

Protokoll: Systematisches Review zur Wirkung von didaktisch-methodischen Ansätzen des sprachsensiblen Unterrichts

Stand der Bearbeitung:

29.04.2020

(formal geänd. 09.09.2020)

Till Woerfel, Martha Höfler, Annika Witte, Anastasia Knaus, Leonie Twente, Rebekka Wanka & Michael Becker-Mrotzek (Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache, Universität zu Köln)

Inhaltverzeichnis

1.	Hintergrund	4
2.	Untersuchungsgegenstand	5
2.1.	Definition des Begriffs „Sprachsensibler Unterricht“	5
2.2.	International bekannte sprachensible Unterrichtsansätze	5
2.3.	Abgrenzung von verwandten sprach- und fachintegrierten Unterrichtsansätzen	7
3.	Forschungsstand	8
4.	Ziel, Art und Nutzen der Untersuchung	10
5.	Methodik	11
5.1.	Leitende Untersuchungsfrage	11
5.2.	Inklusionskriterien	12
5.3.	Systematische Studiensuche	16
5.3.1.	Suchbegriffentwicklung	16
5.3.2.	Suchorte	17
5.3.3.	Suchstrategie	19
5.4.	Screening	21
5.4.1.	Screening-Prozess	21
5.4.2.	Aufbereitung der ermittelten Dokumente	24
5.5.	Kodierung und Studienbewertung	24
5.5.1.	Kodierung der Studieninformationen	24
5.5.2.	Beurteilung der internen und externen Validität	24
5.6.	Narrative Forschungssynthese	28
6.	Finanzierung/Auftraggeber	29
7.	Projektteam und Kooperationspartner	29
	Referenzen	31
	Appendix	38

Kontakt

Dr. Martha Höfler

Wissenschaftliche Mitarbeiterin

Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache

Telefon: 49 (0) 221 470 7870

E-Mail: martha.hoefler@mercator.uni-koeln.de

Über das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache

Das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache ist ein durch die Stiftung Mercator initiiertes und gefördertes Institut der Universität zu Köln. Es will sprachliche Bildung verbessern. Um dieses Ziel zu erreichen, erforscht und entwickelt es innovative Konzepte, Maßnahmen und Instrumente für sprachliche Bildung. Es bildet regional Lehramtsstudierende aus sowie bundesweit Pädagoginnen und Pädagogen in Kitas, Schulen und der Erwachsenenbildung fort und bereitet wissenschaftliche Erkenntnisse gezielt für Entscheidungsträger in Bildungspolitik und -verwaltung sowie Bildungspraxis auf. Mit seiner Forschung und seinen wissenschaftlichen Serviceleistungen zu sprachlicher Bildung in einer mehrsprachigen Gesellschaft trägt das Mercator-Institut zu mehr Chancengleichheit im Bildungssystem bei.

www.mercator-institut-sprachfoerderung.de

1. Hintergrund

Die internationalen Schulleistungsstudien in den 2000er Jahren haben der allgemeinen und der Fachöffentlichkeit sehr deutlich vor Augen geführt, dass das deutsche Bildungssystem nicht den hohen Erwartungen entspricht. Vor allem PISA 2000 wurde von vielen aus den folgenden Gründen als Schock wahrgenommen:

- a) Die Leseleistungen der deutschen Schülerinnen und Schüler lagen deutlich unter dem internationalen Durchschnitt.
- b) In kaum einem anderen Land ist der Abstand zwischen den leistungsstarken und leistungsschwachen Schülerinnen und Schüler so groß ist wie in Deutschland.
- c) In kaum einem anderen Land ist Schulerfolg so eng mit dem Sozialstatus verbunden.
- d) Die Leseleistungen korrelieren hochgradig mit den Leistungen in anderen Fächern.

Mit Blick auf die sprachliche Bildung hatte dies folgende Konsequenzen: Es wurden Bildungsstandards für das Fach Deutsch eingeführt, Bund und Länder haben zahlreiche Initiativen zur Leseförderung gestartet, zuletzt das Programm „Bildung durch Sprache und Schrift – BiSS“ und schließlich hat die Idee der durchgängigen Sprachbildung Eingang in die Curricula gefunden, vor allem in Form des sprachsensiblen Unterrichts. Ein solcher Unterricht unterstützt Schülerinnen und Schüler im Fachunterricht dabei, fachliche Inhalte sprachlich zu durchdringen, bildungssprachliche Fertigkeiten auszubauen und sukzessive die ihnen gestellten Aufgaben im Unterricht selbstständig zu bewältigen (Becker-Mrotzek & Woerfel, im Erscheinen).

Die Bestrebungen von Bildungspolitik und -administration, sprachsensiblen Unterricht in der Praxis zu implementieren, werden zwar aus dem theoretischen Erklärungswissen abgeleitet, dass Sprache und Lernen untrennbar miteinander verbunden sind; sie beziehen sich aber aktuell kaum auf Erkenntnisse, die, wie von der Kultusministerkonferenz (2016, S. 16) gewünscht, aus qualitätssichernden Verfahren hervorgehen. In einer evidenzorientierten Bildungspolitik hat empirisch gesichertes Wissen besonderes Gewicht (Gräsel, 2010). Sofern bislang überhaupt empirische Argumente in der Nutzendebatte zu sprachsensiblen Unterricht eine Rolle spielen, beziehen sich diese auf einzelne, aus der Studienlage herausgegriffene Untersuchungen. Um Bildungsmaßnahmen aus wissenschaftlicher Perspektive zu begründen, braucht es einen umfassenden und systematischen Überblick über die vorliegenden empirischen Erkenntnisse. Vor allem im schulischen Kontext lässt sich im Hinblick auf Sprachbildungsmaßnahmen „ein Missverhältnis zwischen der wachsenden Anzahl dieser Maßnahmen und gesicherten Erkenntnissen bezüglich ihrer Wirksamkeit feststellen“ (Busse, 2019, S. 14). Paetsch, Wolf, Stanat und Darsow (2014, S. 336) sprechen gar von einer „klaffende[n] Forschungslücke“ hinsichtlich der untersuchten Wirkung sprachfördernder bzw. sprachbildender Angebote. Schlussendlich wissen wir heute noch nicht sicher, ob das, was bereits in Lehrplänen und Curricula verankert ist, tatsächlich wirkt.

Es bedarf der forschungsgeliteten Klärung dieses Sachverhalts, um die Bestrebungen der Schulpolitik und -administration wissenschaftlich zu reflektieren und, in der Folge, zu stützen oder entsprechend der vorliegenden Evidenz anzupassen. Auch für die Professionalisierung von Lehrkräften sind wissenschaftliche Erkenntnisse zur Wirkung von sprachsensiblen Unterricht bedeutsam. Nur durch das Wissen darüber, welche sprachsensiblen Ansätze nachweislich den Lernerfolg begünstigen, können Lehrkräfte im Rahmen von Aus-, Fort- und Weiterbildung entsprechende Kompetenzen erwerben und Schülerinnen und Schüler effektiv beim Lernen unterstützen.

In diesem Zusammenhang soll im Rahmen eines systematischen Reviews die Frage geklärt werden, ob sprachensible Unterrichtsansätze im Primar- und Sekundarbereich nachweislich wirkungsvoll sind.

2. Untersuchungsgegenstand

2.1. Definition des Begriffs „Sprachsensibler Unterricht“

Der Begriff ‚Sprachsensibler Unterricht‘ steht für ein Prinzip der Unterrichtsgestaltung bei sprachlich heterogenen Lerngruppen. Darunter werden unterschiedliche Ansätze gefasst, deren Ziel es ist, Schülerinnen und Schüler im Fachunterricht dabei zu unterstützen, fachliche Inhalte sprachlich zu durchdringen sowie allgemein- und fachsprachliche Fertigkeiten auszubauen, um schulische Aufgaben und Anforderungen im Fachunterricht bewältigen zu können (Becker-Mrotzek & Woerfel, im Erscheinen).

Innerhalb heterogener Lerngruppen zielen diese Ansätze insbesondere auf Schülerinnen und Schüler, die wegen unzureichender sprachlicher Fähigkeiten und Fertigkeiten dem Unterricht nicht folgen können und somit die Lernziele nur schwer oder gar nicht erreichen (z. B. Kinder und Jugendliche, die aus Familien mit geringem Zugang zu Bildung kommen, oder die erst im schulpflichtigen Alter beginnen, die Schul-/Unterrichtssprache zu lernen). In Abgrenzung zu sprachlichen Lernhilfen zur Erreichung ausschließlich sprachlicher Lernziele ist das wesentliche Element sprachsensiblen Unterrichts der gezielte Einbezug sprachlicher Mittel und kommunikativer Verfahren, um die fachliche Lernleistung von Schülerinnen und Schülern zu steigern und dadurch fachliches und sprachliches Lernen zu verknüpfen.

Die Unterstützung erfolgt über den Einsatz unterschiedlicher Methoden, die einzelne Vorgehensweisen darstellen und Medien bzw. sprachliche Werkzeuge beinhalten können (z. B. den Einsatz von Wortschatzkarten in der Einzelarbeit). Ein sprachsensibler Unterrichtsansatz umfasst dabei mehr als die einmalige Anwendung einer einzelnen sprachsensiblen Methode. Als Gesamtorientierung didaktisch-methodischen Handelns (Tenberg, 2006, S. 188) integriert er eine Variation von (mindestens zwei) methodischen Vorgehensweisen. Sprachensible Unterrichtsansätze enthalten konkrete Vorgaben zur Gestaltung von bestimmten Unterrichtsinteraktionen und beziehen z. T. auch allgemeine didaktische Prinzipien zur ganzheitlichen Unterrichtsplanung ein.

Im Folgenden werden die international bekanntesten sprachsensiblen Unterrichtsansätze skizziert. Zu ihnen zählen (sprachfokussierendes) *Scaffolding* (Gibbons, 2002) und das *Sheltered Instruction Observation Protocol (SIOP)* (Echevarría, Richards-Tutor, Canges & Francis, 2011; Echevarría, Vogt & Short, 2000). Auch der *Translanguaging*-Ansatz nach García (2009) wird in der Mehrsprachigkeitsdidaktik mittlerweile sprachsensibel umgesetzt.

2.2. International bekannte sprachensible Unterrichtsansätze

Scaffolding und Scaffolding-ähnliche Ansätze

Der Begriff *Scaffolding* wurde zunächst im Kontext von Eltern-Kind-Interaktionen verwendet (Wood, Bruner & Ross, 1976). Dabei standen anfangs weniger sprachliche Elemente im Vordergrund, sondern die Aufgabenbewältigung bzw. die Kontrolle des Frustrationsgrades bei der Bear-

beitung von Aufgaben durch ein Kind. Hier werden durch den erwachsenen Interaktionspartner Teilschritte ausgeführt, um dem Kind eine Hilfestellung zu geben, Aufgaben eigenständig zu lösen (Bruner, 1978, S. 19). Diese Strategie lässt sich insofern auf den Sprachaneignungsprozess übertragen, als einem Kind sprachliche Unterstützung solange angeboten wird, bis es zum selbständigen sprachlichen Handeln in der Lage ist. In der Psychologie wird das Konzept des *Scaffolding* entsprechend stärker mit kognitionspsychologischen Aspekten in Verbindung gebracht (z. B. im Sinne konstruktivistisch orientierter Lehr-Lernansätze) und somit stehen nicht-sprachliche Elemente stärker im Vordergrund.

Hier lassen sich weitere Ansätze nennen, die dem beschriebenen *Scaffolding*-Ansatz nach Bruner ähneln, so z. B. in den Bildungswissenschaften der *Content Enhancement*-Ansatz (Bulgren, Lenz, Schumaker, Deshler & Marquis, 2002), in den Kognitionswissenschaften und Bildungswissenschaft der *Cognitive Apprenticeship*-Ansatz (Chiu, Chou & Liu, 2002) sowie in der Psychologie (und den Bildungswissenschaften) die Ansätze des selbstregulierten Lernens (Azevedo, Cromley, Winters, Moos & Greene, 2005), des *Dynamic Assessment* (Swanson & Lussier, 2001) und der *Involvement Strategies* (Turner, Meyer, Cox, Logan, DiCintio & Thomas, 1998).

Demgegenüber stehen in der Zweitspracherwerbsdidaktik sprachliche Elemente stärker im Vordergrund. Dem Konzept des *Scaffolding* wird eine praktische Bedeutung für das sprachliche Lernen zugeschrieben, über das Lernende fachliche Hürden im Unterricht nehmen und Aufgaben selbstständig lösen können. Hier ist besonders das Modell nach Gibbons (u. a. Gibbons, 2002; 2009) hervorzuheben, das im Rahmen des sogenannten *Makro-Scaffolding* eine Sprachstandsermittlung sowie eine Bedarfsanalyse noch vor der Unterrichtsplanung bzw. Auswahl passender, individueller sprachlicher Unterstützungen für den Unterricht vorsieht. Analog zum Erstspracherwerb lässt sich in der Unterrichtsinteraktion (*Mikro-Scaffolding*) der Auf- und Abbau von sprachlichen Hilfen als zentrales Element des Modells von Gibbons herausstellen. Mit der Verbindung von fachlichen und sprachlichen Lernzielen wird aber auch der Lernprozess fokussiert und somit nicht-sprachliche Aspekte der Aufgabenbewältigung bzw. des selbständigen Problemlösens berücksichtigt.

Sheltered Instruction Observation Protocol (SIOP)

SIOP stellt ein Konzept zum integrierten Fach- und Sprachenlernen dar und ist in ein zertifiziertes Lehrerfortbildungsprogramm in den USA integriert.¹ Das *SIOP*-Modell eignet sich zur Unterrichtsplanung, -gestaltung sowie -evaluation in allen Fächern. Ähnlichkeiten zum *Scaffolding*-Ansatz finden sich bei der Zielsetzung von *SIOP*, in der die aufgaben- und lernzielorientierte Vermittlung der Fachinhalte bei gleichzeitiger Förderung der (fach)sprachlichen Kenntnisse im Vordergrund steht (s. Beese, 2010). *SIOP* richtet sich zunächst primär an Schülerinnen und Schüler, für die die Unterrichtssprache nicht Erst-, sondern Zweitsprache ist.² Gleichwohl findet *SIOP*-basierter Unterricht nicht in gesonderten Unterrichtseinheiten für Zweitsprachenlernende statt, sondern wird in

¹ <https://www.pearsonschool.com/index.cfm?locator=PS2rWo> [letzter Zugriff 12.6.2019]

² Siehe aber Beese (2010, S. 1) zum Einsatz von *SIOP* im regulären Fachunterricht: „Ein Unterricht, in dem gezielt fachliche Inhalte und Fertigkeiten zusammen mit den hierfür erforderlichen fachsprachlichen Kompetenzen entwickelt werden, kommt letztlich aber allen Schülerinnen und Schülern zu Gute. *SIOP*-basierter Unterricht ist daher kein Spezialunterricht für Zweitsprachenlernende, sondern regulärer Fachunterricht, gehalten von den entsprechenden Fachlehrerinnen und -lehrern.“

den regulären Fachunterricht integriert (Beese, 2010).

Der Unterricht wird auf Grundlage eines festgelegten Kriterienkatalogs (*Sheltered Instruction Observation Protocol*) durchgeführt, mit dessen Hilfe fachliche und sprachliche Lernziele bei der inhaltlichen und methodischen Unterrichtsplanung berücksichtigt werden. Das Protokoll setzt sich aus acht Komponenten mit insgesamt 30 Unterpunkten zusammen. Zu den acht Komponenten (1) Unterrichtsvorbereitung, (2) Aufbau von Hintergrundwissen, (3) verständlicher Input, (4) Strategien, (5) Interaktion, (6) Anwendung, (7) Umsetzung der Stunde und (8) Wiederholung und Leistungskontrolle) gibt es eine ausführliche Beschreibung, die anhand von beispielhaften Unterrichtsanalysen konkretisiert wird. Ergänzend dazu existiert eine Reihe von frei zugänglichen Unterrichtseinheiten, welche die praktische Umsetzung des Protokolls vereinfachen sollen.

Translanguaging

Der *Scaffolding*-Ansatz wird mittlerweile auch in der Mehrsprachigkeitsdidaktik („multilingual pedagogy“) aufgegriffen (Bredthauer, 2019). Im *Translanguaging*-Ansatz nach García (2009) wird von einer „Gesamtsprachigkeit“ der Lernenden ausgegangen (Gantefort & Oroquieta Sánchez, 2015, S. 25) und die vorhandenen mehrsprachigen Ressourcen von Schülerinnen und Schülern in der Unterrichtskommunikation berücksichtigt (García & Flores, 2012, S. 243). In einem sprachsensiblen *Translanguaging*-Unterrichtskonzept werden im Sinne des Makro-*Scaffolding* sprachliche und fachliche Lernziele miteinander verbunden (Henderson, 2017, S. 301) und ausgehend von den sprachlichen Voraussetzungen der Schülerschaft Unterstützungen ausgewählt (Daniel, Jiménez, Pray & Pacheco, 2017, S. 14). Als Unterstützung bei der Aufgabenlösung werden die individuellen Familiensprachen (L1) im Unterricht eingesetzt. Die L1 wird entweder einzeln im Selbstgespräch oder dialogisch innerhalb einer Gruppe verwendet und ermöglicht das Verständnis und die Entwicklung komplexer Sachverhalte („*languaging*“), die Grundlage für die Erstellung eines mündlichen oder schriftlichen Produkts in der Unterrichtssprache sind (Swain & Lapkin, 2013, S. 122f.). Der Abbau dieser Unterstützung erfolgt durch ein von der Lehrkraft gesteuertes Bestärken, die Unterrichtssprache als Interaktionsmedium zu verwenden (Swain & Lapkin, 2013, S. 122). *Translanguaging* wird zudem als durchgehendes Konzept angesehen, das je nach (neuem) Bedarf im Sinne der Zone der nächsten Entwicklung nach Vygotskij (1975) eingesetzt wird (Swain & Lapkin, 2013, S. 122).

Ein nach dem *Translanguaging*-Ansatz geplanter und durchgeführter Unterricht unterscheidet sich somit von dem *Scaffolding*-Ansatz nach Gibbons allein in der Nutzung der Ressource Mehrsprachigkeit: „*Sprachsensibler Fachunterricht*‘ ist demnach nicht einsprachig, sondern mehrsprachig zu denken“ (Gantefort & Oroquieta Sánchez, 2015, S. 25).

2.3. Abgrenzung von verwandten sprach- und fachintegrierten Unterrichtsansätzen

Neben den in 2.2 dargestellten Ansätzen sehen auch die Ansätze des *Content and Language Integrated Learning (CLIL)* und der *Content-Based Instruction (CBI)* eine Verknüpfung von fachlichem und sprachlichem Lernen vor (zum Verhältnis der beiden Konzepte s. Cenoz, 2015). Die Zielsprache in *CLIL*- und *CBI*-Konzepten ist in der Regel jedoch eine andere als die Amts-/Unterrichtssprache. Diese kann eine Fremdsprache (z. B. Englisch in Deutschland), aber auch eine Zweit- (Französisch in Kanada, Baskisch in Spanien) oder Herkunftssprache sein. Klassische Beispiele für den *CBI*-

Ansatz sind Immersionsprogramme in Kanada, in denen Schülerinnen und Schüler mit Englisch als Erstsprache in ihrer Zweitsprache Französisch unterrichtet werden. Der *CLIL*-Ansatz bleibt zudem häufig auf ein Fach und den Fremdsprachenkontext beschränkt; entweder liegt der Fokus stärker auf der Vermittlung von inhaltlich-fachlichen Kompetenzen durch eine weitere (Fremd-)Sprache oder der Fokus liegt auf der Vermittlung einer weiteren (Fremd-)Sprache über den fachlichen Inhalt (Massler, Stotz & Queisser, 2014; Rösch, 2013, S. 32).

Seit der Entstehung von *CLIL* lag und liegt der Fokus bis heute nicht auf der Ausbildung sprachlicher Fähigkeiten und Fertigkeiten in der Amts-/Unterrichtssprache (z. B. Deutsch in Deutschland), sondern auf der Ausbildung von fremdsprachlichen Fähigkeiten (z. B. Englisch in Deutschland). Bei *CBI* liegt der Fokus auf dem Ausbau von Kompetenzen in einer Zweitsprache, z. B. des Französischen in Kanada oder des Baskischen in Spanien, unter der Voraussetzung, dass die Erstsprache in der Gesellschaft (oder Familie) gefestigt genug ist (z. B. Englisch in Kanada oder Spanisch in Spanien). Deshalb stehen *CLIL* und *CBI* nicht im Fokus dieses Reviews.

3. Forschungsstand

Sprachsensibles Lernen wird seit den 90er-Jahren zunehmend wissenschaftlich aufgegriffen (z. B. Macken-Horarik, 1996; 1998). Die Forschung zu sprachsensiblen Unterricht findet bis heute insbesondere im anglosächsischen Raum statt (z. B. Gibbons, 2002; Williams, 1994). In diesem multidisziplinären Forschungsfeld gibt es eine Reihe an theoretischen und konzeptuellen Arbeiten (García & Wei, 2018; Leisen, 2011). Die empirische Forschung befindet sich jedoch bislang noch in einem eher frühen Entwicklungsstadium.

Primärstudien in der Wirkungsforschung

Für die angestrebte Untersuchung ist die Frage der Wirkung von sprachsensiblen Ansätzen relevant. Es gibt eine Reihe von empirischen Studien, die die Wirkung solcher Ansätze prüfen und dabei auch positive Effekte auf Fachlernen nahelegen – z. B. Echevarría, Short & Powers (2006) sowie McIntyre, Kyle, Chen, Muñoz und Beldon (2010) zu *SIOP*, Schüler-Meyer et al. (2019) zum *Translanguaging*; McKenzie (2011) sowie Meneses, Escobar und Véliz (2018) zum *Scaffolding*, und auch zu weniger bekannten Ansätzen, wie zum *Thinking-Together-Program*, das den Nutzen gesprochener Sprache im Fachunterricht hervorhebt und Schülerinnen und Schüler dabei unterstützt, ihre sprachlichen Fähigkeiten als Werkzeug für gemeinschaftliches fachliches Denken einzusetzen (Mercer, Dawes, Wegerif & Sams, 2004). Bislang stehen solche Primärstudien aber vielfach noch unverbunden im Forschungsfeld. Daraus ergeben sich auf den ersten Blick auch widersprüchliche Aussagen, etwa dann, wenn die Untersuchung von McIntyre et al. (2010) eine positive Wirkung von *SIOP* auf die Lernleistung bestätigt, während Vidot (2011) keinen signifikanten Effekt feststellt.

Solche auf den ersten Blick divergente Ergebnisse können auch daraus hervorgehen, dass die primäre Wirkungsforschung eine starke Heterogenität in der Beschaffenheit ihrer Studien aufweist. So können schon die sprachsensiblen Interventionen selbst, die dort untersucht werden, sehr vielfältig sein, etwa im Hinblick auf die didaktischen Ebenen, auf denen sie sich verorten (z. B.

Ebene der Unterrichtsplanung oder der Unterrichtsinteraktion), im Hinblick auf den didaktisch-methodischen Ablaufplan sowie im Standardisierungsgrad des Vorgehens. Darüber unterscheiden sich Wirkungsstudien in den jeweiligen Untersuchungssettings sowie auch hinsichtlich der jeweils fokussierten Untersuchungspopulationen. Dasselbe trifft auf die untersuchten Outcomes zu, die im Feld empirischer Wirkungsforschung zu sprachsensiblen Unterricht je nach Erkenntnisinteresse durch verschiedene Indikatoren des sprachlichen und/oder des fachlichen Lernens belegt werden. Um die Komplexität sprachsensibler Ansätze abzubilden, sind Mixed-Method-Ansätze und Multilevel-Analysen sinnvoll und werden von Forschenden auch angewendet (z. B. Echevarría et al., 2011). Daher sind es nicht zuletzt auch die unterschiedlichen Studiendesigns der Primärstudien, die eine Vergleichbarkeit der Studien und damit der Studienergebnisse einschränken können. Aus diesem Grunde können die Ergebnisse einzelner Primärstudien nicht ohne Weiteres aggregiert werden (zur ausführlichen Diskussion des Forschungsstands s. Höfler & Woerfel, i. Vorb.a).

Systematische Übersichtsstudien in der Wirkungsforschung

Einzelne nicht-systematische Literaturüberblicke fassen je einen kleinen Ausschnitt der empirischen Studienlage zu sprachsensiblen Unterrichtsansätzen zusammen (Echevarría, 2012; Echevarría & Short, 2011). Überblicksarbeiten, welche die Evidenz zur Wirkung sprachsensibler Interventionen systematisch zusammenfassen (systematische Reviews) und auf die Herausforderung der Aggregation heterogener Primärstudien angemessen reagieren können, liegen bislang nur wenige vor. Das hat auch eine Vorabuntersuchung des Projektteams bestätigt, die mittels strukturierter Abfragen in der Abstract- und Zitationsdatenbank *Scopus* sowie mittels manueller Suche in den speziell für systematische Reviews ausgelegten Online-Bibliotheken vom EPPI-Centre, dem What Works Clearinghouse (WWC) und der Campbell Collaboration durchgeführt wurde. Eine systematische Überblicksarbeit liefert etwa ein „Intervention Report“ (What Works Clearinghouse, 2009; 2013), der nach der Wirkung von *SIOP* bei Schülerinnen und Schülern fragt, die Englisch als Zweitsprache erlernen. Der Report trifft jedoch keine zusammenfassende Wirkungsaussage, weder in der Ursprungsarbeit im Jahr 2009 noch im späteren Update von 2013. Grund dafür ist die fehlende Passung der gefundenen *SIOP*-Studien hinsichtlich der gesetzten Inklusionskriterien und/oder der Evidenzstandards.

Zu *Scaffolding* konnten zwei systematische Reviews identifiziert werden (Swanson & Lussier, 2001; van de Pol, Volman & Beishuizen, 2010). Diese haben jedoch keinen Bezug zum sprachsensiblen Unterrichten. Sie verstehen *Scaffolding* allgemein als Lernunterstützung durch die Bereitstellung bestimmter intellektueller Aufgaben und nicht als Unterstützung durch sprachbezogene Lernhilfen, bei denen sprachliches Lernen zentrales Instrument in der Vermittlung von Lehrinhalten ist (zur Unterscheidung s. auch Michell & Sharpe, 2005, S. 32). Deshalb treffen sie den Fokus der vorliegenden Untersuchung nicht.

Systematische Übersichtsarbeiten, die mehrere sprachensible Unterrichtsansätze zusammenfassen, oder gar ihre gesamte Vielfalt einbeziehen, liegen bislang nicht vor.

4. Ziel, Art und Nutzen der Untersuchung

Ziel der Untersuchung

Ziel der Untersuchung ist eine Forschungssynthese zur Wirkung sprachsensibler Unterrichtsansätze im Primar- und Sekundarbereich. Wirkung („effectiveness“) meint hier das Vermögen von sprachsensiblen Interventionen, auf das Erleben und Verhalten bezogene, schulisch relevante Lernleistungen unter Realbedingungen in Lehr-Lern-Kontexten positiv zu beeinflussen. In diesem Zusammenhang soll folgende Frage beantwortet werden:

Sind sprachensible Unterrichtsansätze im Primar- und Sekundarbereich nachweislich wirkungsvoll?

Art der Untersuchung

Mittels eines systematischen Reviews wird dafür systematisch die Evidenz der Studien zusammengetragen, die die Wirkung bzw. Wirksamkeit³ von sprachsensiblen Unterrichtsansätzen prüfen. In Folge der Qualitätsbewertung dieser Studien lassen sich Aussagen darüber treffen, inwiefern sprachensible Unterrichtsansätze im Primar- und Sekundarbereich wirkungsvoll sind. Das Review zielt auf eine narrative Forschungssynthese (Popay, Roberts, Sowden, Petticrew, Arai, Rodgers & Britten, 2006; Snilstveit, Oliver & Vojtkova, 2012), die aufgrund der Heterogenität der Studienlage, vor allem im Hinblick auf die Operationalisierung und Messung der abhängigen und unabhängigen Variablen in Wirkungsstudien zu sprachsensiblen Unterricht (s. 3), sinnvoll ist.

Hierfür werden nach festgelegten Einschlusskriterien empirische Studien systematisch ermittelt, relevante Untersuchungen identifiziert und deren Qualität bewertet. Anschließend werden die Studien in der narrativen Forschungssynthese inhaltlich zueinander in Beziehung gesetzt und diskutiert.

Nutzen der Untersuchung

Das Review leistet einen Beitrag dazu, dass Befunde einzelner Studien im wissenschaftlichen und fachpolitischen Kontext nicht mehr als Maß für die generelle Wirkung einer Bildungsmaßnahme, sondern differenziert unter Berücksichtigung der umfassenden Erkenntnislage angeführt werden. In der Lehrerbildung können somit sprachensible Konzepte nicht nur beispielbasiert, sondern auch auf der Grundlage nachgewiesener Wirkung vermittelt werden (und die sprachsensiblen

³ Die Begriffe Wirkung und Wirksamkeit werden in der Anwendung als forschungsmethodische Konzepte hier nach Pant (2014) definiert, d. h., die Untersuchung der Wirkung („effectiveness“) von Interventionen fokussiert Interventionseffekte unter Realbedingungen; die Untersuchung der Wirksamkeit („efficacy“) von Interventionen fokussiert Interventionseffekte unter Idealbedingungen. Die Erforschung pädagogischer Interventionen findet im Regelfall unter Realbedingungen statt (vgl. Pant, 2014, S. 95). Zwar können Interventionsstudien unter (eingeschränkt) kontrollierten und standardisierten Bedingungen (z. B. randomisierte Kontrollstudien, d. h. RCTs), insbesondere bei genau definierten pädagogischen Interventionen, durchgeführt werden (vgl. Pant, 2014, S. 95). Im Review wird aber auf die Effekte sprachsensibler Ansätze im schulischen Alltag geschlossen, d. h. auf ihre Wirkung („effectiveness“) unter Realbedingungen.

Methoden entsprechend gewichtet werden). Wenn vermehrt sprachensible Ansätze eingesetzt werden, deren Wirkung empirisch gesichert ist, wird der Zuwachs sprachlicher Kompetenzen der Schülerinnen und Schüler befördert. Zuletzt kann die Wissenschaft ermittelte Forschungslücken für die Initiierung neuer Forschungsprojekte nutzen.

Verwertungsplan

Die Ergebnisse werden in einschlägigen deutsch- und englischsprachigen Fachzeitschriften publiziert und in Transferformaten für Bildungsverwaltung und -praxis aufbereitet.

5. Methodik

5.1. Leitende Untersuchungsfrage

Um die Forschungsfrage für die Untersuchung zu schärfen, wird das PICOS-Schema (Liberati, Altman, Tetzlaff, Mulrow, Gøtzsche, Ioannidis, Clarke, Devereaux, Kleijnen & Moher, 2009) angewendet, das seinen Ursprung in der evidenzbasierten Medizin hat und genutzt wird, um eine recherchierbare Fragestellung zu formulieren. Mit diesem Schema sollen die Kategorien

- Untersuchungspopulation (participants: P),
- Intervention (intervention: I),
- Vergleich (comparator: C),
- Outcome (outcome: O) und
- Studiendesign (study design: S)

entsprechend des Erkenntnisinteresses näher bestimmt werden. Die für dieses Review relevante Fragestellung nach PICOS lautet wie folgt:

Belegen Wirkungsstudien zu sprachsensiblen Unterrichtsansätzen (S), dass Schülerinnen und Schüler des Primar- und Sekundarbereichs (P), die sprachsensibel unterrichtet werden (I), bessere sprachliche und/oder fachliche Lernziele (O) erreichen, als Schülerinnen und Schüler, die nicht sprachsensibel unterrichtet werden (C)?

Sprachensible Unterrichtsansätze haben das Ziel, sprachliches Lernen vor allem zum Zweck der Erreichung fachlicher Lernziele zu unterstützen. Da der Forschungsstand zeigt, dass Wirkungsstudien zu sprachsensiblen Unterricht das Outcome sowohl mit Indikatoren sprachlichen und/oder fachlichen Lernens operationalisieren (vgl. Höfler & Woerfel, i. Vorb.a), ist die PICOS-Kategorie (O) entsprechend definiert.

Um die Fragestellung zu beantworten, braucht es folglich empirische Studien, die prüfen, ob sprachensible Unterrichtsansätze tatsächlich einen Lernzuwachs bewirken, wie im Fachdiskurs aufgrund von theoretischem Erklärungswissen angenommen wird (s. 1).

5.2. Inklusionskriterien

Die Einschlusskriterien zur Inklusion von Studien beziehen sich auf die PICOS-Kategorien der Untersuchungsfrage (s. 5.1) und definieren diese spezifischer. Mithilfe des PICOS-Schemas werden alle Kategorien abgedeckt, die in systematischen Reviews als obligatorisch für die Definition von Inklusionskriterien angesehen werden (Khan, Kunz, Kleijnen & Antes, 2004; Petticrew & Roberts, 2006, S. 75).

Unter Einbezug des Forschungsstandes, der Informationen zur Beschaffenheit von relevanten Wirkungsstudien liefert (s. 3), wurden im ersten Schritt vorläufige Inklusionskriterien entwickelt. Um die Kriterien zu schärfen, fand in einem zweiten Schritt die Pilotierung der vorläufigen Kriterien anhand von Prototypenstudien statt. Die Studien, die Prototypen für die Beantwortung der Untersuchungsfrage darstellen, bilden unterschiedliche sprachensible Interventionen ab, die in Wirkungsstudien fokussiert werden - sowie auch eine Vielfalt an Studiendesigns (s. 3). Sie wurden mittels einer freien Internetrecherche und einer strukturierten Abfrage in der Literaturlieferantenbank ERIC ermittelt und hinsichtlich der vorläufigen Inklusionskriterien ausgewertet. Folgende Prototypenstudien wurden identifiziert: Echevarría, Short & Powers (2006); Härtig & Stosik (2015); Schüler-Meyer et al. (2019); Short, Fidelman & Lougoit (2012); Vidot (2011). Diese internationalen Studien beziehen sich auf die international bekannten sprachsensiblen Ansätze *SIOF* und *Translanguaging*, aber auch auf sogenannte *No-Name*-Ansätze, wobei einige direkt auf der Ebene der Unterrichtsinteraktion ansetzen und andere, ganzheitlicher, die Ebene der Unterrichtsplanung einbeziehen. Ferner decken die Prototypen experimentelle, quasi-experimentelle und Beobachtungsstudien ab, nutzen rein quantitative oder Mixed-Method-Designs sowie quer- und Längsschnitterhebungen, mit unterschiedlich langer Studiendauer und unterschiedlich vielen Messzeitpunkten.

Im Folgenden sind die im Rahmen der Pilotierung geschärften Inklusionskriterien aufgelistet (vgl. auch Tabelle 1).

Population (P)

Eingeschlossen werden Studien, die Schülerinnen und Schüler im Alter von 5-20 Jahren untersuchen und dadurch der mit Deutschland vergleichbaren Primar- und Sekundarstufe zuzuordnen sind⁴. Sie bilden die Grundgesamtheit des vorliegenden Reviews.

Die Studien fokussieren Personen, die zum Zeitpunkt der Untersuchung keine diagnostizierte, klinisch relevante Störung oder Beeinträchtigung aufweisen, die eine spezifische sonderpädagogische oder therapeutische Unterstützung beim Lernen erfordert.

Intervention (I)

Es werden Studien eingeschlossen, die (mindestens) eine Lehrintervention fokussieren, die explizit darauf abzielt, durch die planvolle und reflektierte Einbindung oder Förderung der Sprache der Lernenden im Fachunterricht ihr fachliches Lernen zu unterstützen. Um das damit angesprochene

⁴ Schülerinnen und Schüler an berufsbildenden Schulen oder Universitäten sind damit ausgeschlossen.

Merkmal der Sprachsensibilität (s. 2.1) zu erfüllen, müssen die Kriterien I1 und I2 zutreffen:

(I1.) Der Fachunterricht findet in der jeweiligen Amtssprache, d. h. in der üblichen Unterrichtssprache statt.

(I2.) Die Lehrintervention forciert die gezielte Einbindung von sprachlichen Mitteln der Lernenden (z. B. ihr sprachliches Wissen oder Handlungsrepertoire, wie Wortschatz, Lese- oder Argumentationstätigkeit) und von kommunikativen Verfahren⁵, um die fachliche Lernleistung zu steigern.

Es werden nur Studien eingeschlossen, bei denen Sprache nicht der zentrale Unterrichtsgegenstand ist (z. B. Deutsch-Unterricht, DaZ-Unterricht, Fremdsprachen-Unterricht). Auch wenn solcher Unterricht sprachsensibel sein kann, fällt er nicht in den Rahmen der Untersuchung.

Die sprachsensiblen Interventionen setzen entweder direkt auf der Ebene der Unterrichtsinteraktion an und geben konkrete Handlungsvorgaben zur Gestaltung von einzelnen Unterrichtssituationen, oder sie sind ganzheitlicher angelegt und beziehen auch die Ebene der Unterrichtsplanung durch die Vorgabe allgemeiner didaktischer Handlungsprinzipien mit ein.

Da die hier relevanten Interventionen einen (sprachsensiblen) Unterrichtsansatz abbilden, umfassen sie mehr als die Anwendung einer einzelnen Methode⁶. Unterrichtsansätze sind Gesamtorientierungen didaktisch-methodischen Handelns, in denen ein begründeter Zusammenhang von Ziel-, Inhalt- und Methodenentscheidungen hergestellt wird (zur Diskussion s. Tenberg, 2006, S. 188).

Die Interventionen kommen im Rahmen von regulärem Schulunterricht, von natürlichen additiven Lehrangeboten (z. B. Nachhilfeunterricht) oder künstlich geschaffenen Lehrexperimenten zum Einsatz.⁷ Im vorliegenden Review wird dann geprüft (s. 5.5.2, 5.2), inwiefern auf institutionelle Lernsettings geschlossen werden kann.

Vergleich (C)

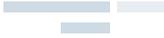
Es werden Studien integriert, die die Wirkung von sprachsensiblen Unterricht (s. Intervention) mit der Wirkung von Lehrsituationen vergleichen, in denen fachlich, aber nicht gezielt sprachsensibel unterrichtet wird.

Es werden nur Studien eingeschlossen, bei denen Sprache nicht der zentrale Unterrichtsgegenstand ist (z. B. Deutsch-Unterricht, DaZ-Unterricht, Fremdsprachen-Unterricht; s. PICOS-Kategorie Intervention).

⁵ Am Beispiel der Scaffolding-Forschung heißt das etwa, dass nur Studien relevant sind, bei denen sprachbezogene Hilfen (scaffolds) im Fokus stehen (s. 3) Studien zu CLIL oder Content Based Instruction (CBI) treffen den Fokus nicht (s. 2.3).

⁶ Eine ‚Methode‘ wird hier als eine einzelne, konkrete und auf gleiche Art reproduzierbare Vorgehensweise verstanden, die den Einsatz von Medien/Werkzeugen beinhalten kann, z. B. den Einsatz von Wortschatzkarten in der Einzelarbeit.

⁷ Inwiefern von Lehrsituationen, die nicht im Rahmen des regulären Schulunterrichts stattfinden, auf die im Review relevante Grundgesamtheit (Schülerinnen und Schüler des Primar- und Sekundarbereichs) geschlossen werden kann, wird im Rahmen der externen Validitätsbeurteilung von eingeschlossenen Studien geprüft (s. 5.5.2).



Diese Vergleichs-Lehrsituationen können im Rahmen von regulärem Schulunterricht, von natürlichen additiven Lehrangeboten (z. B. Nachhilfeunterricht) oder von künstlich geschaffenen Lehrexperimenten stattfinden.⁸

Outcome (O)

Relevant sind Studien, die das Outcome über Indikatoren des sprachlichen und/oder fachlichen Lernens erfassen. Das können Indikatoren der Fachsprache sein (z. B. Fachtermini der Mathematik), Indikatoren der sprachlichen, fachunabhängigen Lernleistung (z. B. Schreibleistung) oder auch konkret fachbezogene Leistungsindikatoren (z. B. Wissen über die Eigenschaften von Schwermetallen im Fach Chemie).

Studiendesign (S)

Es werden experimentelle, quasi-experimentelle und Beobachtungsstudien einbezogen.⁹

Die Studien treffen inferenzstatistische Aussagen zur Wirkung bzw. Wirksamkeit von sprachsensiblen Unterrichtsansätzen (s. Outcome) auf sprachliches und/oder fachliches Lernen (s. Outcome). Es handelt sich demnach um explanative empirische Studien, die entweder rein quantitativ angelegt sind oder einen Mixed-Methods-Ansatz aufweisen, sofern der beschriebene Wirkungszusammenhang mittels inferenzstatistischer Verfahren erfolgt.

Um Rückschlüsse auf die Wirkung bzw. Wirksamkeit von sprachsensiblen Unterrichtsansätzen zu ziehen, vergleichen die Studien die Daten der Stichprobe/n, die sprachsensibel unterrichtet wurde/n (s. Intervention), mit dem Datensatz von mindestens einer Stichprobe (Kontrollgruppe¹⁰), die dadurch definiert ist, dass sie nicht gezielt sprachsensibel unterrichtet wurde (s. **Vergleich**).

Studien, die im Querschnitt Daten erheben bzw. analysieren, erfassen bestimmte Variablen bei einer einzelnen Stichprobe einmalig, zu einem einzigen Zeitpunkt. Längsschnitterhebungen dagegen erfassen bestimmte Variablen bei ein und derselben Stichprobe mindestens zu zwei Zeitpunk-

⁸ s. Fußnote 7

⁹ Bei einer experimentellen Studie (randomisierte, kontrollierte Studie: RCT) wird mit exakt vergleichbaren Gruppen gearbeitet, die durch eine zufällige Zuordnung der Untersuchungsobjekte zu den Gruppen gebildet werden (Randomisierung). Die unabhängige Variable wird systematisch bei den jeweiligen Gruppen eingesetzt. Bei einer quasi-experimentellen Studie (nicht-randomisierte kontrollierte Studie: NRCT) wird dagegen auf Gruppen zurückgegriffen, die nicht zufällig zusammengestellt, sondern vorgefunden oder anderweitig gebildet wurden (keine Randomisierung). Diese werden jedoch so systematisch unterschiedlich behandelt wie im echten Experiment (Döring & Bortz, 2016, S. 193f.). Eine Beobachtungsstudie (auch „nicht-experimentelle Studie“) greift auf vorgefundene Gruppen zurück (keine Randomisierung) und betrachtet deren vorgefundene Unterschiede in der unabhängigen Variable (keine experimentelle Variation, d. h. keine aktive Manipulation der unabhängigen Variable/n) (Döring & Bortz, 2016, S. 194).

¹⁰ Eine Kontrollgruppe im hier verstandenen Sinne kann sowohl in (Quasi-)Experimenten als auch in Beobachtungsstudien eine Rolle spielen. Wenn in (quasi-)experimentellen Studien Interventionen im Rahmen der Untersuchung eingesetzt werden, die den Inklusionskriterien nach PICOS C (s. Vergleich) entsprechen, können Kontrollgruppen auch Interventionsgruppen sein.

ten, was einen Prä-Post-Vergleich möglich macht (Döring & Bortz, 2016, S. 210f.). In der Unterrichtsforschung ist ein Prä-Post-Vergleich nicht immer umsetzbar und manchmal wird mittels eines „independent measures design“ bzw. „between-subjects design“ grundsätzlich auf eine Messwiederholung verzichtet und sich auf die Post-Messung beschränkt. Auch wenn Studien mit querschnittlichem Erhebungsdesign Einschränkungen in der internen Validität aufweisen, können sie Hinweise auf die Wirkung von sprachsensiblen Interventionen liefern und Erkenntnisse aus längsschnittlichen Erhebungen stützen. Vor allem vor dem Hintergrund des Review-Ziels, die vorliegenden Daten zur Wirkung von sprachsensiblen Unterrichtsansätzen zusammenzutragen, werden deshalb Studien mit querschnittlichem Erhebungsdesign, neben solchen mit längsschnittlichen Erhebungen, in das Review integriert, sofern sie die Inklusionskriterien (s. 5.2) erfüllen. Das schließt klassische Querschnittstudien (verstanden als korrelative Studien mit einem Messzeitpunkt ohne Kontrollgruppe) aus. Die eingehende Qualitätsbeurteilung der integrierten Studien erfolgt zu einem späteren Zeitpunkt des Review-Verfahrens (s. 5.5.2).

In Tabelle 1 sind die Inklusionskriterien der jeweiligen PICOS-Kategorien nochmals zusammengefasst.

Tabelle 1 Zusammenfassende Übersicht zu den Inklusionskriterien

PICOS-Kategorie	Inklusionskriterien
Population (P)	<ul style="list-style-type: none"> ▪ Schülerinnen und Schüler im Alter von 5-20 Jahren ▪ ohne spezifischen sonderpädagogischen/therapeutischen Unterstützungsbedarf im Lernen
Intervention (I)	<ul style="list-style-type: none"> ▪ gezielte Einbindung sprachlicher Mittel und kommunikativer Verfahren zur Steigerung der fachlichen Lernleistung ▪ Zielsprache ist die jeweilige Amtssprache ▪ didaktische Vorgaben entweder ausschließlich auf Ebene der Unterrichtsinteraktion oder zusätzlich auf Ebene der Unterrichtsplanung ▪ bildet Unterrichtsansatz ab ▪ findet im Rahmen von regulärem Schulunterricht, von natürlichen additiven Lehrangeboten oder künstlich geschaffenen Lehrexperimenten statt ▪ findet im Rahmen von Fachunterricht statt, bei dem Sprache nicht selbst der zentrale Unterrichtsgegenstand ist
Vergleich (C)	<ul style="list-style-type: none"> ▪ fachrelevantes Lernen ohne gezielt angewendeten sprachsensiblen Ansatz ▪ findet im Rahmen von regulärem Schulunterricht, natürlichen additiven Lehrangeboten oder künstlich geschaffenen Lehrexperimenten statt ▪ findet im Rahmen von Fachunterricht statt, bei der Sprache nicht

	selbst der zentrale Unterrichtsgegenstand ist
Outcome (O)	<ul style="list-style-type: none"> ▪ Indikatoren sprachlichen und/oder fachlichen Lernens
Studiendesign (S)	<ul style="list-style-type: none"> ▪ experimentelles, quasiexperimentelles oder Beobachtungsdesign ▪ inferenzstatistische Aussage zur Wirkung von sprachsensibler Intervention (s. I) auf das Outcome (s. O) ▪ mindestens eine Vergleichsgruppe (s. C) ▪ längsschnittliches oder auch querschnittliches Erhebungsdesign (sofern Post-Interventionsmessung)

5.3. Systematische Studiensuche

Die zur Beantwortung der Forschungsfragen erforderlichen einschlägigen Studien, die die Grundlage der narrativen Synthese bilden, werden über die im Folgenden dargestellte Suchstrategie ermittelt. Bei der Entwicklung der Suchstrategie und der anschließenden Suchabfragen orientieren wir uns an den Richtlinien der Cochrane-Collaboration zur Erstellung von systematischen Reviews im Gesundheitswesen (O'Connor, Green & Higgins, 2011; Lefebvre, Mannheimer & Glanville, 2011), Richtlinien der Campbell Collaboration (Kugley, Wade, Thomas, Mahood, Jørgensen, Hammerstrøm & Sathe, 2017) und des EPPI-Centre (Gough, Oliver & Thomas, 2017) zur Erstellung von systematischen Reviews im Bildungswesen sowie den PRISMA-Richtlinien, welche zur Dokumentation von systematischen Reviews und Meta-Analysen im Gesundheitswesen entwickelt wurde (Liberati et al., 2009).

5.3.1. Suchbegriffentwicklung

Zunächst wird eine Stichwortsammlung¹¹ angelegt, die als Grundlage für die Entwicklung von Suchanweisungen in elektronischen Datenbanken dient. Um für die Suchanweisungen möglichst passende Stichwörter zu finden, werden diese auf verschiedene Weise ermittelt (s. Twente & Worfel, 2020 für eine ausführlichere Darstellung). Alle ermittelten Stichwörter werden in Tabelle 5-Tabelle 8 aufgeführt.

1. **Auswahl von Stichwörtern durch das Review-Team:** Gemeinsam mit Wissenschaftlerinnen und Wissenschaftlern mit inhaltlicher Expertise sowie Vertreterinnen und Vertretern der pädagogischen Praxis werden deutsch- und englischsprachige Stichwörter ausgewählt.
2. **Erweiterung der Stichwörter mittels Text mining-Verfahren:** Damit die Auswahl der Stichwörter für die Suchanweisungen in den festgelegte elektronischen Datenbanken nicht ausschließlich auf subjektiven Entscheidungen von Einzelpersonen beruhen, werden Text-mining Verfahren

¹¹ Unter Stichwort verstehen wir („freie“) Suchbegriffe, die von Autorinnen und Autoren verwendet oder festgelegt werden (z. B. bei einer freien Suchabfrage oder bei der Festlegung eigener Suchbegriffe bei einem publizierten Aufsatz). In Abgrenzung dazu stellen Schlagwörter („feste“) Suchbegriffe aus von Bibliothekarinnen und Bibliothekaren bzw. Informationswissenschaftlerinnen und -wissenschaftlern kontrolliertem und in der Regel fest definiertem Vokabular dar (z. B. die Einträge in Literaturdatenbanken zugewiesenen Suchbegriffe aus einem datenbankinternen Thesaurus).

ren angewandt¹², um die Stichwortauswahl zu optimieren. Hierfür werden auf der Grundlage der in 1. festgelegten Stichwörter Suchanweisungen in der allgemeinen Datenbank *Scopus* und den Fachdatenbanken *ERIC* und *FIS Bildung* durchgeführt: Die somit ermittelten Dokumente werden mit den bibliographischen Informationen sowie den Abstracts und den in der Datenbank vergebenen Stichwörtern/Schlüsselbegriffen exportiert und dienen als Datengrundlage für eine semi-automatisierte Optimierung der Stichwörter. Mittels der *R*-Pakete *litsearchr*¹³ (Grames, 2019) sowie *quanteda* (Benoit, Watanabe, Wang, Nulty, Obeng, Müller & Matsuo, 2018) werden in der freien Software- und Programmierumgebung *R* (R Core Team, 2013) automatisch Kollokationen der Stichwörter bestimmt, die besonders repräsentativ für den Inhalt einschlägiger Texte sind.

3. **Erweiterung der Stichwörter mittels einschlägiger Studien:** Zusätzlich zu 1. und 2. werden aus festgelegten einschlägigen Wirkungsstudien zum sprachsensiblen Unterricht Stichwörter aus Titeln und Abstracts extrahiert. Hierfür wurden 14 einschlägige englischsprachige und sieben deutschsprachige Dokumente aus dem Forschungsstand (inkl. der Prototypenstudien zur Schärfung der Inklusionskriterien) festgelegt.¹⁴ Die ermittelten Stichwörter werden dem Stichwortdatensatz hinzugefügt und dienen der Optimierung der Suchanweisungen in elektronischen Datenbanken.

5.3.2. Suchorte

Für die Ermittlung von Dokumenten wurden Suchorte festgelegt, deren Umfang so breit wie möglich (im Rahmen der Durchführbarkeit) und nötig (im Rahmen der thematischen Interdisziplinarität) ist, um eine Einseitigkeit der recherchierten Daten ausschließen zu können und das Verzerrungspotenzial niedrig zu halten (s. Tabelle 4 für eine ausführliche Darstellung).

Datenbanken zur Identifizierung konventionell publizierter Literatur

Die Auswahl der Datenbanken (s. u.) entspricht dem multidisziplinären Forschungsfeld (Sprachwissenschaften, Bildungswesen, Psychologie) sowie dem Umfang der enthaltenen Literatur > 300.000 Einträge. Die Datenbanken enthalten konventionell publizierte wissenschaftliche Literatur, z. B. in Form von Beiträgen in Sammelbänden oder Fachzeitschriften, von Monographien oder Dissertationen.

¹² Der Vorteil dieses Ansatzes ist, dass für die Suchabfragen potentielle Stichwörter auf der Grundlage eines größeren Datensatzes – von anderen Wissenschaftlerinnen und Wissenschaftlern verfassten Texten – hinsichtlich ihrer Bindungsstärke bzw. Relevanz innerhalb eines Kollokations-Netzwerks bestimmt werden und nicht ausschließlich (wie in Schritt 1) auf individuellen und somit eher verzerrten Entscheidungen von Einzelpersonen beruhen (vgl. Grames, Stillman, Tingley & Elphick, 2019).

¹³ Weitere Informationen und Updates s. Projektseite: <https://elizagrames.github.io/litsearchr/#/about> [letzter Zugriff am 19.1.2019].

¹⁴ Echevarría et al. (2011); Echevarría, Short & Powers (2006); Große (2015); Härtig & Stosik (2015); Hayes, Rueda & Chilton (2009); Honigsfeld & Cohan (2006); Lang (2016) McIntyre et al. (2010); McKenzie (2011); Meneses et al. (2018); Mercer et al. (2004); Merritt et al. (2016); Rous (2016); Schüler-Meyer et al. (2019); Schwippert et al. (2016), Short, Fidelman & Lougoit (2012); Sumfleth et al. (2013); Wessel & Prediger (2012); Vidot (2011); Whittier & Robinson (2007); Zwiers (2006).

- *Scopus*: Internationale Abstract- und Zitationsdatenbank, die bibliographische Einträge peer-reviewter Literatur aus unterschiedlichen Fachbereichen listet;
- *ERIC*: Fachdatenbank mit dem Schwerpunkt Pädagogik (überwiegend englischsprachige Literatur);
- *FIS Bildung*: Fachdatenbank, die alle Teilbereiche des Bildungswesens umfasst (u. a. englisch- und deutschsprachige Literatur);
- *PSYINDEX*: Fachdatenbank mit dem Schwerpunkt Psychologie, einschließlich psychologisch relevanter Aspekte aus Nachbardisziplinen (u. a. englisch- und deutschsprachige Literatur aus Deutschland, Österreich und der Schweiz).
- *PsycInfo*: Fachdatenbank für Psychologie und für die Psychologie relevante Gebiete, u. a. Linguistik und Erziehungswissenschaften;

Dankenbanken, Fachforen und Fachverteiler zur Identifizierung grauer Literatur

Die Datenbanken, Fachforen und Fachverteiler (s.u.) bieten Zugang zur Ermittlung grauer Literatur¹⁵. Sie dienen zur Identifizierung von nicht publizierten Manuskripten, vorläufigen Ergebnissen aus laufenden Wirkungsstudien, Arbeitspapieren und Berichten zu abgeschlossenen und laufenden Projekten (s. 5.3.3). (Bachelor- und Masterarbeiten werden nicht berücksichtigt.)

- *OpenGrey*: Europäische Datenbank, die neben Technologie, biomedizinische Wissenschaft und Wirtschaft auch Sozial- und Geisteswissenschaften umfasst (Listung erfolgt in Englisch).
- *The LINGUIST List*: Internationales Forum für Austausch von Informationen und Beiträgen aller Art im Fachbereich Linguistik
- *ResearchGate*: Internationales Netzwerk und Datenbank für Forscher aus allen Bereichen der Wissenschaft (Dokumentenserver für fachliche/wissenschaftliche Open-Access-Publikationen aller Art)
- *Academia.edu*: Dokumentenserver, Verteilstation und Selbstverwaltungsprogramm für Akademiker aus allen Bereichen der Wissenschaft (enthält internationale Open-Access Papers)
- Fachverteiler Gesellschaft für Angewandte Linguistik (GAL),
- Fachverteiler DaZ-Portal (Universität Jena)
- Fachverteiler Sektion Interkulturelle und International Vergleichende Erziehungswissenschaft (SIIVE),

¹⁵ Unter grauer Literatur wird hier Literatur verstanden, die nicht in den konventionellen wissenschaftlichen Publikationsorganen (Buch, Zeitschrift) veröffentlicht ist bzw. nicht durch das kommerzielle Verlagswesen kontrolliert wird, weil sie entweder i) in einem noch un abgeschlossenen Zustand ist (z. B. Ergebnisse nicht abgeschlossener Projekte), ii) die Vorgaben eines konventionellen Publikationsorgans (noch) nicht erfüllt hat, iii) privat/in-house ohne kommerzielle Absicht verlegt bzw. online veröffentlicht wird (vgl. Kugley, Wade, Thomas, Mahood, Jørgensen, Hammerstrøm & Sathe, 2017, S. 17; Petticrew & Roberts, 2006, S. 280; Paez, 2017, S. 233–234).

Fachzeitschriften

Die Fachzeitschriften (s.u.) sind im Hinblick auf ihre disziplinäre Passung und Qualitätskriterien (peer-reviewed) ausgewählt (s. Tabelle 4).

- Educational Psychology Review
- Fremdsprache Deutsch - Zeitschrift für die Praxis des Deutschunterrichts
- International Journal of Multilingualism
- International Multilingual Research Journal
- Journal of Educational Psychology
- Journal of Immersion and Content-Based Language Education
- JoWR (Journal of Writing Research)
- Language and Education
- Language Learning
- Latin and American Journal of Content and Language Integrated Learning (LACLIL)
- Learning & Instruction
- Praxis Deutsch
- Studies in Second Language Acquisition
- Teaching and Teacher Education
- The Journal of Educational Research
- The Language Learning Journal (LLJ)
- The Modern Language Journal
- Zeitschrift für Fremdsprachenforschung
- Zeitschrift für Erziehungswissenschaft
- Zeitschrift für Unterrichtswissenschaft
- Zielsprache Deutsch

5.3.3. Suchstrategie

Sprache der Dokumente

Wie der Forschungsstand zeigt (s. 33), gibt es Untersuchungen, die sich direkt auf das deutsche Primar- und Sekundarschulsystem beziehen – der Forschungsschwerpunkt zu sprachsensiblen Unterricht liegt aber im überwiegend englischsprachig geprägten Raum. Die systematische Suche nach Wirkungsstudien nimmt Dokumente in den Blick, die in den Sprachen Deutsch oder Englisch verfasst wurden, um die nationale und die internationale Forschung einzubeziehen.

Fokussierte Jahresspanne

Da sich in den 1990er Jahren die Auseinandersetzung mit Unterrichtspraktiken, die sprachliches und fachliches Lernen verknüpfen, zu einem eigenen Forschungszweig entwickelte (s. 3), wird der entsprechende Veröffentlichungszeitraum publizierter Dokumente festgelegt auf 1990 bis 2019¹⁶. Unveröffentlichte Dokumente, die z. B. im Rahmen grauer Literatur recherchiert werden, müssen im selben Zeitraum erstellt bzw. verfasst worden sein.

Vorgehen bei der Suche in Datenbanken für konventionell publizierte Literatur

Für eine strukturierte Abfrage in den Fachdatenbanken werden die in 5.3.1 festgelegten und den PICOS Kategorien zugewiesenen Stichwörter mit den Thesaurieinträgen bzw. Wortlisten (FIS Bildung) der ausgewählten Datenbanken abgeglichen. Daraufhin werden verschiedene Suchanweisungen mit Schlagwörtern und/oder Stichwörtern getestet. Ziel ist es, möglichst alle im Hinblick der Forschungsfrage relevanten Dokumente (Sensitivität) zu finden und gleichzeitig möglichst wenige Gesamttreffer pro Datenbank zu erzielen (Präzision). Um ein adäquates Verhältnis zwischen Sensitivität und Präzision einer Suchanweisung festzulegen, wird angestrebt, die maximale Trefferzahl der in einer Datenbank verzeichneten einschlägigen Studien zu ermitteln, bei gleichzeitig niedriger Gesamttrefferzahl.

Die Suche erfolgt durch die datenbankspezifischen Suchmasken oder durch manuelle Verwendung boolescher Operatoren (AND, OR, NOT) und wird z. B. durch Trunkierung der Stichwörter zusätzlich modifiziert. In der Kombination der Suchbegriffe wird Bezug auf Begriffe der PICOS-Kategorien Population, Intervention, Outcome und Studiendesign genommen (die Kategorie der Vergleichsgruppe bringt in diesem Zusammenhang keinen Mehrwert für die Suche).

Sprachsensible Ansätze werden folgendermaßen abgedeckt:

1. mittels Eigennamen oder einschlägigen Bezeichnungen sprachsensibler Interventionen (z. B. „sprachsensibler Unterricht“ oder „SIOP“) (d. h. ausschließlich über die PICOS-Kategorie Intervention) oder
2. mittels sprachbezogener Unterrichtsansätze (z. B. „bilingualer Unterricht“ oder „translanguaging“) kombiniert mit fachlichem Lernen (z. B. „Chemieunterricht“) bzw. mit Bildungserfolg (z. B. „schulischer Erfolg“) (d. h. mittels einer Kombination aus PICOS-Intervention und -Outcome)¹⁷ oder
3. mittels „scaffolding“ kombiniert mit sprachlichem Lernen (z. B. „Lesekompetenz“) und fachlichem Lernen bzw. Bildungserfolg (d. h. mittels einer Kombination aus PICOS-Intervention und

¹⁶ Nach dem Screening ist eine Nachsuche angedacht, die insbesondere die Möglichkeit für im Laufe des Reviews neu publizierte Studien, nach 2019, berücksichtigt.

¹⁷ Die Verknüpfung von Sprach- und Fachlernen ist zentrales Merkmal sprachsensibler Ansätze. Um sprachensible Interventionen ohne Eigennamen und fachdiskursferne Ansätze abzudecken, werden sprachbezogene Interventionen mit Fachlernen (z. T. auch abgedeckt über Indikatoren des allgemeinen Bildungserfolgs) kombiniert.

-Outcome)¹⁸

Die über diese Strategie ermittelten Dokumente werden zusätzlich präzisiert, indem die Stichwörter, die sprachensible Ansätze abbilden sollen, mit den PICOS-Kategorien Population und Studiendesign zwingend zusammen auftreten müssen. Gesucht wird nach folgendem Grundschema:

(Population AND Intervention (1 OR 2 OR 3) AND Studiendesign).

Sämtliche Suchanweisungen, Datum der Abfragen, verwendete Filter und Suchtrefferanzahl werden protokolliert.

Vorgehen bei der Suche in Datenbanken für graue Literatur

Zur Ermittlung relevanter grauer Literatur werden drei Wege verfolgt: Zunächst wird auf den akademischen Plattformen linguistlist.org und ResearchGate.net sowie über die Mailinglisten (s. 5.3.2) ein deutsch- und englischsprachiger *Call for preprints* geschaltet (s. S. 48), der sich an universitäre Einrichtungen und weitere wissenschaftliche Institutionen richtet und dazu aufruft, noch nicht publizierte Manuskripte und vorläufige Ergebnisse aus laufenden Wirkungsstudien zur Verfügung zu stellen. Zudem wird die Datenbank Opengrey nach unveröffentlichter Literatur durchsucht. Alle bis 30.11.2019 ermittelten Dokumente werden berücksichtigt.

Vorgehen bei der Suche in Fachzeitschriften

Ergänzend wird in dem Jahrgang 2019 einschlägiger Zeitschriften manuell/per Handsuche gesucht, um aktuelle, in den Datenbanken und Fachforen (noch) nicht erfasste Dokumente zu ermitteln.

5.4. Screening

5.4.1. Screening-Prozess

Die für die Beantwortung der Forschungsfragen relevanten Dokumente werden mittels eines mehrstufigen Screening-Verfahrens in der web-basierte Software EPPI-Reviewer 4 identifiziert, in dem die Einschlusskriterien angewendet werden. Vier Reviewerinnen und Reviewer screenen die Dokumente auf Titel-/Abstract-Ebene sowie die hierdurch ermittelten Dokumente auf Volltext-Ebene.

Duplikatsentfernung und Abstract-/Titelscreening

Im ersten Schritt werden zunächst Duplikate der ermittelten Dokumente entfernt. Im zweiten Schritt wird die Rating-Übereinstimmung (IRR) der Screenenden auf Titel-/Abstract-Ebene mittels der hierfür bereitgestellten Anwendung des EPPI-Reviewers geprüft. Im dritten Schritt werden

¹⁸ Da der Begriff des Scaffolding vielfach auch allgemein als Lernunterstützung benutzt wird, wird dieser hier mit Begriffen des sprachlichen Lernens und des fachlichen Lernens (bzw. allgemeiner Bildungserfolg) kombiniert. Damit erhöht sich die Wahrscheinlichkeit, sprachbezogenes Scaffolding zu identifizieren.

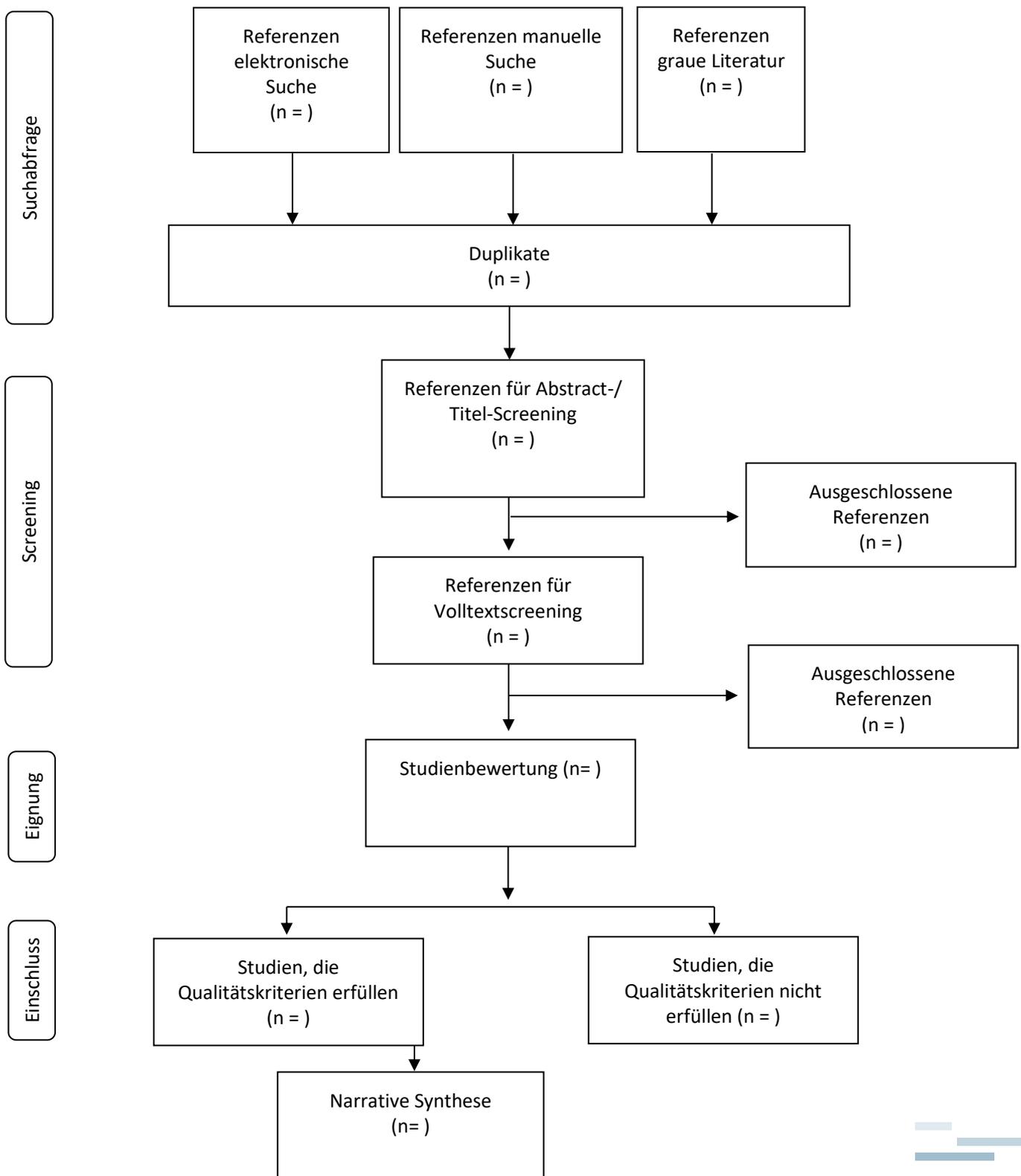
Titel und Abstract jedes Treffers hinsichtlich festgelegter Inklusionskriterien gescreent (single). Dokumente, welche die Inklusionskriterien nicht treffen, werden unter Nennung eines Ausschlusskriteriums (z. B. „ausgeschlossen wegen Intervention“, „ausgeschlossen wegen Studiendesign“ usw.) nicht in das folgende Volltextscreening eingeschlossen. Wenn eine Untersuchung von einer Person des Projektteams selbst durchgeführt wurde, wird das Dokument einer externen Person zum Screening übertragen. In Fällen, in denen die im Titel/Abstract gegebenen Informationen für einen definitiven Ausschluss nicht genügen, wird das Dokument in das Volltextscreening eingeschlossen.

Volltextscreening

Zunächst wird die Rating-Übereinstimmung (IRR) der Screenenden auf Volltext-Ebene mittels der hierfür bereitgestellten Anwendung des EPPI-Reviewers geprüft. Im nächsten Schritt werden die im Titel-/Abstract-Screening als relevant identifizierten Dokumente hinsichtlich der Einschlusskriterien im Volltext gescreent (single) und mit Begründung ein- oder ausgeschlossen. In allen Schritten des Screening-Prozesses werden die Namen der Screenenden und die Ausschlussgründe festgehalten.

Flowdiagramm

Für eine Dokumentation des gesamten Screeningprozesses wird das PRISMA-Flowdiagramm (Moher, Liberati, Tetzlaff, Altman & The PRISMA Group, 2009) adaptiert.



5.4.2. Aufbereitung der ermittelten Dokumente

Ermittelte Dokumente werden zunächst in das Literaturverwaltungsprogramm Zotero importiert. Dabei werden folgende Daten, die je nach Dokumententyp variieren können, automatisch oder manuell erfasst: Autorenschaft, Publikationsjahr, -art, und -sprache, Titel, Abstract, Sammelbandtitel, Herausgebende, Reihe, Band, Jahreszahl, DOI, ISBN, ISSN, Quelle/Suchorttyp, URL, Datum des Zugriffs.

Anschließend werden die Datensätze im RIS-Format in das webbasierte Softwareprogramm EPPI-Reviewer 4 (EPPI-Centre) importiert, das verschiedene Anwendungen für das Verwalten und Analysieren von Daten in systematischen Übersichtsarbeiten bietet.

5.5. Kodierung und Studienbewertung

Nach dem Screening werden die Studien, die alle Inklusionskriterien treffen, kodiert und anschließend bewertet.

5.5.1. Kodierung der Studieninformationen

Vier Personen aus dem Review-Team teilen die integrierten Studien untereinander auf und extrahieren mit einem Kodierschema deskriptive Daten aus den jeweils zugeteilten Studien. Die Kodierinnen und Kodierer prüfen im Anschluss daran paarweise die extrahierten Studieninformationen des Kodierpartners auf Vollständigkeit.

Die Kodieritems enthalten Informationen zum Inhalt und zur Beschaffenheit der Studien. Genauer erfassen sie allgemeine Informationen zur Publikation (Titel, Quelle, etc.) sowie zum Studienergebnis. Weitere Items beziehen sich auf die PICOS-Kategorien, die der leitenden Untersuchungsfrage und den abgeleiteten Inklusionskriterien zugrunde liegen, d. h. auf die Untersuchungspopulation, auf die Intervention, auf die Vergleichsgruppe, auf das Outcome und auf das Studiendesign. Für die Pilotierung der Kodieritems dienten die Prototypenstudien, die bereits zur Pilotierung der Inklusionskriterien genutzt wurden (s. 5.2).

Anhand derjenigen Studieninformationen, die in der Gesamtbetrachtung Rückschlüsse auf die Studienqualität zulassen (z. B. zum Aufbau der Studie), werden die Studien in einem zweiten Schritt bewertet.

5.5.2. Beurteilung der internen und externen Validität

Im Anschluss an die Kodierung der Studieninformationen wird zur Erfassung der „Güte der aus den Studien gezogenen Schlussfolgerungen“ (Döring & Bortz, 2016, S. 42) die externe und interne Validität bewertet. Mit Petticrew und Roberts (Petticrew & Roberts, 2006, S. 280) wird interne Validität hier verstanden als „extent to which the design and conduct of a study eliminates the possibility of bias“ und die externe Validität als „extent to which a study can be generalized to other populations, places, settings and times“ (Petticrew & Roberts, 2006, S. 279).

Interne Validität

Um Kriterien für das vorliegende Review zu bestimmen, wurden über eine freie Internet- und Literaturrecherche zunächst Quellen identifiziert, die Kriterien enthalten, die für die Validitätsbewertung des fokussierten Studientypus brauchbar sein können — d. h. für quantitative Interventionsstudien mit (quasi)-experimentellem oder Beobachtungsdesign (s. Inklusionskriterien, 5.2). Folgende Quellen wurden einbezogen:

What Works Clearinghouse (2017); Valentine und Cooper (2008); Jadad et al. (1996); National Institute for Health and Excellence (2012); Pawson et al. (2003); Puffer et al. (2003); Crombie (1996), EPPI-Centre (2013); Gough et al. (2017); Higgins, Lasserson, Chandler, Tovey und Churchill (2016), Cook und Campbell (1983), CASP (2018); Centre for Reviews and Dissemination (2009); Petticrew und Roberts (2006); Rutter, Francis, Coren und Fisher (2010); Joanna Briggs Institute (2017).

Im nächsten Schritt wurde geprüft, inwiefern diese Quellen Qualitätskriterien angeben, die folgende Bedingungen erfüllen:

- passend für pädagogische Interventionsstudien¹⁹
- gleichermaßen anwendbar auf experimentelle, quasi-experimentelle und Beobachtungsdesigns
- Balance zwischen Typ I-Fehler (die Intervention wird empfohlen, obwohl sie nicht wirkt) und Typ II-Fehler (die Intervention wird nicht empfohlen, obwohl sie wirkt)²⁰
- Einheitlichkeit auf horizontaler Betrachtungsebene (Streuung der Inhalte) und auf vertikaler Betrachtungsebene (Betrachtungstiefe)
- möglichst hoher Konkretisierungsgrad²¹
- nützlich für narrative Forschungssynthese

Da keine der oben angegebenen Quellen ein Kriterienset enthält, das diese Bedingungen umfassend erfüllt, wurden im nächsten Schritt zum Bezugsrahmen des Review-Projekts passende Bewertungskriterien für die interne Validität entwickelt. Für die Pilotierung der Kriterien dienten die Prototypenstudien, die bereits zur Pilotierung der Inklusionskriterien genutzt wurden (s. 5.2). Zur ausführlichen Diskussion der Bewertungskriterien und ihrer Operationalisierungen, siehe Höfler und Woerfel (i. Vorb.b). Tabelle 2 zeigt das entwickelte Kriterienset.

¹⁹ Viele Kriteriensets entstammen der klinischen Interventionsforschung, weshalb bei der Übertragbarkeit Vorsicht geboten ist (Petticrew & Roberts, 2006, S. 128).

²⁰ Siehe dazu auch Petticrew und Roberts (2006, S. 128).

²¹ Da sich Studien sehr stark in den Details ihres Aufbaus bzw. ihrer Durchführung unterscheiden, kann die Formulierung von Validitätskriterien oftmals nur annähernd konkret sein.

Tabelle 2 Bewertungskriterien der internen Validität (IV)

Bewertungskriterien Interne Validität (IV)
IV 1: Ist die Population der Interventionsstudie beschrieben, indem relevante Informationen/Daten (Alter, Geschlecht) ²² angegeben werden?
IV 2: Ist das Setting der Intervention beschrieben, indem relevante Informationen/Daten (Unterrichtsregion, Lernsetting, Fachbereich) ²³ angegeben werden?
IV 3: Ist die Intervention selbst beschrieben, indem Informationen zu Ziel, Ablauf und Intensität (Dauer, Anzahl an Wiederholungen, Länge des Zeitintervalls zwischen Wiederholungen) angegeben werden?
IV 4: Konstruktvalidität aV ²⁴ : Repräsentieren die verwendeten Outcome-Messinstrumente inhaltlich die interessierenden theoretischen Outcome-Konstrukte?
IV 5: Konstruktvalidität uV ²⁵ : Repräsentieren die durchgeführten Interventionen inhaltlich die interessierenden theoretischen Interventionskonzepte?
IV 6: Qualität der Erhebung aV und ggf. uV: Wurden die Variablen (aV/ggf. uV) mit einer hohen Wahrscheinlichkeit adäquat, das heißt durch validierte Erhebungsinstrumente/-verfahren, erhoben?
IV 7: Sind die Gruppen, die miteinander verglichen werden, ähnlich, d. h. gleichen sie sich z. B. in Alter, Geschlecht und Unterrichtskontext (Unterrichtsregion, Lernsetting, Fachbereich) bzw. wurden die Unterschiede methodisch gesteuert?
IV 8: Wurden die Gruppen ähnlich behandelt (allgemeines Interventionsziel, Interventionsintensität und -zeitraum, Berücksichtigung von Behandlungseffekten) bzw. wurden die Unterschiede methodisch gesteuert?
IV 9: Sind die angewendeten Outcome-Messinstrumente bei den Gruppen gleich?
IV 10: Wenn es im Laufe der Untersuchung einen Drop-Out-Anteil in der Untersuchungspopulation gab, wurden die Unterschiede in den Gruppen (prä- und post-aV) in der Ergebnisinterpretation berücksichtigt?

²² Die Populationsmerkmale sind orientiert an Bezugspunkten der externen Validität, s. S. 23

²³ Die Settingmerkmale sind orientiert an den Bezugspunkten der externen Validität s. S. 23

²⁴ Abhängige Variable

²⁵ Unabhängige Variable

<p>IV 11: Ist die Studie lang genug angelegt (Länge zwischen uV und aV), um Veränderungen im Outcome festzustellen?</p>
<p>IV 12: Ist die Stichprobe groß genug, um Interventionseffekte festzustellen ($n \geq 20$)</p>
<p>IV 13: Wurde die Outcome-Bewertung blind zum Gruppenstatus der untersuchten Personen vorgenommen?</p>

Bevor die Studien auf Basis der o.a. Kriterien bewertet werden, wird die Rating-Übereinstimmung (IRR) von vier Bewerterinnen und Bewertern des Reviewteams hinsichtlich der Qualitätskriterien mittels der hierfür bereitgestellten Anwendung des EPPI-Reviewers geprüft. Die Studienbewertung findet dann entlang der Kriterien statt. Falls im Rahmen dieses Verfahrens Divergenzen hinsichtlich der internen Validität einzelner Studien auftreten, wird eine externe Bewerterin oder ein externer Bewerter hinzugezogen. Ihre/seine Entscheidung gibt den Ausschlag.

Das Review verzichtet auf ein quantitatives Ranking der internen Validität integrierter Studien, weil dieses im wissenschaftlichen Diskurs kritisch betrachtet wird (Petticrew & Roberts, 2006, S. 129f.; Valentine & Cooper, 2008, S. 132f.). Die Beurteilung der internen Validität der Wirkungsstudien erfolgt im Review-Projekt narrativ, basierend auf einer tabellarischen Aufschlüsselung der Ergebnisse je Kriterium und Studie.

Externe Validität

Neben der Prüfung der internen Validität (s. S. 25) ist die externe Validitätsprüfung wesentlich, um die in das Review integrierten Studien zu bewerten: „[...] one at least needs also to know whether the study results are likely to be generalizable“ (Petticrew & Roberts, 2006, S. 130). Vor allem bei pädagogischen Wirkungsstudien ist zu prüfen, bei welchen Personen, in welchen Kontexten, welche Effekte geprüft und belegt sind. Dies ist bedeutsam, um Empfehlungen für eine differentielle Bildungspraxis abzuleiten.

Die Determinanten, die eine generelle Wirkung von sprachsensiblen Interventionen potenziell einschränken können, sind mannigfaltig und können deshalb nicht vollständig in der Studienbewertung berücksichtigt werden. Die Prüfung der externen Validität kann sich nur auf solche Aspekte beziehen, zu denen Studien mit hoher Wahrscheinlichkeit Informationen/Daten liefern und zu denen die Forschung ernstzunehmende Hinweise auf moderierende Einflüsse bereitstellt.

Insgesamt werden hier sieben potenzielle Determinanten der Generalisierbarkeit der Wirkung von sprachsensiblen Unterrichtsansätzen berücksichtigt. Die Determinanten sind in der folgenden Darstellung unter die drei Inhaltsbereiche („Subgroups“) subsumiert, die nach Valentine und Cooper (Valentine & Cooper, 2008, S. 140) für eine Beurteilung der Generalisierbarkeit zu berücksichtigen sind:

Subgroups of Participants:

- Alter
- Geschlecht

Subgroups of Settings:

- Unterrichtsregion
- Lernsetting
- Fachbereich

Subgroups of Outcomes:

- Lernbereich
- (Lern-)Leistungsindikator

Die mit Bezug auf diese Determinanten entwickelten Kriterien zur Bewertung der externen Validität sind in Tabelle 3 dargestellt (zur ausführlichen Begründung der potenziell moderierenden Wirkung der gewählten Determinanten Höfler & Woerfel, i. Vorb.b).

Tabelle 3 Bewertungskriterien der externen Validität (EV)

	Bewertungskriterien EV
EV 1	Inwiefern ist die Studie generalisierbar auf unterschiedliche Altersgruppen?
EV 2	Inwiefern ist die Studie generalisierbar auf Schülerinnen und Schüler?
EV 3	Inwiefern ist die Studie generalisierbar auf unterschiedliche regionale Unterrichtsregionen?
EV 4	Inwiefern ist die Studie generalisierbar auf unterschiedliche Lernsettings?
EV 5	Inwiefern ist die Studie generalisierbar auf unterschiedliche Fachbereiche?
EV 6	Inwiefern ist die Studie generalisierbar auf fachliches und sprachliches Lernen?
EV 7	Inwiefern ist die Studie generalisierbar auf unterschiedliche (Lern-)Leistungsindikatoren?

Im Review erfolgt zunächst die Bewertung der externen Validität der einzelnen integrierten Studien. Eine Gesamtbewertung zu einer Studie, z. B. über die Beurteilung der Items E1 bis E7 (vgl. Tabelle 3) mittels einer zweistufigen Bewertungsskala (z. B. trifft zu/trifft nicht zu), wird jedoch nicht durchgeführt. Erst in der Gesamtbewertung der Generalisierbarkeit *aller zu einem Ansatz vorliegenden Studien* lässt sich die Frage beantworten, inwiefern die Verallgemeinerbarkeit der Wirkung eines Ansatzes einzuschätzen ist. Deshalb werden die Ergebnisse nach didaktisch-methodischem sprachsensiblen Ansatz und der dazu vorliegenden Evidenz tabellarisch geclustert und abschließend narrativ diskutiert. Es findet dabei keine Gewichtung der einzelnen Items statt, da diesen qualitativ unterschiedlichen Aspekten zugrunde liegen, die sich schwer quantifizieren lassen.

5.6. Narrative Forschungssynthese

In der narrativen Forschungssynthese werden die über das Screening ermittelten Studien zusam-

menfassend dargestellt und kritisch diskutiert, welche Aussagen auf Basis der aktuellen empirischen Studienlage zur Wirkung von sprachsensiblen Ansätzen getroffen werden können.

6. Finanzierung/Auftraggeber

Das systematische Review wird aus laufenden Mitteln des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache finanziert. Das Mercator-Institut ist ein durch die Stiftung Mercator initiiertes und gefördertes Institut der Universität zu Köln. Es will sprachliche Bildung verbessern. Um dieses Ziel zu erreichen, erforscht und entwickelt es innovative Konzepte, Maßnahmen und Instrumente für sprachliche Bildung. Es bildet regional Lehramtsstudierende aus sowie bundesweit Pädagoginnen und Pädagogen in KiTas, Schulen und der Erwachsenenbildung fort und bereitet wissenschaftliche Erkenntnisse gezielt für Entscheidungsträger in Bildungspolitik und -verwaltung sowie Bildungspraxis auf. Mit seiner Forschung und seinen wissenschaftlichen Serviceleistungen zu sprachlicher Bildung in einer mehrsprachigen Gesellschaft trägt das Mercator-Institut zu mehr Chancengleichheit im Bildungssystem bei.

Das Mercator-Institut war und ist in dem Themenbereich des geplanten systematischen Reviews aktiv:

- Förderung externer Projekte (z. B. Miteinander, Ruhr Universität Bochum)
- Forschung/Publicationen (Kooperationsprojekt mit Universität Duisburg-Essen)
- Lehre (Modul „Deutsch für Schülerinnen und Schüler mit Zuwanderungsgeschichte“; Weiterbildungsstudium DaZ, Fort-Weiterbildungsmaßnahmen)
- Beratungen (Orientierungsrahmen DaZ; Bund-Länder-Verbundprojekt „Bildung durch Sprache und Schrift“; BISS)

Es bestehen keinerlei Interessenskonflikte, die die Unabhängigkeit der Autorinnen und Autoren einschränken könnten.

7. Projektteam und Kooperationspartner

Projektleitung

Prof. Dr. Michael Becker-Mrotzek (Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache)

Dr. Martha Höfler (operative Projektleitung seit April 2020, Mercator-Institut)

Dr. Till Woerfel (operative Projektleitung August 2018 bis April 2020, Mercator-Institut)

Projektteam

Dr. Martha Höfler

Anastasia Knaus (bis August 2019)

Rebekka Wanka

Dr. Annika Witte

Dr. Till Woerfel

(alle Mercator-Institut)

Beratung

Methodik

Jolika Schulte (Universität Vechta)

Prof. Dr. Karin Zimmer (Universität Vechta/DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Inhalt

Dr. Christoph Gantefort (Mercator-Institut)

Prof. Dr. Michael Becker-Mrotzek (Mercator-Institut)

Praxis

Dr. Iris Günther (Lehrerin und Lehrkraft für besondere Aufgaben, Mercator-Institut)

Dr. Monika Socha (wissenschaftliche Mitarbeiterin, Mercator-Institut)

Nadine Steckenborn (abgeordnete Lehrkraft, Mercator-Institut)

Dr. Peter Weber (abgeordnete Lehrkraft, Mercator-Institut)

Technische Unterstützung

Leonie Twente (WHB, Mercator-Institut)

Susana Mogollón (WHB, Mercator-Institut)

Wissenschaftskommunikation

Frauke König (Stabsstelle Kommunikation, Mercator-Institut)

Anna Niewerth (Stabsstelle Kommunikation, Mercator-Institut)

Referenzen

- Azevedo, Roger; Cromley, Jennifer G.; Winters, Fielding I.; Moos, Daniel C. & Greene, Jeffrey A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional Science*, 33 (5/6), 381–412.
- Becker-Mrotzek, Michael & Woerfel, Till (im Erscheinen). Sprachsensibler Unterricht und Deutsch als Zweitsprache als Gegenstand der Lehrerbildung. In Colin Cramer, Martin Drahm, Johannes König, Martin Rothland & Sigrid Blömeke (Hrsg.), *Handbuch Lehrerbildung*. Bad Heilbrunn: Klinkhardt.
- Beese, Melanie (2010). *Sheltered Instruction Observation Protocol. The SIOP Model – ein Modell zum integrierten Fach- und Sprachenlernen in allen Fächern mit besonderem Fokus auf Zweitsprachenlernende*. Duisburg: Universität Duisburg-Essen. Verfügbar unter: <http://www.uni-due.de/imperia/md/content/prodaz/siop.pdf>.
- Benoit, Kenneth; Watanabe, Kohei; Wang, Haiyan; Nulty, Paul; Obeng, Adam; Müller, Stefan & Matsuo, Akitaka (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3 (30), 774. doi:10.21105/joss.00774
- Bredthauer, Stefanie (2019). Sprachvergleiche als multilinguale Scaffolding-Strategie. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 24 (1), 127–143.
- Bruner, Jérôme Seymour (1978). The role of dialogue in language acquisition. In Anne Sinclair, Robert J. Jarvella & Willem Johannes Maria Levelt (Hrsg.), *The child's conception of language* (2. Auflage). Berlin: Springer.
- Bulgren, Janis A.; Lenz, B. Keith; Schumaker, Jean B.; Deshler, Donald D. & Marquis, Janet G. (2002). The use and effectiveness of a comparison routine in diverse secondary content classrooms. *Journal of Educational Psychology*, 94 (2), 356–371. doi:10.1037/0022-0663.94.2.356
- Busse, Vera (2019). Umgang mit Mehrsprachigkeit und sprachsensibler Unterricht aus pädagogischer Sicht: Ein einführender Überblick. In Martin Butler & Juliana Goschler (Hrsg.), *Sprachsensibler Fachunterricht: Chancen und Herausforderungen aus interdisziplinärer Perspektive* (S. 1–34). Wiesbaden: Springer Fachmedien Wiesbaden. doi:10.1007/978-3-658-27168-8
- Cenoz, Jasone (2015). Content-based instruction and content and language integrated learning: the same or different? *Language, Culture and Curriculum*, 28 (1), 8–24. doi:10.1080/07908318.2014.1000922
- Centre for Reviews and Dissemination (2009). *Systematic Reviews. CRD's guidance for undertaking reviews in health care*. York: University of York. Verfügbar unter: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [7.5.2020].
- Chiu, Mei-Hung; Chou, Chin-Cheng & Liu, Chia-Ju (2002). Dynamic processes of conceptual change: Analysis of constructing mental models of chemical equilibrium. *Journal of Research in Science Teaching*, 39 (8), 688–712. doi:10.1002/tea.10041
- Cook, T.D. & Campbell, D.T. (1983). The design and conduct of true experiments and quasi-experiments in field settings. In M.D. Dunnette (Hrsg.), *Handbook of Industrial and Organizational Psychology* (3. Auflage, S. 223–326). Hoboken, NJ: Wiley-Interscience.
- Critical Appraisal Skills Programme (2018). CASP Appraisal Checklists Verfügbar unter: <https://casp-uk.net/casp-tools-checklists/> [17.10.2019].
- Crombie, Ian K. (1996). *The Pocket Guide to Critical Appraisal. A Handbook for Health Care Professionals*. London, UK: BMJ Publishing Group.
- Daniel, Shannon M.; Jiménez, Robert T.; Pray, Lisa & Pacheco, Mark B. (2017). Scaffolding to make translanguaging a classroom norm. *TESOL Journal*, e361, 1–21. doi:10.1002/tesj.361

- Döring, Nicola & Bortz, Jürgen (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Berlin, Heidelberg: Springer.
- Echevarría, Jana (2012). *Effective Practices for Increasing the Achievement of English Learners*. Center for Research and Reform in Education. Verfügbar unter: <http://www.cal.org/create/publications/briefs/pdfs/effective-practices-for-increasing-the-achievement-of-english-learners.pdf> [16.4.2019].
- Echevarría, Jana; Richards-Tutor, Catherine; Canges, Rebecca & Francis, David (2011). Using the SIOP Model to Promote the Acquisition of Language and Science Concepts with English Learners. *Bilingual Research Journal*, 34 (3), 334–351. doi:10.1080/15235882.2011.623600
- Echevarría, Jana & Short, Deborah J. (2011). *The SIOP Model: A Professional Development Framework for a Comprehensive School-Wide Intervention*. The Center for Research on the Educational Achievement and Teaching of English Language Learners. Verfügbar unter: <http://www.cal.org/create/publications/briefs/pdfs/professional-development-framework.pdf> [13.3.2019].
- Echevarría, Jana; Short, Deborah & Powers, Kristin (2006). School Reform and Standards-Based Education: A Model for English-Language Learners. *The Journal of Educational Research*, 99 (4), 195–211. doi:10.3200/JOER.99.4.195-211
- Echevarría, Jana; Vogt, MaryEllen & Short, Deborah J. (2000). *Making content comprehensible for English learners: the SIOP Model*. Boston: Pearson Studium.
- EPPI-Centre (2013). Guidelines for extracting data and quality assessing primary studies in educational research Version 0.9.7. London, UK. Verfügbar unter: <https://eppi.ioe.ac.uk/cms/Resources/Tools/tabid/184/Default.aspx#Guidelines> [3.7.2019].
- Gantefort, Christoph & Oroquieta Sánchez, Maria Jose (2015). Translanguaging-Strategien im Sachunterricht der Primarstufe: Förderung des Leseverstehens auf Basis der Gesamtsprachigkeit. *Transfer Forschung ↔ Schule.*, 1 (1), 24–37.
- García, Ofelia (2009). *Bilingual education in the 21st century: A global perspective*. Malden, MA; Oxford: Wiley-Blackwell.
- García, Ofelia & Flores, Nelson (2012). Multilingual pedagogies. In Marilyn Martin- Jones, Adrian Blackledge & Angela Creese (Hrsg.), *The Routledge handbook of multilingualism* (S. 232–246). London ; New York: Routledge.
- García, Ofelia & Wei, Li (2018). Translanguaging. *The Encyclopedia of Applied Linguistics* (S. 1–7). American Cancer Society. doi:10.1002/9781405198431.wbeal1488
- Gibbons, Pauline (2002). *Scaffolding Language, Scaffolding Learning: Teaching English Language Learners in the Mainstream Classroom*. Portsmouth, NH: Heinemann.
- Gibbons, Pauline (2009). *English learners, academic literacy, and thinking: Learning in the challenge zone*. Portsmouth, NH: Heinemann.
- Gough, David; Oliver, Sandy & Thomas, James (Hrsg.). (2017). *An Introduction to Systematic Reviews* (2nd edition.). Los Angeles: SAGE.
- Grames, Eliza (2019). Introduction to litsearchr with an example of writing a systematic review search strategy for Black-backed Woodpecker occupancy of post-fire forest systems Verfügbar unter: https://elizagrames.github.io/litsearchr/introduction_vignette_v010.html [3.1.2019].
- Grames, Eliza M.; Stillman, Andrew N.; Tingley, Morgan W. & Elphick, Chris S. (2019). An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution*, 10 (10), 1645–1654. doi:10.1111/2041-210X.13268

- Gräsel, Cornelia (2010). Stichwort: Transfer und Transferforschung im Bildungsbereich. *Zeitschrift für Erziehungswissenschaft*, 13 (1), 7–20. doi:10.1007/s11618-010-0109-8
- Große, Maria (2015). Pons Latinus - Modellierung eines sprachsensiblen Lateinunterrichts. *Berlin-Brandenburger Beiträge zur Bildungsforschung 2015. Herausforderungen, Befunde und Perspektiven interdisziplinärer Bildungsforschung*. Frankfurt am Main: Peter Lang.
- Härtig, Hendrik & Stosik, Tina (2015). Wortschatztraining im Physikunterricht. *Der mathematische und naturwissenschaftliche Unterricht*, 68 (3), 155–158.
- Hayes, Katherine; Rueda, Robert & Chilton, Susan (2009). Scaffolding language, literacy, and academic content in English and Spanish: The linguistic highway from Mesoamerica to Southern California. *English Teaching: Practice and Critique*, 8 (2), 137–166.
- Henderson, Kathryn I. (2017). Ofelia García, Susana Ibarra & Kate Seltzer. (2016). The Translanguaging Classroom: Leveraging Student Bilingualism for Learning. Philadelphia, PA: Les Editions CEC. *Journal of Immersion and Content-Based Language Education*, 5 (2), 300–303. doi:10.1075/jicb.5.2.08hen
- Higgins, Julian PT; Lasserson, Toby; Chandler, Jackie; Tovey, David & Churchill, Rachel (2016). *Methodological Expectations of Cochrane Intervention Reviews (MECIR)*. London, UK: Cochrane. Verfügbar unter: <http://community.cochrane.org/sites/default/files/uploads/MECIR%20PRINTED%20BOOK%20LET%20FINAL%20v1.01.pdf> [22.10.2019].
- Höfler, Martha & Woerfel, Till (i. Vorb.a). Heterogenität und Komplexität von Wirkungsstudien zu sprachsensiblen Unterrichtsansätzen
- Höfler, Martha & Woerfel, Till (i. Vorb.b). Studienbewertung in systematischen Reviews: Kriterien zur Beurteilung von Interventionsstudien in der (sprachlichen) Bildungsforschung
- Honigfeld, Andrea & Cohan, Audrey (2006). Lesson study meets SIOP: Linking two successful professional development models. *Faculty Works: Education*, 17.
- Jadad, Alejandro R.; Moore, R. Andrew; Carroll, Dawn; Jenkinson, Crispin; Reynolds, D. John M.; Gavaghan, David J. & McQuay, Henry J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17 (1), 1–12. doi:10.1016/0197-2456(95)00134-4
- Joanna Briggs Institute (2017). *Critical Appraisal Checklist for Randomized Controlled Trials*. Joanna Briggs Institute. Verfügbar unter: https://joannabriggs.org/sites/default/files/2019-05/JBI_RCTs_Appraisal_tool2017_0.pdf [7.5.2020].
- Khan, Khalid S.; Kunz, Regina; Kleijnen, Jos & Antes, Gerd (Hrsg.). (2004). *Systematische Übersichten und Meta-Analysen: Ein Handbuch für Ärzte in Klinik und Praxis sowie Experten im Gesundheitswesen*. Berlin Heidelberg: Springer-Verlag. Verfügbar unter: <https://www.springer.com/de/book/9783540439363> [7.6.2019].
- Kugley, Shannon; Wade, Anne; Thomas, James; Mahood, Quenby; Jørgensen, Anne-Marie Klint; Hammerstrøm, Karianne & Sathe, Nila (2017). Searching for studies: a guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews*, 13 (1), 1–73. doi:https://doi.org/10.4073/cmg.2016.1
- Lang, Martin (2016). Förderung der fachspezifischen Schreibkompetenzen im Technikunterricht. *Befähigung zu gesellschaftlicher Teilhabe. Beiträge der fachdidaktischen Forschung*. Münster; New York: Waxmann.
- Lefebvre, Carol; Mannheimer, Eric & Glanville, Julie (2011). Searching for studies. In Peter Higgins & Sally Green (Hrsg.), *Cochrane Handbook of Systematic Reviews of Intervention. Version 5.1.0 (updated March 2011)*. The Cochrane Collaboration. Verfügbar unter: www.handbook.cochrane.org. [21.1.2019].
- Leisen, Josef (2011). Praktische Ansätze schulischer Sprachförderung – Der sprachensible Fachunterricht. Expertentagung der Hanns-Seidel-Stiftung in Kooperation mit dem Staatsinstitut für Schulqualität und Bildungsforschung (ISB) und dem Bayerischen Staatsministerium für

- Unterricht und Kultus. Gehalten auf der „Begegnen, Verstehen, Zukunft sichern – Beiträge der Schule zu einem gelungenen Miteinander“, Bildungszentrum Kloster Banz. Verfügbar unter: https://www.hss.de/fileadmin/media/downloads/Berichte/111027_RM_Leisen.pdf [1.6.2019].
- Liberati, Alessandro; Altman, Douglas G.; Tetzlaff, Jennifer; Mulrow, Cynthia; Gøtzsche, Peter C.; Ioannidis, John P.A.; Clarke, Mike; Devereaux, Philip J.; Kleijnen, Jos & Moher, David (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*, 62 (10), e1–e34. doi:10.1016/j.jclinepi.2009.06.006
- Macken-Horarik, Mary (1996). Literacy and learning across the curriculum: Towards a model of register for secondary school teachers. In Hasan Ruqaiya & Williams Geoff (Hrsg.), *Literacy in Society* (S. 232–278). London, UK: Longman.
- Macken-Horarik, Mary (1998). Exploring the requirements of critical school literacy: a view from two classrooms. In Frances Christie & Ray Misson (Hrsg.), *Literacy and Schooling*. London, UK: Routledge. doi:10.4324/9780203019559-9
- Massler, Ute; Stotz, Daniel & Queisser, Claudia (2014). Assessment instruments for primary CLIL: the conceptualisation and evaluation of test tasks. *The Language Learning Journal*, 42 (2), 137–150. doi:10.1080/09571736.2014.891371
- McIntyre, Ellen; Kyle, Diane; Chen, Cheng-Ting; Muñoz, Marco & Beldon, Scott (2010). Teacher Learning and ELL Reading Achievement in Sheltered Instruction Classrooms: Linking Professional Development to Student Development. *Literacy Research and Instruction*, 49 (4), 334–351. doi:10.1080/19388070903229412
- McKenzie, Lolita D. (2011). *Scaffolding English Language Learners' Reading Performance*. Minneapolis, MN: Walden University. Verfügbar unter: <https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=1954&context=dissertations> [24.4.2019].
- Meneses, Alejandra; Escobar, José-Pablo & Véliz, Soledad (2018). The effects of multimodal texts on science reading comprehension in Chilean fifth-graders: text scaffolding and comprehension skills. *International Journal of Science Education*, 40 (18), 2226–2244. doi:10.1080/09500693.2018.1527472
- Mercer, Neil; Dawes, Lyn; Wegerif, Rupert & Sams, Claire (2004). Reasoning as a Scientist: Ways of Helping Children to Use Language to Learn Science. *British Educational Research Journal*, (3), 359–378.
- Merritt, Eileen G.; Palacios, Natalia; Banse, Holland; Rimm-Kaufmann, Sara E. & Leis, Micela (2016). Teaching Practices in Grade 5 Mathematics Classrooms with High-Achieving English-Learner Students. *The Journal of Educational Research*, 110 (1), 17–31. doi:<https://doi.org/10.1080/00220671.2015.1034352>
- Michell, Michael & Sharpe, Tina (2005). Collective instructional scaffolding in English as a Second Language classrooms. *Prospect*, 20 (1), 31–58.
- Moher, David; Liberati, Alessandro; Tetzlaff, Jennifer; Altman, Douglas G. & The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine*, 6 (7). doi:10.1371/journal.pmed.1000097
- National Institute for Health and Care Excellence [NICE] (2012). *The guidelines manual*. London, UK: National Institute for Health and Clinical Excellence. Verfügbar unter: <https://www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-pdf-2007970804933> [7.5.2020].
- O'Connor, Denise; Green, Sally & Higgins, Peter (2011). Defining the review question and developing criteria for including studies. In Peter Higgins & Sally Green (Hrsg.), *Cochrane Handbook of Systematic Reviews of Intervention. Version 5.1.0 (updated March 2011)*. The Cochrane Collaboration. Verfügbar unter: www.handbook.cochrane.org. [21.1.2019].

- Paetsch, Jennifer; Wolf, Katrin M.; Stanat, Petra & Darsow, Annkathrin (2014). Sprachförderung von Kindern und Jugendlichen aus Zuwandererfamilien. *Zeitschrift für Erziehungswissenschaft*, 17 (2), 315–347. doi:10.1007/s11618-013-0474-1
- Paez, Arsenio (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, 10 (3), 233–240. doi:10.1111/jebm.12266
- Pant, Hans Anand (2014). Aufbereitung von Evidenz für bildungspolitische und pädagogische Entscheidungen: Metaanalysen in der Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*, 17 (4), 79–99. doi:10.1007/s11618-014-0510-9
- Pawson, Ray; Boaz, Annette; Grayson, Lesley; Long, Andrew & Barnes, Colin (2003). SCIE Knowledge review 03: Types and quality of knowledge in social care. Social Care Institut for Excellence. Verfügbar unter: <https://www.scie.org.uk/publications/knowledgereviews/kr03.asp> [7.5.2020].
- Petticrew, Mark & Roberts, Helen (2006). *Systematic reviews in the social sciences: a practical guide*. Malden, MA ; Oxford: Wiley-Blackwell.
- Popay, Jennie; Roberts, Helen; Sowden, Amanda; Petticrew, Mark; Arai, Lisa; Rodgers, Mark & Britten, Nicky (2006). *Guidance on the conduct of narrative synthesis in systematic Reviews. A Product from the ESRC Methods Programme. Version 1.* (Report). Swindon: Economic and Social Research Council. Verfügbar unter: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.3100&rep=rep1&type=pdf> [10.10.2018].
- Puffer, Suezann; Torgerson, David & Watson, Judith (2003). Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ : British Medical Journal*, 327 (7418), 785–789.
- R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Verfügbar unter: <http://www.R-project.org/> [18.1.2019].
- Rösch, Heidi (2013). Integrative Sprachbildung im Bereich Deutsch als Zweitsprache (DaZ). In Charlotte Röhner & Britta Hövelbrinks (Hrsg.), *Fachbezogene Sprachförderung in Deutsch als Zweitsprache: Theoretische Konzepte und empirische Befunde zum Erwerb bildungssprachlicher Kompetenzen* (S. 18–36). Weinheim: Beltz Juventa.
- Rous, Meike (2016). *Fachsprache im Biologieunterricht:: Förderung von konzeptuellem Lernen und Textverstehen durch fachspezifisch-sprachsensible Aufgaben*. Berlin: Logos Berlin.
- Rutter, Deborah; Francis, Jennifer; Coren, Esther & Fisher, Mike (2010). *SCIE systematic research reviews: guidelines* (2. Auflage). London, UK: Social Care Institute for Excellence.
- Schüler-Meyer, Alexander; Prediger, Susanne; Kuzu, Taha; Wessel, Lena & Redder, Angelika (2019). Is Formal Language Proficiency in the Home Language Required to Profit from a Bilingual Teaching Intervention in Mathematics? A Mixed Methods Study on Fostering Multilingual Students' Conceptual Understanding. *International Journal of Science and Mathematics Education*, 17 (2), 317–339. doi:10.1007/s10763-017-9857-8
- Schwippert, Knut; Neumann, Astrid & Leiß, Dominik (2016). Das Projekt „Fach-an-Sprache-an-Fach“. Aufbau bildungssprachlicher Prozeduren durch adaptive Aufgaben im Deutsch- und Mathematikunterricht. *Qualifizierung für sprachliche Bildung. Programme und Projekte zur Professionalisierung von Lehrkräften und pädagogischen Fachkräften*. Münster; New York: Waxmann.
- Sekretariat der Kultusministerkonferenz (Hrsg.). (2016). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Köln: Wolters Kluwer. Verfügbar unter: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf [6.6.2019].

- Short, Deborah J.; Fidelman, Carolyn G. & Louguit, Mohammed (2012). Developing Academic Language in English Language Learners Through Sheltered Instruction. *TESOL Quarterly*, 46 (2), 334–361. doi:10.1002/tesq.20
- Snilstveit, Birte; Oliver, Sandy & Vojtkova, Martina (2012). Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. *Journal of Development Effectiveness*, 4 (3), 409–429. doi:10.1080/19439342.2012.710641
- Sumfleth, Elke; Kobow, Iwen; Tunalı, Nermin & Walpuski, Maik (2013). Fachkommunikation im Chemieunterricht. In Michael Becker-Mrotzek, Karen Schramm, Eike Thürmann & Helmut Vollmer (Hrsg.), *Sprache im Fach: Sprachlichkeit und fachliches Lernen* (S. 255–276). Münster New York: Waxmann.
- Swain, Merrill & Lapkin, Sharon (2013). A Vygotskian sociocultural perspective on immersion education: The L1/L2 debate. *Journal of Immersion and Content-Based Language Education*, 1 (1), 101–129. doi:10.1075/jicb.1.1.05swa
- Swanson, H. Lee & Lussier, Catherine M. (2001). A Selective Synthesis of the Experimental Literature on Dynamic Assessment. *Review of Educational Research*, 71 (2), 321–363. doi:10.3102/00346543071002321
- Tenberg, Ralf (2006). *Didaktik lernfeldstrukturierter Unterrichts: Theorie und Praxis beruflichen Lernens und Lehrens*. Bad Heilbrunn: Klinkhardt.
- Turner, Julianne C.; Meyer, Debra K.; Cox, Kathleen E.; Logan, Candice; DiCintio, Matthew & Thomas, Cynthia T. (1998). Creating contexts for involvement in mathematics. *Journal of Educational Psychology*, 90 (4), 730–745. doi:10.1037/0022-0663.90.4.730
- Twente, Leonie & Woerfel, Till (2020). Implementing Text Mining to Optimize Search Strategies in Systematic Reviews [Poster]. Gehalten auf der LERN-Jahrestagung 2020 (abgesagt), Frankfurt a. M. Verfügbar unter: https://www.researchgate.net/publication/340061079_Implementing_Text_Mining_to_Optimize_Search_Strategies_in_Systematic_Reviews_Der_Einsatz_von_Text_Mining-Verfahren_zur_Optimierung_von_Suchanweisungen_in_systematischen_Reviews [8.5.2020].
- Valentine, Jeffrey C. & Cooper, Harris (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: the Study Design and Implementation Assessment Device (Study DIAD). *Psychological Methods*, 13 (2), 130–149. doi:10.1037/1082-989X.13.2.130
- van de Pol, Janneke; Volman, Monique & Beishuizen, Jos (2010). Scaffolding in Teacher–Student Interaction: A Decade of Research. *Educational Psychology Review*, 22 (3), 271–296. doi:10.1007/s10648-010-9127-6
- Vidot, Jose L. (2011). *The Efficacy of Sheltered Instruction Observation Protocol (SIOP) in Mathematics Instruction on English Language Learner Students*. Minneapolis, MN: Walden University. Verfügbar unter: <https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=1942&context=dissertations> [17.4.2019].
- Vygotskij, Lev S. (1975). *Thought and language* (12. Auflage). Cambridge, Mass: MIT Press.
- Wessel, Lena & Prediger, Susanne (2012). Fach- und sprachintegrierte Förderung für mehrsprachige Lernende am Beispiel von Anteilen und Brüchen. *Beiträge zum Mathematikunterricht 2012 Digital. Vorträge auf der 46. Tagung für Didaktik der Mathematik*. Gesellschaft für Didaktik der Mathematik. Verfügbar unter: <https://doi.org/10.17877/DE290R-1238>
- What Works Clearinghouse (2009). *Sheltered Instruction Observation Protocol (SIOP)*. (What Works Clearinghouse Intervention Report). Institute of Education Sciences. Verfügbar unter: <https://files.eric.ed.gov/fulltext/ED505688.pdf> [7.6.2020].
- What Works Clearinghouse (2013). *Sheltered Instruction Observation Protocol (SIOP)*. (What Works Clearinghouse Intervention Report). Institute of Education Sciences. Verfügbar un-

- ter: https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_siop_022013.pdf
[7.6.2020].
- What Works Clearinghouse (Hrsg.). (2017). What Works Clearinghouse™ Standards Handbook (Version 4.0) Verfügbar unter:
https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf
f [13.11.2019].
- Whittier, L. Elena & Robinson, Michael (2007). Teaching Evolution to Non-English Proficient Students by Using Lego Robotics. *American Secondary Education*, 35 (3), 19–28.
- Williams, Cen (1994). *Arfarniad o Ddulliau Dysgu ac Addysgu yng Nghyd-destun Addysg UwchraddDdwyeithog, [An evaluation of teaching and learning methods in the context of bilingualsecondary education]*. Bangor: University of Wales.
- Wood, David; Bruner, Jérôme Seymour & Ross, Gail (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, 17 (2), 89–100. doi:10.1111/j.1469-7610.1976.tb00381.x
- Zwiers, Jeff (2006). Integrating academic language, thinking, and content: Learning scaffolds for non-native speakers in the middle grades. *Journal of English for Academic Purposes*, 5 (4), 317–332. doi:10.1016/j.jeap.2006.08.005

Appendix

Tabelle 4 Suchorte und datenbankspezifische Einschlusskriterien

Suchort <i>Name der Datenbank/des Journals/ URL)</i>	Suchorttyp	Vorgehen
Scopus	<i>allgemeine Datenbank</i>	Strukturierte Abfragen
Education Resource Information Centre (ERIC)	<i>Fachdatenbank</i>	Strukturierte Abfragen
PSYINDEX	<i>Fachdatenbank</i>	Strukturierte Abfragen
PsycINFO	<i>Fachdatenbank</i>	Strukturierte Abfragen
FIS Bildung	<i>Fachdatenbank</i>	Strukturierte Abfragen
ResearchGate.net Academia.edu	<i>akademische Plattformen</i>	Manuelle Suche (auf der Grundlage der die Inklusionskriterien erfüllenden Dokumente)
Linguist List ResearchGate.net	<i>akademische Plattformen</i>	„Call for preprints“: Abfrage nicht publizierter Manuskripte, vorläufiger Ergebnisse noch nicht abgeschlossener Projekte
EPPI-Centre	<i>Institutsinterne Datenbank</i>	Suche nach Systematic Reviews und Meta-Analysen (Vorrecherche)
Campbell Collaboration	<i>Institutsinterne Datenbank</i>	Suche nach Systematic Reviews und Meta-Analysen (Vorrecherche)

What Works Clearing House	<i>Institutsinterne Datenbank</i>	Suche nach Systematic Reviews („Intervention reports“) (Vorrecherche)
Opengrey	<i>graue Literatur Datenbank</i>	Strukturierte Abfragen nach unveröffentlichten Manuskripten, vorläufigen Ergebnissen noch nicht abgeschlossener Projekte
<p>Educational Psychology Review</p> <p>Fremdsprache Deutsch - Zeitschrift für die Praxis des Deutschunterrichts</p> <p>International Journal of Multilingualism</p> <p>International Multilingual Research Journal</p> <p>Journal of Educational Psychology</p> <p>Journal of Immersion and Content-Based Language Education</p> <p>JoWR (Journal of Writing Research)</p> <p>Language and Education</p> <p>Language Learning</p> <p>Latin and American Journal of Content and Language Integrated Learning (LACLIL)</p> <p>Learning & Instruction</p> <p>Praxis Deutsch</p>	<i>Handsuche/manuelle Zeitschriftensuche</i>	Suche in den aktuellsten Ausgaben (2019) einschlägiger Zeitschriften, die in den o.g. Datenbanken ggf. noch nicht erfasst sind.

Studies in Second Language Acquisition Teaching and Teacher Education The Journal of Educational Research The Language Learning Journal (LLJ) The Modern Language Journal Zeitschrift für Fremdsprachenforschung Zeitschrift für Erziehungswissenschaft Zeitschrift für Unterrichtswissenschaft Zielsprache Deutsch		
---	--	--

Tabelle 5 Entwicklung der Suchstrategie/Optimierung Suchbegriffe (Englisch)

PICOS-Kategorie	Quelle	Stichwörter
Welche Probanden- gruppe soll in wel- chem Setting berück- sichtigt werden (Pop- ulation)	1	primary; pupil; school; secondary; student
	2	“elementary education”; “elementary school students”; “elementary school”; “elementary students”; “ell stu- dents”; “english language learner”; “high school students”; “high schools”; “junior high schools”; “middle school students”; “middle school”; “primary education”; “primary school”; “public school”; “school students”; “second- ary education”; “secondary school”
	3	“bilingual students”; classes; classrooms; “diversely populated school”; elementary; ELL; “English; language learners”; “English Language Learners”; “English Only students”; “fifth-grade classrooms”; “German schools”; Grade; “language minority students”; “lifescience classes”; “mainstream classrooms”; “Mathematics Class- rooms”; “Multilingual Students”; “native Spanish speakers”; “Non-English Proficient Students”; “U.S. Schools”; “English as a second language classes”; “rural high school”
Was soll getestet worden sein (Inter- vention)?	1	"content and language learning"; "integrated strategies in classroom"; "language and subject learning"; "lan- guage based learning"; "language objectives and content learning; "language promotion"; Scaffolding; "sheltered Instruction Observation Protocol"; SIOP; "thinking together"
Welche Intervention ist Gegenstand der Untersuchung?	2	“academic language instruction”; “instructional approach”; “interaction based strategies in classroom”; “lan- guage learning while teaching content”; “sheltered English”; “sheltered immersion”; “sheltered instruction”
	3	“attention to vocabulary building”; “bilingual teaching”; “instructional models”; “instructional practices”; “lan- guage education”; “language integrated learning”; “Mathematics Instruction”; “reciprocal teaching”; “scaffolding teaching technique”; “sheltered instruction observation protocol”; “sheltered instruction”; “talk-based activity”; “teaching content curriculum”; “Teaching Evolution”; “translanguaging practices”; “use language”

Nach welchen Ergebnissen soll gesucht werden (Outcome)?	1	Art; biology; chemistry; Civics; "content lesson"; "content objectives"; Economics; Ethics; "educational progress"; "language competence"; "language proficiency"; "language skills"; Geography; grammar; History; Humanities; lexic; Math; morpheme; Physics; science; semantic; sentence; Social Studies"; syntax; technology; vocabulary; word
	2	"academic achievement"; "academic content"; "academic language"; "academic literacy"; "academic performance"; "achievement gain"; "communicative competence"; "comprehension skills"; "content knowledge"; "language awareness"; "language proficiency"; "language skills"; "listening comprehension"; "listening skills"; "literacy development"; "literacy education"; "literacy skills"; "metalinguistic awareness"; "oral proficiency"; "oral reading"; "reading achievement"; "reading comprehension"; "reading fluency"; "reading improvement"; "reading skills"; "student learning"; "student performance"; "vocabulary development"; "vocabulary knowledge"; "vocabulary learning"; "writing achievement"; "writing skills"; "phonological awareness"; "student achievement"; "vocabulary skills"
	3	"academic English"; "Academic Progress"; "academic register"; "academic success"; "academic understanding"; "acquisition of knowledge"; "Acquisition of Language and Science Concepts"; "background knowledge"; "communicative skills"; "conceptual understandings of fractions"; "content area"; "content curriculum"; "develop students' skills"; "development of understanding"; "discursive process"; "Educational Achievement"; "English scores"; "formal language proficiency"; "language development"; "learn Science"; "learning gains"; "learning of science"; Mathematics; "Oral Language"; "principles of evolution"; "professional development"; "Reading Performance"; science; "scientific concepts"; "scientific understanding"; "state science standards"; summarizing; talk; "teacher learning"; "technical register"; "understand mathematics"; "value-added metrics"; "ways of reasoning"
Soll nach einem bestimmten Studierende-	1	affect; cause; cohort; controlled; correlation; "cross section"; effect; efficacy; enhance; experiment; foster; impact; improve; increase; inference; influence; intervention; "mixed method"; observational; post; predict; profit;

sign gesucht werden (Study Design)?		promote; significance
	2	“case studies”; “effect size”; “ethnographic study”; “experimental design”; “longitudinal study”; “mixed methods”; “multivariate analysis”; “participant observation”; “positive effects”; “positive impact”; “predictor variables”; “pretests posttests”; “program effectiveness”; “program evaluation”; “qualitative study”; “quantitative research”; “quantitative study”; “quasiexperimental design”; “regression analysis”; “regression statistics”; “research design”; “research study”; “qualitative research”; “significant difference”; “significant gains”; “significant increase”; “statistical analysis”; “statistically significant”; “test results”; “test scores”
	3	ANCOVA; analysis; “analysis of covariance”; “analysis of variance”; assess; compared; controls; data; evaluate; examine; explore; findings; gains; “mean score”; measured; method; “performed better”; posttest; pretest; “qualitative and quantitative data”; “quasi experimental design”; “randomized control trial”; research; results; sample; score; study; treatment; validity

Anm: **(1)** Vom Review-Team und Expertinnen und Experten ausgewählte Stichwörter **(2)** Über Text-mining-Verfahren ausgewählte Stichwörter **(3)** über manuelles Prototypen-Titel-/Abstractscreening ausgewählte Stichwörter

Tabelle 6 Entwicklung der Suchstrategie/Optimierung Suchbegriffe (Deutsch)

PICOS-Kategorie	Quelle	Stichwörter
<i>Welche Probandengruppe soll in welchem Setting berücksichtigt werden (Population)?</i>	1	Abiturient; bilingual; DaZ; „Deutsch als Erstsprache“; Gesamtschule; Grundschule; Gymnasium; Hauptschule; Klasse; Lerner; Mehrsprachig; Oberstufe; Primarbereich; Primarstufe; Realschule; Schule; Schüler; Sekundarbereich; Sekundarschule; Unterstufe
	2	„Deutsch als Zweitsprache“; „Englisch als Zweitsprache“; Schuljahr; „Zweitsprache Deutsch“
	3	"Deutsch als Erstsprache"; Lernende; "muttersprachliche Mitschüler"; "nichtdeutsche Herkunftssprache"; Regelunterricht
<i>Was soll getestet worden sein (Intervention)? Welche Intervention ist Gegenstand der Untersuchung?</i>	1	Immersion; „integrative Sprachförderung“; reziprokes Lehren; Scaffolding; SIOP; „sprachaufmerksamer Unterricht“; „sprachbewusster Unterricht“; Sprachbildung; „sprachliche Bildung“; „sprachliche Förderung“; „sprachsensibler Unterricht“; Translanguaging
	2	„bilinguale Erziehung“; „bilingualer Unterricht“
	3	„durchgängige Sprachbildung“; „Erlernen der Fachsprache“; „fachbezogenes textsortenbasiertes Schreiben“; „fachintegrierte Sprachbildung“; „fachintegrierte Sprachförderung“; fachspezifisch; „fachspezifisch sprachsensiblen Förderung“; „fachsprachliche Prozesse“; Fach- und Sprachförderung; „fach- und sprachintegriertes Förderkonzept“; „Schreiben im Fachunterricht“; „sprachbildende Materialien“; „sprachensible Aufgaben“; „sprachensible Unterstützung“; „Sprach- und Mathematikdidaktik“; Wortschatztraining; Wortschatzübung
<i>Nach welchen Ergebnissen soll gesucht werden (Outcome)?</i>	1	Bildungssprache; „bessere Leistung“; Biologie; Chemie; Fachsprache; Fachwissen; „fachwissenschaftliche Kompetenz“; Grammatik; Hörverstehen; Lernziele; Lesefähigkeit; Lesekompetenz; Lexikon; Literalität; Mathe; Morphem; Mündlichkeit; „phonologische Bewusstheit“; Physik; Satz; Schreibfähigkeit; Schreibkompetenz; Schriftlichkeit; Schriftsprache; „schulischer Erfolg“; Semantik; Sprachbewusstheit; Sprachfähigkeit; Sprachkompetenz; Syn-

		tax; Technik; Wort; Wortschatz
	2	„biologische Fachsprache“; „mündlicher Ausdruck“; „naturwissenschaftlicher Unterricht“; „verbale Kommunikation“
	3	„adressatenbezogene argumentative Schreibaufgaben“; „allgemeine wie fachbezogene Sprachkompetenz“; Argumentieren; „bessere Leistung“; Biologieunterricht; Chemieunterricht; Fach; Fachkommunikation; Fachkompetenzen; „fachliche und sprachliche Herausforderungen“; „fachlicher Kontext“; „fachliches Lernen“; „fachspezifische Schreibkompetenz“; „fachsprachliche Fähigkeiten“; Fachtexte; Fachunterricht; Fachwörter; Kompetenzteilbereich; „konzeptuelles Verständnis“; „literale Kompetenzen“; Leistungen; „Lernerfolg im Fachunterricht“; Lernleistung; Mathematikleistungen; Mathematikunterricht; mathematisch; „naturwissenschaftliche Fächer“; Physikunterricht; Physikverständnis; „schreibfördernde Kompetenzen“; Schriftsprache; „schulischer Erfolg“; Sprachbewusstsein; „Sprache“; „sprachliches Lernen“; Technikunterricht; Textverstehen
<i>Soll nach einem bestimmten Studiendesign gesucht werden</i> (Study Design)?	1	Analyse; Beobachtungsstudie; besser; Effekt; Einfluss; Ergebnis; Evaluationsforschung; evaluieren; Experiment; fördern; Inferenz; Intervention; Kohorten; kontrollieren; korrelieren; Längsschnitt; Maßnahme; messen; mixed-method; post; profitieren; quantitativ; Querschnitt; randomisiert; Signifikanz; Statistik; Stichprobe; Studie; Test; Triangulation; unterstützen; untersuchen; wirken
	2	„empirische Forschung“; „empirische Untersuchung“; „quantitative Forschung“
	3	Analyse; auswerten; beforschen; beobachten; „Daten erheben“; Entwicklungsforschungsprojekt; Erprobung; evidenzbasiert; Forschungsdesign; Forschungsergebnisse; Kontrollgruppendesign; Methodentriangulation; „Pre-test-Posttest-Plan“; überprüft; „vergleichende Beobachtung“; „Wissenschaftlich überprüfte Entwicklung von Lehr-Lernkonzepten“; Zwischentests

Anm: (1) Vom Review-Team und Expertinnen und Experten ausgewählte Stichwörter (2) Über Text-mining-Verfahren ausgewählte Stichwörter (3) über manuelles Prototypen-Titel-/Abstractscreening ausgewählte Stichwörter

Tabelle 7 Entwicklung der Suchstrategie/Optimierung Suchbegriffe (NOT-Begriffe Englisch)

PICOS-Kategorie	NOT-Stichwörter
<i>Welche Probanden- gruppe soll in wel- chem Setting berück- sichtigt werden (Po- pulation)</i>	adult; apprenticeship; autism; "career training"; disorders; dyslexic; "foreign language class"; "foreign language learn"; "foreign language teach"; "further education"; "in firm training"; "in plant training"; "in service education"; "in service train- ing"; "industrial training"; "job training"; "out of school education"; parent; preschooler; university; varsity; vocational; "with disability"
<i>Was soll getestet worden sein (Inter- vention)? Welche Intervention ist Gegenstand der Untersuchung?</i>	"foreign language education"
<i>Nach welchen Ergeb- nissen soll gesucht werden (Outcome)?</i>	
<i>Soll nach einem be- stimmten Studiende- sign gesucht werden (Study Design)?</i>	"discussion paper"; "narrative review"; "literature review"

Anm: (1) Vom Review-Team und Expertinnen und Experten ausgewählte Stichwörter (2) Über Text-mining-Verfahren ausgewählte Stichwörter (3) über manuelles Prototypen-Titel-/Abstractscreening ausgewählte Stichwörter

Tabelle 8 Entwicklung der Suchstrategie/Optimierung Suchbegriffe (NOT-Begriffe Deutsch)

PICOS-Kategorie	NOT-Stichwörter
Welche Probanden- gruppe soll in wel- chem Setting berück- sichtigt werden (Pop- ulation)	Ausbildung; Autismus; „außerschulische Bildung“; Behinderte; beruflich; Berufsausbildung; Berufsbildung ; betrieblich; Dyslexie; Eltern; Erwachsene; Fortbildung; Fremdsprache; Hochschule; Industrie; Krankheit; Legastheniker; „Lese-Rechtschreib-Schwäche“; LRS; „mit Behinderung“; Störung; Universität; Vorschule; Weiterbildung
Was soll getestet worden sein (Inter- vention)? Welche Intervention ist Gegenstand der Untersuchung?	DaZ-Unterricht; Deutsch als Zweitsprache Unterricht; Fremdsprache
Nach welchen Ergeb- nissen soll gesucht werden (Outcome)?	
Soll nach einem be- stimmten Studiende- sign gesucht werden (Study Design)?	Diskussionspapier; Literaturüberblick; Übersichtsarbeit

Anm: (1) Vom Review-Team und Expertinnen und Experten ausgewählte Stichwörter (2) Über Text-mining-Verfahren ausgewählte Stichwörter (3) über manuelles Prototypen-Titel-/Abstractscreening ausgewählte Stichwörter

Call for Preprints (9.8.-30.11.2019)

Das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache sucht (noch) nicht publizierte Studien zur Wirkung von sprachsensiblen Unterrichtsansätzen (SIOP, Scaffolding etc.).

Das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache führt ein systematisches Review zur Wirkung didaktisch-methodischer Ansätze des sprachsensiblen Unterrichts (z. B. SIOP, Scaffolding) durch. In dem Review wird beabsichtigt, den bisherigen Forschungsstand zu vorab definierten Kriterien zusammenzustellen, zu bewerten und in einer Forschungssynthese zusammenzufassen und zu diskutieren.

Ergänzend zu den hierfür erfolgten Recherchen in elektronischen Datenbanken sucht das Projekt-Team (noch) nicht publizierte Manuskripte/Arbeitspapiere/Projektberichte/Dissertationen (keine BA-/MA-Arbeiten), die (noch) nicht in Datenbanken zu finden sind.

Gesucht werden (quasi-)experimentelle oder Beobachtungsstudien im Kontrollgruppendesign, in denen

- die Wirkung didaktisch-methodischer Ansätze des sprachsensiblen Unterrichts
- im Primar- oder Sekundarschulalter

statistisch untersucht wird.

Wenn Sie eine solche Studie durchgeführt haben oder wenn erste Ergebnisse einer aktuell noch laufenden Untersuchung vorliegen, würden wir uns über eine Zusendung der Dokumente sehr freuen. Die übermittelten Studien werden vom Projekt-Team überprüft und, sofern sie die Einschlusskriterien des Reviews treffen, in die Forschungssynthese des systematischen Reviews einbezogen, d. h. zusammengefasst und diskutiert. Die Publikation der Ergebnisse ist für 2021 vorgesehen. Selbstverständlich werden Ihre Einsendungen ausschließlich im Rahmen des Reviews verwendet und nicht an Dritte weitergegeben.

Bitte schicken Sie Ihre Einsendungen bis zum 30.9.2019 an Leonie Twente (leo-nie.twente@mercator.uni-koeln.de). Bei Rückfragen wenden Sie sich bitte an Till Woerfel (till.woerfel@mercator.uni-koeln.de).

Vielen Dank für Ihre Unterstützung!

Do you have unpublished studies examining the effectiveness of language integrated strategies (SIOP, scaffolding) in classrooms?

The Mercator Institute for Literacy and Language Education at the University of Cologne, Germany, is conducting a systematic review on the effectiveness of language integrated strategies (e.g. scaffolding, Sheltered Instruction Observation Protocol (SIOP)) in classrooms. In this review, it is intended to collate, critically appraise and synthesize existing research evidence ac-

according to pre-defined criteria.

To complement our electronic database search, we are looking for manuscripts/working papers/project reports/dissertations (except for BA-/MA-thesis) that have not (yet) been published or submitted for publication (and are not (yet) indexed in databases).

We are interested in (quasi-)experimental and observational studies using a control group design that (statistically) examine

- the effectiveness of concepts of instruction that integrate language support and subject teaching
- for children of primary or secondary school age.

If you have carried out this type of study or if first results of an ongoing study are available, we would like to kindly request the document. Submitted studies will be reviewed by the project team; studies that match the review inclusion criteria will be included in the final review synthesis (i.e. summarized and discussed). Publication of the results is planned for 2021.

Of course, your submissions will only be used within the scope of the review and will not be passed to third parties.

Please send documents by September 30th, 2019 to Leonie Twente at leonie.twente@mercator.uni-koeln.de.

If you have any questions, please contact Till Woerfel at till.woerfel@mercator.uni-koeln.de.

Thank you for your support!