$). Die Testleistungen der Kinder im Leseverständnis, in den arithmetischen Mathematikaufgaben und den mathematischen IQB-Aufgaben ließen sich durch keinen der beiden Wortschatztest bedeutsam vorhersagen.

Die Ergebnisse für den Längsschnitt von der dritten zur vierten Klassenstufe ergaben allerdings ein etwas anderes Muster: Hier hatte der bildungssprachliche Wortschatz einen signifikanten Effekt auf die Noten der Kinder in Mathematik ($\beta = -.64, p < .01$) und im Sachunterricht ($\beta = -.58, p < .05$). Auch für die Testleistungen in den mathematischen IQB-Aufgaben erweist sich der bildungssprachliche Wortschatz als signifikanter Prädiktor ($\beta = .52, p < .01$); der Effekt des allgemeinen Wortschatzes auf diese Leistungen verschwand bei zusätzlicher Berücksichtigung des bildungssprachlichen Wortschatzwissens. Ebenso zeigte sich ein schwacher Effekt des bildungssprachlichen Wortschatzes auf die arithmetischen Fertigkeiten der Kinder ($\beta = .39, p = .08$).

Diese Ergebnisse legen nahe, dass der erfasste allgemeine bildungssprachliche Wortschatz insbesondere für den Sachunterricht und mathematische IQB-Aufgaben bedeutsam ist und vor allem ab der dritten Klasse relevant wird. In der zweiten Klasse hatte hingegen nur der allgemeine Wortschatz einen Effekt auf die Noten in Mathematik und im Sachunterricht ein Jahr später. Dies legt nahe, dass zum Ende der Grundschulzeit der bildungssprachliche gegenüber dem allgemeinen Wortschatz an Bedeutung gewinnt.

Die hier beschriebenen Ergebnisse sind allerdings für alle drei Skalen noch vorläufig, da die Stichproben, die in die Analysen einbezogen werden konnten, zurzeit noch vergleichsweise klein sind (s. oben). Die bildungssprachliche Wortschatzskala wurde in einer weiteren Studie bei Viertklässlern eingesetzt und geprüft. Hier zeigten sich signifikante Zusammenhänge der Skala mit allen untersuchten Schulnoten (Lesen, Schreiben, Rechnen, Sachunterricht); der allgemeine Wortschatz (PPVT) hatte nach Aufnahme der BiSpra Wortschatzaufgaben keinen signifikanten Effekt mehr (Schuth, Köhne & Weinert, im Druck). Wie bereits beschrieben, wird es durch ein Linkingverfahren möglich sein, die Analysen mit komplexeren multivariaten Verfahren anhand einer größeren Stichprobe zu wiederholen und zu erweitern.

Auch wenn die berichteten Befunde als vorläufig zu betrachten sind, zeigt sich trotz der noch geringen Größe der in die Analysen einbezogenen Stichprobe, dass die Aufgaben insgesamt gute Kennwerte aufweisen und sich die vorhergesagten prädiktiven Zusammenhänge mit Testleistungen und Noten weitgehend bestätigen. Dies ist selbst dann der Fall, wenn etablierte Instrumente zur Erfassung sprachlicher Kompetenzen, die nicht speziell auf bildungssprachliche Kompetenzen zielen (aber natürlich auch entsprechende Anforderungen enthalten können), gleichzeitig in die statistische Modellierung der Zusammenhänge aufgenommen werden. Die Ergebnisse können daher als Hinweis darauf interpretiert werden, dass

die erfassten bildungssprachlichen Fähigkeiten bereits in der Grundschule einen bedeutenden Einfluss auf den Schulerfolg haben. Erste empirische Hinweise deuten zudem darauf hin, dass der Einfluss des familiären sozio-ökonomischen Hintergrunds auf die Schulnoten vor allem auch durch die bildungssprachlichen Fähigkeiten der Kinder vermittelt wird.

1.3 Weiterführende Analysen im Rahmen von BiSpra

Der Zusammenhang zwischen dem sozio-ökonomischen (SES) und bildungsbezogenen familiären Hintergrund von Schülerinnen und Schülern und ihrem schulischen Kompetenzerwerb (z. B. Richter, Kuhl & Pant, 2012) sowie ihren sprachlichen Fähigkeiten (z. B. Weinert & Ebert, 2013) ist empirisch gut belegt. Da zu erwarten ist, dass insbesondere bildungssprachliche Fähigkeiten substanziell mit dem familiären und bildungsbezogenen Hintergrund der Schülerinnen und Schüler assoziiert sind, wurden im BiSpra-Projekt verschiedene Maße zur Bestimmung des SES bzw. des kulturellen Kapitals eingesetzt. Ein häufig benutzter Indikator für das kulturelle Kapital ist der familiäre Buchbestand (Paulus, 2009), der mithilfe der „Bücherfrage“ zumeist durch Eltern (z. B. in IGLU 2011; Tarelli, Wendt, Bos & Zylowski, 2012), zum Teil aber auch schon durch Grundschul Kinder (z. B. in TIMSS 2011; Bos, Wendt, Köller & Selter, 2012) eingeschätzt wird. Allerdings ist unklar, ob Grundschul Kinder bereits dazu in der Lage sind, reliable Angaben zum familiären Buchbestand zu machen. Daher sind wir im Rahmen des BiSpra-Projekts der Frage nachgegangen, ob bzw. inwieweit die Schätzungen bezüglich der Anzahl der Bücher zu Hause von Grundschulkindern und ihren Eltern übereinstimmen; speziell haben wir geprüft, ob die Übereinstimmung bei Schülerinnen und Schülern der vierten Jahrgangsstufe höher ist als bei Schülerinnen und Schülern der dritten Jahrgangsstufe (Pagel, Schönhoff, Domenech & Heppt, 2016). Des Weiteren wurde untersucht, ob sich der Zusammenhang zwischen dem familiären Buchbestand und den bildungssprachlichen Fähigkeiten der Schülerinnen und Schüler unterscheidet, je nachdem, ob die Bücherfrage durch Kinder der dritten oder der vierten Klasse oder durch ihre Eltern beantwortet wurde. Damit soll die Studie Hinweise darauf geben, ab welcher Klassenstufe sich die Bücheraufgabe als valider Indikator für den sozio-ökonomischen Status eignet (und zum Beispiel bei Fehlen einer Elternangabe als Proxy verwendet werden kann).

Die durchgeführten Analysen basieren auf Daten des 2. Messzeitpunkts. Die Analytestichprobe bestand aus 715 Kindern der dritten ($n = 344$, $M_{Alter} = 8.84$, $SD = 0.54$) und vierten Jahrgangsstufe ($n = 371$, $M_{Alter} = 9.88$, $SD = 0.56$) und ihren Eltern. Die Bücherfrage wurde sowohl bei den Eltern als auch bei den Schülerinnen und Schülern mithilfe einer fünfstufigen Skala erhoben. Zusätzlich zur sprachlichen Formulierung erhielten die Kinder für jede der fünf Kategorien eine bildliche Illustration. Die Einschätzungen der Bücherfrage der Schülerinnen und Schüler und ihrer Eltern wurden miteinander sowie mit den geschätzten Personenfähigkeiten (WLE) für die Hörverstehensaufgaben korreliert. Es zeigte sich, dass die Angaben der Eltern und der Schülerinnen und Schüler zwar substanziell korrelieren ($r = .50$, $p < .001$), jedoch nicht so hoch, wie aufgrund der identischen Itemformulierung zu erwarten wäre. Es zeigte sich außerdem, dass die Eltern- und Schülerangaben mit zunehmender Klassenstufe stärker übereinstimmen, die Schätzungen der Schülerinnen und Schüler also genauer werden. Allerdings ergaben sich daraus keine differenziellen Zusammenhänge mit der Hörverstehensleistung.

Da elektronische Bücher (E-Books) auch in Deutschland zunehmend an Bedeutung gewinnen (Berg, 2015), wurde anhand der Daten des 2. Messzeitpunkts außerdem untersucht, ob

E-Books zusätzlich in die Bücherfrage bei den Eltern aufgenommen werden sollten, um die Validität des Indikators sicher zu stellen (Pagel & Heppt, 2016). Die Analytestichprobe bestand in diesem Fall aus 845 Eltern, die den Elternfragebogen beantwortet hatten, und ihren Kindern. Zusätzlich zur Anzahl gedruckter Bücher gaben die Eltern auf einer 5-stufigen Skala an, wie viele E-Books sie besitzen. Es zeigte sich eine signifikante Korrelation ($r = .27$, $p < .001$) zwischen den beiden Angaben. Im Gegensatz zu der traditionellen Bücherfrage gab jedoch der Großteil der Eltern an, weniger als 10 E-Books zu besitzen (64.9%). Weitere 19.4% der Eltern beantworteten die Frage gar nicht. Nur die Angaben zur Anzahl gedruckter Bücher korrelierten signifikant mit den Leistungen der Kinder in den Hörverstehensaufgaben. In einer hierarchischen Regressionsanalyse mit dem bildungssprachlichen Hörverstehen als abhängiger Variable und der Anzahl an gedruckten Büchern und E-Books als unabhängigen Variablen ergab sich ein signifikanter Effekt für die Anzahl gedruckter Bücher, jedoch keine inkrementelle Validität für die Anzahl an E-Books. Die Ergebnisse deuten darauf hin, dass E-Books bislang nicht so häufig genutzt werden, wie vermutet werden könnte. Ihre Verbreitung auf dem Buchmarkt scheint deshalb die Validität der Bücherfrage zurzeit nicht einzuschränken. Aufgrund des auffälligen Antwortverhaltens (viele fehlende Werte) und der hohen Wahrscheinlichkeit dafür, dass E-Books in naher Zukunft an Bedeutung gewinnen werden, sind die Ergebnisse für die Bildungsforschung dennoch von Bedeutung.

Des Weiteren wurde untersucht, inwiefern die Entwicklung der bildungssprachlichen Kompetenzen mit dem schulischen Lernen zusammenhängt. Um diese Frage zu beantworten, wurde geprüft, ob die Fähigkeiten in den beiden sprachkomponenten-bezogenen bildungssprachlichen Maßen stärker mit der Klassenstufe oder mit dem Alter der Kinder zusammenhängen. Zweitens wurde untersucht, ob es insbesondere bildungssprachliche Maße sind, die - im Vergleich zu allgemeinen Sprachmaßen - mit der Klassenstufe zusammenhängen. Zur Erfassung des allgemeinen Wortschatzwissens wurde der PPVT eingesetzt und zur Messung allgemeiner Grammatikfähigkeiten der TROG-D. Für alle Kompetenzmaße wurden Personenfähigkeitsparameter (WLE) berechnet und in den Analysen verwendet. Die Stichprobe bestand aus $N = 837$ Kindern der Klassenstufen 2 und 3. Da der Unterschied zwischen Klassenstufe und Alter untersucht werden sollte, wurden Kinder, die im Vergleich zum Altersdurchschnitt der entsprechenden Klassenstufe überdurchschnittlich jung bzw. alt waren, aus den Analysen ausgeschlossen. Die Analysen basieren auf den Daten der Klassenstufen 2 und 3 zum ersten Messzeitpunkt und sind daher querschnittlich.

Zur Beantwortung der Hypothesen wurden zunächst Rank-Order-Korrelationen berechnet zwischen den Variablen Klassenstufe, Alter, dem Verständnis von Satzverbindungen und dem allgemeinen Grammatikverständnis, respektive dem bildungssprachlichen Wortschatzwissen und dem allgemeinen Wortschatzwissen. Die Ergebnisse zeigen, dass die Korrelationen zwischen allen Sprachmaßen und der Klassenstufe ($r_{\text{Satzverbindungen}} = .42^{***}$; $r_{\text{allgemeine Grammatik}} = .26^{***}$; $r_{\text{bildungssprachlicher Wortschatz}} = .61^{***}$; $r_{\text{allgemeinen Wortschatz}} = .44^{***}$) numerisch höher ausfielen als zwischen den Sprachmaßen und dem Alter ($r_{\text{Satzverbindungen}} = .32^{***}$; $r_{\text{allgemeine Grammatik}} = .17^*$; $r_{\text{bildungssprachlicher Wortschatz}} = .54^{***}$; $r_{\text{allgemeinen Wortschatz}} = .39^{***}$). Zudem waren die bildungssprachlichen Maße signifikant stärker mit der Klassenstufe korreliert als die generellen Sprachmaße (Fisher's $z_{\text{Satzverbindung-Grammatik}} = -3.14^{***}$; $z_{\text{allgemeines- bildungssprachliches Wortschatzwissen}} = -4.82^{***}$).

Anschließend wurde für jedes Sprachmaß als abhängige Variable eine ANOVA berechnet. Alter und Klassenstufe wurden als unabhängige Variable eingesetzt. Es zeigte sich, dass der

Effekt der Klassenstufe auf die bildungssprachlichen Maße (Satzverbindungen: $F = 25.82^{***}$, $\eta^2 = .16$; Wortschatz: $F = 18.56^{***}$, $\eta^2 = .12$) signifikant höher war als auf die allgemeinen Sprachmaße (Grammatik: $F = 20.39^{***}$, $\eta^2 = .07$; Wortschatz: $F = 21.68^{***}$, $\eta^2 = .03$). Das Alter der Kinder trug bei keinem der untersuchten Sprachmaße signifikant zur Erklärung der Leistungsvarianz der Kinder bei.

Zusammenfassend kann festgehalten werden, dass vor allem die Klassenstufe einen signifikanten Effekt auf die Fähigkeiten im Verständnis von Satzverbindungen und das bildungssprachlichen Wortschatzwissen hat – das Alter der Kinder hingegen weniger. Im Vergleich zu einem allgemeinen Grammatiktest und einem allgemeinen Wortschatztest ist der Effekt der Klassenstufe auf die beiden bildungssprachlichen Maße signifikant größer. Die Ergebnisse deuten auf einen Einfluss der schulischen Bildung auf die bildungssprachlichen Fähigkeiten hin und betonen damit die Wichtigkeit, Hindernisse bei deren Erwerb bereits in der Grundschule zu adressieren. Anschließende Forschungsfragen an diese Ergebnisse könnten darauf abzielen, spezifische Komponenten des schulischen Lernens zu identifizieren, die diesen Effekt erzeugen.

2 Voraussichtlicher Nutzen, insbesondere der Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans

Es wurden keine Patente angemeldet oder erteilt.

Die im Projekt entwickelten Aufgaben sowie die Forschungsergebnisse sind sowohl für Wissenschaftlerinnen und Wissenschaftler als auch für Praktikerinnen und Praktiker, die sich mit Sprachdiagnostik in der Grundschule beschäftigen, von erheblichem Interesse. Dies umfasst Sprachdidaktiker, Sprachwissenschaftler, empirische Bildungsforscher ebenso wie pädagogische Psychologen und Lehrkräfte in Grundschulen.

Die in der ersten Projektphase begonnene Entwicklung von testgeeigneten Aufgaben zur Erfassung schulrelevanter sprachlicher Kompetenzen ist sowohl für Forschungszwecke als auch für die Bildungspraxis von großem Nutzen, da bislang keine geeigneten Instrumente zur Bestimmung bildungssprachlicher Fähigkeiten vorliegen. Forschungsbemühungen zur Entwicklung reliabler und valider Diagnoseinstrumente sind insbesondere deshalb von besonderer Bedeutung, da eine gezielte Diagnostik schulrelevanter sprachlicher Kompetenzen die Grundlage einer effektiven Sprachförderung darstellt. In diesem Sinne besteht ein erhebliches Interesse an den im Projekt entwickelten Aufgaben. Erste Lizenzverträge zur Nutzung der BiSpra-Aufgaben in anderen Forschungsprojekten (u. a. im Rahmen der Forschungsinitiative „Bildung durch Sprache und Schrift“ (BiSS)), wurden bereits abgeschlossen. Darüber hinaus dienen die im Projekt entwickelten Hörverstehensaufgaben als Grundlage zur Entwicklung eines Testinstruments, mit dem sich erstsprachliche Kompetenzen bei Schülerinnen und Schülern mit russischer bzw. türkischer Herkunftssprache erfassen lassen. Dieses Testinstrument wird bei Schülerinnen und Schülern der 2. Jahrgangsstufe im Nationalen Bildungspanel (NEPS) eingesetzt.

3 Darstellung der wichtigsten Positionen des zahlenmäßigen Nachweises und der Notwendigkeit und Angemessenheit der geleisteten Arbeiten

Im Berliner Teilprojekt fielen die zahlenmäßig größten Ausgaben für die im Projekt beschäftigten Mitarbeiterinnen an. Neben einer wissenschaftlichen Mitarbeiterstelle (TV-L – E 13, Stellenumfang: 50%, Gesamtsumme: ca. 94.000 €), waren drei studentische Mitarbeiterinnen im Projekt beschäftigt (Stellenumfang: zwischen 41 und 80h/Monat, Gesamtsumme: ca. 57.000 €). Der Stellenumfang war insgesamt äußerst gering bemessen und für die Bewältigung der Aufgaben (u. a. Entwicklung weiterer Aufgaben zur Erfassung des bildungssprachlichen Hörverstehens, Planung und Durchführung einer Pilotierungsstudie an mehreren Berliner Schulen, Vorbereitung der Datenerhebung im Rahmen der Validierungsstudie, Dateneingabe, -aufbereitung und -auswertung sowie Dokumentation, Budgetplanung und Drittmittelverwaltung etc.) und die Erreichung der Projektziele notwendig und angemessen.

Im Bamberger Teilprojekt fielen die zahlenmäßig höchsten Ausgaben für die im Projekt beschäftigten Mitarbeiterinnen sowie für die Durchführung der Erhebungen der Validierungsstudie durch ein externes Erhebungsinstitut (IEA DPC) an. Die wissenschaftlichen Mitarbeiterinnen (1,5 TV-L E13-Stellen; Gesamtkosten: ca. 263.000 €) wurden - über die drei Jahre Projektlaufzeit verteilt - von insgesamt zehn studentischen Mitarbeiterinnen unterstützt (Stellenumfang: ca. zwischen 20 und 60h/Monat, Gesamtsumme: ca. 47.700 €). Die Personalkosten waren zentral für die Erreichung der Projektziele und unter Berücksichtigung der vielfältigen, in kurzem Zeitraum notwendigen Entwicklungs-, Erprobungs-, Auswertungs-, Organisations- und Dokumentationsarbeiten eher knapp angesetzt (u. a. Erstellung eines bildungssprachlichen Wortpools; Entwicklung und Erprobung verschiedener Aufgabenformate und Items zur Erfassung bildungssprachlichen Wortschatzwissens und Konnektorenverständnisses; Planung und Durchführung einer Pilotierungsstudie an Horten in Nürnberg sowie (anteilig in enger Kooperation mit dem Teilprojekt Berlin) Vorbereitung der Datenerhebung im Rahmen der Validierungsstudie (Dateneingabe, -aufbereitung und -auswertung, Dokumentation, Vorbereitung des quer- und längsschnittlichen Linking der verschiedenen Instrumentenversionen). Die von den Projektmitarbeiterinnen (beide Standorte) vorbereiteten Erhebungen der Validierungsstudie (Messwiederholungsstudie mit zwei Kohorten) wurden von einem externen Erhebungsinstitut durchgeführt (Gesamtausgaben: 226.100 €). Um überregional Kompetenzdaten (und ergänzende Befragungsdaten) in hoher Qualität erheben zu können, ist die Einbindung eines professionellen, erfahrenen und verlässlichen Partners, der über die notwendige Expertise, Logistik sowie einen entsprechenden Personalstamm für die Realisierung solcher Erhebungen verfügt – mit Blick auf Zeit, Kosten und Datenqualität – zur Erreichung entsprechender Ziele notwendig und sinnvoll.

4 Während der Durchführung des Vorhabens dem ZE bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Während der Durchführung des Vorhabens wurden im deutschen Sprachraum keine anderen Testverfahren zur Erfassung bildungssprachlicher Fähigkeiten bei Kindern im Grundschulalter vorgelegt. Studien zu Zusammenhängen zwischen bildungssprachlichen Fähigkeiten und Schulerfolg wurden – soweit uns bekannt ist – insbesondere in den an BiSpra

beteiligten Arbeitsgruppen durchgeführt. Allerdings beschäftigen sich auch die Arbeitsgruppen um Prof. Dr. Astrid Rank (Universität Regensburg) und Prof. Dr. Anja Wildemann (Universität Koblenz-Landau) mit der Operationalisierung von Bildungssprache. In ihren Arbeitsgruppen wurden zwei Ratingskalen entwickelt, mit denen sich die bildungssprachlichen Mittel identifizieren lassen, die von Kindern im Elementar- und Primarbereich produziert werden (vgl. Fornol, Heppt, Sutter, Hartinger, Rank & Wildemann, 2015).

5 Erfolgte und geplante Veröffentlichungen der Ergebnisse^{17 18}

Zeitschriftenartikel

- * Dragon, N., Berendes, K., Weinert, S., Heppt, B. & Stanat, P. (2015). Ignorieren Grundschul Kinder Konnektoren? — Untersuchung einer bildungssprachlichen Komponente. *Zeitschrift für Erziehungswissenschaft*, 18(4), 803-825.
- Haag, N., Heppt, B., Stanat, P., Kuhl, P. & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction*, 28, 24-34.
- Heppt, B., Henschel, S. & Haag, N. (2016). Everyday and academic language comprehension: Investigating their relationships with school success and challenges for language minority learners. *Learning and Individual Differences*. Advance online publication. doi: 10.1016/j.lindif.2016.01.004
- Heppt, B., Haag, N., Böhme, K. & Stanat, P. (2015). The role of academic-language features for reading comprehension of language-minority students and students from low-SES families. *Reading Research Quarterly*, 50(1), 61–82.
- * Heppt, B., Stanat, P., Dragon, N., Berendes, K. & Weinert, S. (2014). Bildungssprachliche Anforderungen und Hörverstehen bei Kindern mit deutscher und nicht-deutscher Familiensprache. *Zeitschrift für Pädagogische Psychologie*, 28(3), 139-149.
- Knoepke, J., Richter, R., Isberner, M.-B., Naumann, J., Neeb, Y. & Weinert, S. (2016). Processing of positive-causal and negative-causal coherence relations in primary school children and adults: A test of the cumulative cognitive complexity approach in German. *Journal of Child Language*. DOI: 10.1017/S0305000915000872
- Schuth, E., Köhne, J. & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning & Instruction*, 49, 157-165. DOI: <http://dx.doi.org/10.1016/j.learninstruc.2017.01.005>

¹⁷ Aufgeführt sind all jene Publikationen, Vorträge und Poster mit Bezug zu BiSpra, die während der zweiten Förderphase (3/2013 – 7/2016) oder im Anschluss daran entstanden sind.

¹⁸ Die mit einem * markierten Veröffentlichungen basieren auf den im BiSpra-Projekt erhobenen Daten.

Buchkapitel

- * Berendes, K., Dragon, N., Weinert, S., Heppt, B. & Stanat, P. (2013). Hürde Bildungssprache? Eine Annäherung an das Konzept „Bildungssprache“ unter Einbezug aktueller empirischer Forschungsergebnisse. In A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik. Interdisziplinäre Perspektiven* (S. 17-41). Münster: Waxmann.
- Fornol, S., Heppt, B., Sutter, S., Hartinger, A., Rank, A. & Wildemann, A. (2015). Entwicklung und Erfassung bildungssprachlicher Merkmale. Forschungserkenntnisse und Perspektiven. In K. Liebers, B. Landwehr, A. Marquardt & K. Schlotter (Hrsg.), *Lernprozessbegleitung und adaptive Lerngelegenheiten im Unterricht der Grundschule. Ergebnisse der 23. Jahrestagung der Kommission Grundschulforschung und Pädagogik der Primarstufe. Band 1* (S. 157-168). Wiesbaden: Springer VS.
- * Köhne, J., Kronenwerth, S., Redder, A., Schuth, E. & Weinert, S. (2015). Bildungssprachlicher Wortschatz – linguistische und psychologische Fundierung und Itementwicklung. In A. Redder, J. Naumann & R. Tracy (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS) – Ergebnisse* (S. 67-92). Münster: Waxmann.
- Redder, A. & Weinert, S. (2013). Sprachliche Handlungsfähigkeiten im Fokus von FiSS. Zur Einleitung in den Sammelband. In A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik - Interdisziplinäre Perspektiven* (S. 7-16). Münster: Waxmann.
- * Schuth, E., Heppt, B., Köhne, J., Weinert, S. & Stanat, P. (2015). Die Erfassung schulisch relevanter Sprachkompetenzen bei Grundschulkindern – Entwicklung eines Testinstruments. In A. Redder, J. Naumann & R. Tracy (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS) – Ergebnisse* (S. 73-92). Münster: Waxmann.
- Weinert, S. (2016). Natürliche Sprache(n) und Formelsprache(n) in der Bildung. In J. Killian, B. Brouër & D. Lüttenberg (Hrsg.), *Handbuch Sprache in der Bildung* (S. 3-24). Berlin: De Gruyter.
- Weinert, S. (in Druck). Sprachliche Bildung – Sprache in der Bildung. In W. Thielmann, C. Trautmann, A. Krause & G. Lehmann (Hrsg.), *Form und Funktion. Festschrift für Angelika Redder zum 65. Geburtstag*. Tübingen: Stauffenburg.

Andere Publikationen (geordnet nach Jahren)

- Ebert, S., Köhne, J. & Weinert, S. (2014). Mit Worten Türen öffnen! *uni.vers*, 42-45.
- Heppt, B. (2016). *Verständnis von Bildungssprache bei Kindern mit deutscher und nicht-deutscher Familiensprache*. Dissertation, Humboldt-Universität zu Berlin. Verfügbar unter: <http://edoc.hu-berlin.de/dissertationen/heppt-birgit-2016-05-23/PDF/heppt.pdf>.

- * Heppt, B., Schuth, E., Weinert, S. & Stanat, P. (2015). Bedeutung und Erfassung bildungssprachlicher Fähigkeiten bei Grundschulkindern. In A. Redder & S. Lambert (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung. Transfer-Perspektiven* (S. 24-26). Hamburg: FiSS-Koordinierungsstelle.
 - * Köhne, J., Schuth, E., Heppt, B., Weinert, S. & Stanat, P. (2014). Bildungssprachliche Kompetenzen von Kindern im Grundschulalter. In A. Redder & S. Lambert (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung. Die Projekte stellen sich vor* (S. 16-19). Hamburg: FiSS-Koordinierungsstelle.
- Redder, A. & Weinert, S. (Hrsg.). (2013). *Sprachförderung und Sprachdiagnostik - Interdisziplinäre Perspektiven*. Münster: Waxmann.
- * Weinert, S., Stanat, P., Redder, A., Dragon, N., Heppt, B. & Uessler, S. (2014). *Verbundvorhaben: Bildungssprachliche Kompetenzen (BiSpra): Anforderungen, Sprachverarbeitung und Diagnostik. Schlussbericht 2013*. Bamberg u.a.: Otto-Friedrich-Universität Bamberg u.a. Verfügbar über die Universitätsbibliothek Hannover und Technische Informationsbibliothek.

Konferenzbeiträge

- * Berendes, K., Dragon, N., Weinert, S., Heppt, B. & Stanat, P. (2013, März). *Alltagssprachliches und bildungssprachliches Wort- und Satzverständnis im Grundschulalter*. Poster präsentiert auf dem 1. Kongress der Gesellschaft für Empirische Bildungsforschung (GEBF), Kiel.
 - * Dragon, N., Berendes, K., Weinert, S., Heppt, B. & Stanat, P. (2013, März). *Konnektorenverständnis im Grundschulalter: Ignorieren Grundschulkinde Konnektoren?* Poster präsentiert auf dem 1. Kongress der Gesellschaft für Empirische Bildungsforschung (GEBF), Kiel.
 - * Dragon, N., Berendes, K., Weinert, S., Heppt, B. & Stanat, P. (2013, Juni). *Assessing academic language comprehension: Do primary school children ignore clause connectors?* Poster presented at the Child Language Seminar (CLS), Manchester, UK.
- Heppt, B., Haag, N., Stanat, P. & Böhme, K. (2013, April). *Reading comprehension in second language learners: The role of academic language features*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), San Francisco, USA.
- Heppt, B., Haag, N., Stanat, P. & Böhme, K. (2013, August). *Academic language features: An obstacle for second language learners' reading comprehension?* Paper presented at the 15th Conference of the European Association for Research on Learning and Instruction (EARLI), Munich, Germany.

- * Heppt, B. & Henschel, S. (2015, September). *Bedeutung familiärer Struktur- und Prozessmerkmale für bildungssprachliches Hörverstehen*. Vortrag auf der 15. Fachgruppentagung Pädagogische Psychologie der DGPs (PAEPS), Kassel.
- Heppt, B., Henschel, S. & Haag, N. (2014, September). *Bildungssprachliche Anforderungen und Leseverstehen bei Kindern im Grundschulalter*. Vortrag auf der 23. Jahrestagung der Kommission Grundschulforschung und Pädagogik der Primarstufe der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE), Leipzig.
- Heppt, B., Henschel, S. & Haag, N. (2015, August). *Can proficiency in everyday language and proficiency in academic language be differentiated?* Paper presented at the 16th Conference of the European Association for Research on Learning and Instruction (EARLI), Limassol, Cyprus.
- * Heppt, B., Köhne, J. & Schuth, E. (2014, November). *Erfassung bildungssprachlicher Fähigkeiten im BiSpra-Projekt*. Vortrag auf der 6. Herbstschule der Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS), Frankfurt am Main.
- * Heppt, B., Schuth, E. & Taraszow, T. (2015, Mai). *Diagnostik schulrelevanter sprachlicher Kompetenzen im Kontext von Mehrsprachigkeit*. Vortrag auf der 3. Internationalen FiSS-Frühjahrstagung, Hamburg.
- * Heppt, B. & Stanat, S. (2016, September). *Bedingungen und Effekte sprachbildenden Unterrichts in der Grundschule*. Vortrag auf der 81. Tagung der Arbeitsgruppe für empirische pädagogische Forschung (AEPF), Rostock.
- * Pagel, L. & Heppt, B. (2016, July). *„What about e-books? – Do electronic books need to be considered in educational research?* Poster presented at the Conference of Junior Researchers of EARLI (JURE), Helsinki, Finland.
- * Pagel, L. & Heppt, B. (2016, September). *Incentivierung durch Verlosung: Effekte auf Rücklauf und Qualität von Elternfragebögen*. Vortrag auf der Pre-Conference der 81. Tagung der Arbeitsgruppe für empirische pädagogische Forschung (AEPF), Rostock.
- * Pagel, L., Schönhoff, K., Domenech, F. & Heppt, B. (2016, März). *„Wie viele Bücher gibt es bei dir zuhause ungefähr?“ – Können Kinder die Bücherfrage im Laufe der Grundschulzeit genauer beantworten?* Poster präsentiert auf dem 4. Kongress der Gesellschaft für Empirische Bildungsforschung (GEBF), Berlin.
- Schuth, E. (2014, März). *BiSpra II: Bildungssprachliche Kompetenzen, Anforderungen, Sprachverarbeitung und Diagnostik*. Vortrag auf der 2. FiSS-Frühjahrstagung, Dortmund.
- Schuth, E., Köhne, J. & Weinert, S. (2016, März). *Academic vocabulary predicts school success*. Vortrag auf dem 4. Kongress der Gesellschaft für Empirische Bildungsforschung (GEBF), Berlin.

Schuth, E., Köhne, J. & Weinert, S. (2015, July). *The development of school relevant language skills in primary school*. Paper presented at the European Congress of Psychology (ECP), Milano, Italy.

Schuth, E., Köhne, J. & Weinert, S. (2014, July). *The role of academic language development in elementary school*. Roundtable discussion at the Junior Researchers of EARLI (JURE), Nikosia, Cyprus.

Eingeladene Vorträge/Workshops

Heppt, B. (2015, Januar). *Alltagssprache und Bildungssprache – Was unterstützt Kinder im Verstehen der Bildungssprache im Jahr vor der Einschulung?* Vortrag im Rahmen des Qualifizierungsprogramms „Alltagsintegrierte Sprachbildung und Sprachförderung für (Kinder-)Tageseinrichtungen im Landkreis Goslar“, Goslar.

* Heppt, B. & Schuth, E. (2015, März). *Erfassung bildungssprachlicher Fähigkeiten bei Kindern im Grundschulalter: Ein Einblick in das Projekt BiSpra*. Workshop auf dem Frankfurter Fachforum „Sprache und Naturwissenschaften, Frankfurt am Main.

* Heppt, B. & Stanat, S. (2016, Juli). *Gibt es Unterschiede in der bildungssprachlichen Entwicklung von ein- und mehrsprachigen Lernenden? Theoretische Annahmen und empirische Befunde*. Vortrag auf der gemeinsamen Tagung des Trägerkonsortiums Bildung durch Sprache und Schrift (BiSS) und der Koordinierungsstelle Mehrsprachigkeit und Bildung (KoMBI) „Sprachliche Bildung in der Einwanderungsgesellschaft“, Köln.

Weinert, S. (2015, Mai). *Warum Sprachförderung und Sprachdiagnostik so wichtig sind: Zur Bedeutung von Sprache für die kindliche Entwicklung*. Vortrag auf der 3. Internationalen FiSS-Frühjahrstagung, Hamburg.

Weinert, S. (2015, April). *Sprache in der Bildung: Bedeutung und Einflussvariablen im Vor- und Grundschulalter*. Vortrag im Rahmen der Vortragsreihe „Sprachliche Bildung & kulturelle Praxis“ am Institut für Bildungswissenschaften, Universität Basel / Fachhochschule Nordschweiz, Basel.

Weinert, S. (2016, August). *Bedeutung von Sprache für die kindliche Entwicklung und schulisches Lernen*. Vortrag bei der 9. SH-Sommeruniversität „Sprache und Kommunikation im Unterricht“, Sankelmark.

Weinert, S. (2016, Oktober). *Language and its acquisition: Influential conditions and impact on child development*. Invited Keynote Lecture at the Annual Conference of the Society for Longitudinal and Life Course Studies (SLLS), Bamberg.

6 Im Bericht zitierte Literatur

- Anstrom, K., DiCerbo, P., Butler, F. A., Katz, A., Millet, J. & Rivera, C. (2010). *A review of the literature on academic English: Implications for K-12 English language learners*. Arlington, VA: The George Washington University.
- Bailey, A. L. & Butler, F. A. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Bailey, A. L., Butler, F. A., Stevens, R. & Lord, C. (2007). Further specifying the language demands of school. In A. L. Bailey (Hrsg.), *The language demands of school. Putting academic English to the test* (S. 103-156). New Haven, CT: Yale University Press.
- Bailey, A. L., Stevens, R., Butler, F. A. & Huang, B. H. (2005). *Using standards and empirical evidence to develop academic English proficiency test items in reading*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Berg, A. (2015). *Studie zur Nutzung von E-Books*. Berlin: Bitkom Research.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (Hrsg.). (2012). *TIMSS 2011: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bulheller, S. & Häcker, H. O. (2003). *Peabody Picture Vocabulary Test. PPVT. Deutschsprachige Fassung des PPVT-III für Jugendliche und Erwachsene*. Frankfurt/M.: Swets.
- Butler, F. A., Bailey, A. L., Stevens, R. & Huang, B. H. (2004). *Academic English in fifth-grade mathematics, science, and social studies textbooks*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Cattell, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest Skala 1 (CFT 1)* (5., rev. Aufl.). Göttingen: Hogrefe.
- Chamot, A. U. & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Reading, MA: Addison-Wesley.
- Coxhead, A. (2000). A new academic wordlist. *TESOL Quarterly*, 34(2), 213-238.
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 121-129.
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters LTD.

- Dragon, N., Berendes, K., Weinert, S., Heppt, B. & Stanat, P. (2015). Ignorieren Grundschul Kinder Konnektoren? — Untersuchung einer bildungssprachlichen Komponente. *Zeitschrift für Erziehungswissenschaft*, 18(4), 803-825.
- Dunn, L. M. & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test - Revised (PPVT-R). Manual for forms L and M*. Circle Pines, MN: American Guidance Service.
- Eckhardt, A. G. (2008). *Sprache als Barriere für den schulischen Erfolg. Potentielle Schwierigkeiten beim Erwerb schulbezogener Sprache für Kinder mit Migrationshintergrund*. Münster: Waxmann.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Fornol, S., Heppt, B., Sutter, S., Hartinger, A., Rank, A. & Wildemann, A. (2015). Entwicklung und Erfassung bildungssprachlicher Merkmale. Forschungserkenntnisse und Perspektiven. In K. Liebers, B. Landwehr, A. Marquardt & K. Schlotter (Hrsg.), *Lernprozessbegleitung und adaptives Lernen in der Grundschule. Forschungsbezogene Beiträge* (S. 157-168). Wiesbaden: Springer.
- Fox, A. V. (2007). *TROG-D. Test zur Überprüfung des Grammatikverständnisses*. Idstein: Schulz-Kirchner.
- Gogolin, I. (2006). Chancen und Risiken nach PISA - über die Bildungsbeteiligung von Migrantenkindern und Reformvorschläge. In G. Auernheimer (Hrsg.), *Schieflagen im Bildungssystem. Die Benachteiligung der Migrantenkinder* (Bd. 2, S. 33-50). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gogolin, I. (2009). Zweisprachigkeit und die Entwicklung bildungssprachlicher Fähigkeiten. In I. Gogolin & U. Neumann (Hrsg.), *Streitfall Zweisprachigkeit – The bilingualism controversy* (S. 263-280). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gogolin, I. & Lange, I. (2011). Bildungssprache und Durchgängige Sprachbildung. In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Mehrsprachigkeit* (S. 107-128). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gogolin, I., Dirim, I., Klinger, T., Lange, I., Lengyel, D., Michel, U. et al. (2011). *Förderung von Kindern und Jugendlichen mit Migrationshintergrund FörMig. Bilanz und Perspektiven eines Modellprogramms*. Münster: Waxmann.
- Granzer, D., Walther, G., Winkelmann, H., Robitzsch, A. & Köller, O. (2008). *Bildungsstandards: Kompetenzen überprüfen und fördern. Grundschule Mathematik 3./4. Schuljahr. Heft 1*. Berlin: Cornelsen.
- Haffner, J., Baro, K., Parzer, P. & Resch, F. (2005). *HRT 1-4. Heidelberger Rechentest - Erfassung mathematischer Basiskompetenzen im Grundschulalter*. Göttingen: Hogrefe.

- Halliday, M. A. K. & Hasan, R. (1989). *Language, context and text: Aspects of language in a social-semiotic perspective* (2nd ed.). Oxford, UK: Oxford University Press.
- Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A. & Kliegl, R. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 62(1), 10-20.
- Heppt, B., Dragon, N., Berendes, K., Stanat, P. & Weinert, S. (2012). Beherrschung von Bildungssprache bei Kindern im Grundschulalter. *Diskurs Kindheits- und Jugendforschung*, 3, 349-356.
- Heppt, B., Stanat, P., Dragon, N., Berendes, K. & Weinert, S. (2014). Bildungssprachliche Anforderungen und Hörverstehen bei Kindern mit deutscher und nicht-deutscher Familiensprache. *Zeitschrift für Pädagogische Psychologie*, 28(3), 139-149.
- Knoepke, J., Richter, R., Isberner, M.-B., Naumann, J., Neeb, Y. & Weinert, S. (2016). Processing and acquisition of positive-causal and negative-causal coherence relations in primary school children and adults: A test of the cumulative cognitive complexity approach in German. *Journal of Child Language*. doi:10.1017/S0305000915000872.
- Köhne, J., Kronenwerth, S., Redder, A., Schuth, E. & Weinert, S. (2015). Bildungssprachlicher Wortschatz – linguistische und psychologische Fundierung und Itementwicklung. In A. Redder, J. Naumann & R. Tracy (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS) – Ergebnisse* (S. 67-92). Münster: Waxmann.
- Kultusministerkonferenz. (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4). Beschluss der Kultusministerkonferenz von 15.10.2004*. München: Luchterhand.
- Lenhard, W. & Schneider, W. (2006). *ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Melchers, P. & Preuß, U. (2009). *Kaufman-Assessment Battery for Children. Deutsche Version (K-ABC)*. Frankfurt: Pearson.
- OECD. (2005). *PISA 2003. Technical Report*. Paris, France: OECD Publishing.
- OECD. (2006). *Where immigrant students succeed - A comparative review of performance and engagement in PISA 2003*. Paris, France: OECD Publishing.
- OECD. (2010). *PISA 2009 results: Overcoming social background: Vol. 2. Equity in learning opportunities and outcomes*. Paris, France: OECD Publishing.
- Pagel, L., Schönhoff, K., Domenech, F. & Heppt, B. (2016, März). „Wie viele Bücher gibt es bei dir zuhause ungefähr?“ – Können Kinder die Bücherfrage im Laufe der

- Grundschulzeit genauer beantworten?* Poster präsentiert auf dem 4. Kongress der Gesellschaft für Empirische Bildungsforschung (GEBF), Berlin.
- Pagel, L. & Heppt, B. (2016, July). „*What about e-books? – Do electronic books need to be considered in educational research?*“ Poster presented at the Conference of Junior Researchers of EARLI (JURE), Helsinki, Finland.
- Paulus, C. (2009). *Die "Bücheraufgabe" zur Bestimmung des kulturellen Kapitals bei Grundschulern*. Universität des Saarlandes, Saarbrücken.
- Richter, D., Kuhl, P. & Pant, H. A. (2012). Soziale Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 191-206). Münster: Waxmann.
- Roßbach, H. G., Tietze, W. & Weinert, S. (2005). *Peabody Picture Vocabulary Test - Revised. Deutsche Forschungsversion des Tests von L. M. Dunn / L. M. Dunn von 1981*: Universität Bamberg, FU-Berlin.
- Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431-459.
- Schleppegrell, M. J. (2004). *The language of schooling. A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum.
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A. & Kliegl, R. (2015). childLex: a lexical database of German read by children. *Behavior Research Methods*, 47(4), 1085-1094.
- Schuth, E., Heppt, B., Köhne, J., Weinert, S. & Stanat, P. (2015). Die Erfassung schulisch relevanter Sprachkompetenzen bei Grundschulkindern – Entwicklung eines Testinstruments. In A. Redder, J. Naumann & R. Tracy (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS) – Ergebnisse* (S. 73-92). Münster: Waxmann.
- Schwippert, K., Wendt, H. & Tarelli, I. (2012). Lesekompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 191-207). Münster: Waxmann.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Tarelli, I., Wendt, H., Bos, W. & Zylowski, A. (2012). Ziele, Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2011). In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von*

Grundschulkindern in Deutschland im internationalen Vergleich (S. 27-67). Münster: Waxmann.

Uessler, S., Runge, A. & Redder, A. (2013). "Bildungssprache" diagnostizieren. Entwicklung eines Instruments zur Erfassung von bildungssprachlichen Fähigkeiten bei Viert- und Fünftklässlern In A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik. Interdisziplinäre Perspektiven* (S. 42-67). Münster: Waxmann.

Weinert, S. & Ebert, S. (2013). Spracherwerb im Vorschulalter: Soziale Disparitäten und Einflussvariablen auf Grammatikerwerb. *Zeitschrift für Erziehungswissenschaft*, 16, 303-332.

Weinert, S., Stanat, P., Redder, A., Dragon, N., Heppt, B. & Uessler, S. (2014). *Verbundvorhaben: Bildungssprachliche Kompetenzen (BiSpra): Anforderungen, Sprachverarbeitung und Diagnostik. Schlussbericht 2013*. Bamberg u.a.: Otto-Friedrich-Universität Bamberg u.a. Verfügbar über die Universitätsbibliothek Hannover und Technische Informationsbibliothek.

Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. A. (2007). *Acer ConQuest Version 2.0: Generalised item response modelling software*. Camberwell, Victoria: ACER Press.