

Bertram, Victoria

Implementationserfolg beim mehrsprachigkeitssensiblen Reziproken Lehren. Untersuchung von Wiedergabetreue und Akzeptanz aus Schüler*innenperspektive

Münster ; New York : Waxmann 2025, 181 S. - (Empirische Erziehungswissenschaft; 86)



Quellenangabe/ Reference:

Bertram, Victoria: Implementationserfolg beim mehrsprachigkeitssensiblen Reziproken Lehren. Untersuchung von Wiedergabetreue und Akzeptanz aus Schüler*innenperspektive. Münster ; New York : Waxmann 2025, 181 S. - (Empirische Erziehungswissenschaft; 86) - URN: urn:nbn:de:0111-pedocs-340330 - DOI: 10.25656/01:34033; 10.31244/9783818850500

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-340330>

<https://doi.org/10.25656/01:34033>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Informationszentrum (IZ) Bildung

E-Mail: pedocs@dipf.de

Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Victoria Bertram

Implementationserfolg beim mehrsprachigkeitssensiblen Reziproken Lehren

**Untersuchung von Wiedergabetreue
und Akzeptanz aus Schüler*innenperspektive**

Empirische Erziehungswissenschaft

herausgegeben von

Rolf Becker, Wilfried Bos, Hartmut Ditton,
Cornelia Gräsel, Eckhard Klieme, Kai Maaz,
Thomas Rauschenbach, Hans-Günther Roßbach,
Knut Schwippert, Ludwig Stecher, Christian Tarnai,
Rudolf Tippelt, Rainer Watermann

Band 86

Victoria Bertram

Implementationserfolg beim mehrsprachigkeitssensiblen Reziproken Lehren

Untersuchung von Wiedergabetreue und
Akzeptanz aus Schüler*innenperspektive



Waxmann 2025
Münster • New York

Die Open-Access-Veröffentlichung des E-Books wurde gefördert durch den Open-Access-Publikationsfonds des DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation in Frankfurt am Main.

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Empirische Erziehungswissenschaft, Band 86

ISSN 1862-2127

Print-ISBN 978-3-8188-0050-5

E-Book-ISBN 978-3-8188-5050-0

<https://doi.org/10.31244/9783818850500>

Waxmann Verlag GmbH, 2025
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Pleßmann Design, Ascheberg
Satz: Roger Stoddart, Münster

Dieses E-Book ist unter der Lizenz CC BY 4.0 DE open access verfügbar.



Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Zitate etc.) sind von der CC-Lizenz ausgenommen und für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen.

Danksagung

An dieser Stelle möchte ich mich herzlich bei allen bedanken, die mich bei der Erstellung dieser Dissertation unterstützt haben. Mein größter Dank gilt meinen Gutachterinnen Dominique Rauch und Jasmin Decristan, die mich mit ihrer kompetenten und äußerst wertschätzenden Betreuung bei allen Fragen unterstützt und mir die Möglichkeit gegeben haben, dieses Thema zu verfolgen. Ein großes Dankeschön möchte ich auch an alle Kolleg*innen und Doktorand*innen am DIPF, dem Leibniz-Institut für Bildungsforschung und Bildungsinformation in Frankfurt am Main, richten, für die engagierte Unterstützung und den Austausch in diversen Kolloquien, Terminen oder bei einer Tasse Kaffee. Ohne die Lehrkräfte und Schüler*innen, die am Projekt meRLe teilgenommen haben, hätte diese Dissertation nicht entstehen können. Daher möchte ich mich auch bei ihnen sehr herzlich bedanken. Zudem geht mein Dank auch an alle studentischen Hilfskräfte, die im Projekt meRLe mitgearbeitet, die Datenerhebung auch in stressigen Zeiten erfolgreich gemeistert und mich mit ihrer sorgfältigen Arbeit bis Projektende begleitet haben. Von ganzem Herzen möchte ich abschließend meinem Mann und meiner Familie danken, die mich bedingungslos unterstützen und mich immer darin bestärkt haben, meine Interessen zu verfolgen.

Zusammenfassung

Zum Umgang mit sprachlicher Heterogenität im Unterricht werden bereits in der Grundschule innovative didaktische Konzepte benötigt, um das gesamte linguistische Repertoire der Schüler*innen in den Lernprozess einzubeziehen und zum Abbau von Bildungsbenachteiligungen mehrsprachiger Schüler*innen beizutragen. Damit innovative Konzepte in der Praxis wirksam umgesetzt werden können, ist die Untersuchung von förderlichen und hinderlichen Faktoren im Implementationsprozess von zentraler Bedeutung. Der Forschungsstand zeigt, dass Interventionen besonders erfolgreich sind, wenn sie mit hoher Wiedergabetreue eingesetzt (Durlak & DuPre, 2008; O'Donnell, 2008) und von Teilnehmer*innen akzeptiert werden (Eckert et al., 2021). Die Wiedergabetreue und Akzeptanz stellen damit zentrale Facetten des Implementationserfolgs dar.

In der schulischen Implementationsforschung ist bislang eine Fokussierung auf die Lehrkräfte zu beobachten, die eine Intervention im Unterricht umsetzen. Die Perspektive der Schüler*innen wird häufig vernachlässigt. Vor dem Hintergrund konstruktivistischer Lerntheorien stellen diese jedoch wichtige Akteur*innen im Implementationsprozess dar, da der Lernerfolg und somit die Wirksamkeit einer Intervention im unterrichtlichen Kontext maßgeblich von individuellen Wahrnehmungs- und Nutzungsprozessen der Schüler*innen abhängt (Schrader et al., 2020; Vieluf et al., 2020). Besonders für die Untersuchung mehrsprachigkeitssensibler Unterrichtskonzepte ist es relevant zu untersuchen, wie Schüler*innen diese innovativen didaktischen Konzepte wahrnehmen und nutzen, da Lehrkräfte der Integration von Mehrsprachigkeit in ihrem Unterricht häufig skeptisch und ängstlich gegenüberstehen und bislang nur wenig über gelingende Implementationsfaktoren bekannt ist (Bredthauer & Engfer, 2018; Hilbe & Kutzelmann, 2017).

Die vorliegende Arbeit knüpft an diese Forschungslücken an und untersucht Facetten des Implementationserfolgs eines innovativen und mehrsprachigkeitssensiblen Unterrichtskonzepts – dem mehrsprachigkeitssensiblen Reziproken Lehren – aus Schüler*innenperspektive. Hierbei wird die kooperative Methode des Reziproken Lehrens, die die Förderung der Deutsch-Lesekompetenz durch die Instruktion von Lesestrategien zum Ziel hat, mit mehrsprachigkeitssensiblen Impulsen verknüpft, sodass alle Sprachen der Schüler*innen aktiv in den Lernprozess integriert werden (Decristan et al., 2022). Anhand der Daten aus dem Projekt meRLe („Förderung der Deutsch-Lesekompetenz durch mehrsprachigkeitssensibles Reziprokes Lehren im Grundschulunterricht“), in dem das innovative Unterrichtskonzept konzipiert und als Unterrichtseinheit mit zwölf Unterrichtsstunden in vierten Klassen implementiert wurde, werden in der vorliegenden Arbeit Fragen der Implementationsforschung und zum Umgang mit sprachlicher Vielfalt im Unterricht beleuchtet. Die Wiedergabetreue und Akzeptanz der Schüler*innen werden dabei als zentrale Facetten von Implementationserfolg untersucht. Als Wiedergabetreue wird die Nutzung von Promptkarten während des Lesestrategietrainings in den Blick genommen, die ein wichtiges Kernelement zur Anwendung der Lesestrategie-

gien im kooperativen Setting darstellen. Die Promptkarten sind bilingual gestaltet, sodass mehrsprachige Schüler*innen diese auf Deutsch und auf ihren Familiensprachen nutzen können. Die Akzeptanz stellt ein Gesamturteil der Intervention dar und wird als affektive und kognitive Dimension konzeptualisiert.

Die Fragestellungen der Arbeit beziehen sich zum einen auf die Bedeutung der Facetten des Implementationserfolgs für den Lernerfolg der Schüler*innen und darauf, wie Unterschiede im Implementationserfolg durch individuelle Merkmale erklärt werden können. Hierbei wird insbesondere die Mehrsprachigkeit der Schüler*innen in den Blick genommen. Zum anderen wird die Aktivierung der Mehrsprachigkeit über die Nutzung der Promptkarten in einer anderen Sprache untersucht und Zusammenhänge zum Lernerfolg sowie individuelle Determinanten der mehrsprachigen Nutzung werden analysiert. Zur Auswertung der Fragestellungen werden Daten aus dem Projekt meRLe von 634 Schüler*innen aus 36 vierten Klassen einbezogen, in denen das mehrsprachigkeitssensible Reziproke Lehren von einer Lehrkraft umgesetzt wurde. 431 mehrsprachige Schüler*innen, die die Promptkarten in ihren Familiensprachen vorliegen hatten, werden zudem differenziert untersucht.

Die zentralen Ergebnisse zeigen, dass die Nutzung der Promptkarten, vermittelt über die affektive Akzeptanz, mit höheren Ausprägungen in der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention einhergeht. Bei mehrsprachigen Schüler*innen zeigen sich zudem höhere Ausprägungen in der affektiven und kognitiven Akzeptanz, die jedoch vermittelt werden durch niedrigere Ausprägungen im Leseverstehen und positivere Einstellungen zu Mehrsprachigkeit. Darüber hinaus können für mehrsprachige Schüler*innen, die die Promptkarten in ihrer Familiensprache genutzt haben, motivationale Potenziale für den Lernprozess abgeleitet werden, da die Aktivierung der Mehrsprachigkeit mit höheren Ausprägungen der Lesefreude am Ende der Intervention einhergeht. Zusätzlich zeigt sich, dass die Nutzung der Promptkarten in einer anderen Sprache durch positive Einstellungen zu Mehrsprachigkeit vorhergesagt wird.

Die Ergebnisse verdeutlichen die zentrale Rolle der Schüler*innen im Implementationsprozess, die in zukünftigen Studien stärkere Berücksichtigung finden sollte. Über Effektivitätsuntersuchungen von Interventionen hinaus können hieraus wichtige Implikationen für die Weiterentwicklung und adaptive Gestaltung von innovativen didaktischen Konzepten abgeleitet werden. Die Ergebnisse leisten zudem einen wichtigen Beitrag zum Umgang mit Mehrsprachigkeit im Unterricht und verdeutlichen die zentrale Rolle eines wertschätzenden Klimas für sprachliche Heterogenität, das im Rahmen mehrsprachigkeitssensibler Unterrichtskonzepte aufgebaut werden sollte.

Inhalt

1	Einleitung	13
2	Mehrsprachigkeit und Reziprokes Lehren	16
2.1	Mehrsprachigkeit im schulischen Kontext	16
2.1.1	Begriffsbestimmung Mehrsprachigkeit	17
2.1.2	Ganzheitliche Perspektive auf Mehrsprachigkeit	18
2.1.3	Einstellungen zu Mehrsprachigkeit	20
2.1.4	Mehrsprachigkeit als Ressource für das Lernen	22
2.2	Mehrsprachigkeitssensibles Unterrichten	23
2.2.1	Theoretische Ansätze und praktische Impulse für einen mehrsprachigkeitssensiblen Unterricht	23
2.2.2	Forschungsergebnisse zu mehrsprachigkeitssensiblen Unterrichtskonzepten	26
2.3	Förderung der Lesekompetenz durch Reziprokes Lehren	29
2.3.1	Facetten von Lesekompetenz	30
2.3.2	Kooperatives Lernen	31
2.3.3	Instruktion von Lesestrategien beim Reziproken Lehren	32
2.4	Mehrsprachigkeitssensibles Reziprokes Lehren	34
3	Implementationserfolg von Interventionen im Unterricht	36
3.1	Implementation von Innovationen	36
3.2	Definitionen von Implementationserfolg	41
3.3	Wiedergabetreue	43
3.3.1	Dimensionen der Wiedergabetreue	44
3.3.2	Untersuchung der Wiedergabetreue durch Schüler*innen	46
3.4	Akzeptanz	47
3.4.1	Akzeptanz als einstellungsähnliches Konstrukt	47
3.4.2	<i>Treatment Acceptability</i> zur Untersuchung von Akzeptanzurteilen im Unterricht	49
3.4.3	Erfassung der Akzeptanz von Interventionen im Unterricht aus Schüler*innenperspektive	51
3.5	Facetten des Implementationserfolgs in Unterrichtsinterventionen: Zusammenhänge und Determinanten	52
3.5.1	Konzeptuelle Modelle der Implementations- und Unterrichtsforschung als theoretischer Rahmen	53
3.5.2	Zusammenhänge zwischen Wiedergabetreue, Akzeptanz und Schüler*innenoutcomes	57
3.5.3	Determinanten für Implementationserfolg im Unterricht	59

4	Fragestellungen	62
4.1	Fragestellung 1: Sagen die Facetten des Implementationserfolgs interventionsspezifische Schüler*innenoutcomes vorher?	63
4.1.1	Direkte Vorhersagekraft der Nutzung der Promptkarten und Akzeptanz für interventionsspezifische Schüler*innenoutcomes	64
4.1.2	Vermittelnde Rolle der Akzeptanz	65
4.2	Fragestellung 2: Welche individuellen Schüler*innenmerkmale sagen die Facetten des Implementationserfolgs vorher?	66
4.2.1	Individuelle Determinanten der Nutzung der Promptkarten.	67
4.2.2	Individuelle Determinanten der Akzeptanz	68
4.3	Fragestellung 3: Sagt die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen interventionsspezifische Schüler*innenoutcomes vorher?	69
4.3.1	Direkte Vorhersagekraft der Nutzung der Promptkarten in einer anderen Sprache für interventionsspezifische Schüler*innenoutcomes	69
4.3.2	Vermittelnde Rolle der Akzeptanz	70
4.4	Fragestellung 4: Welche individuellen Schüler*innenmerkmale sagen die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen vorher?	71
5	Methodisches Vorgehen	72
5.1	Datengrundlage.	72
5.1.1	Intervention zum mehrsprachigkeitssensiblen Reziproken Lehren	72
5.1.2	Design und Datenerhebung.	76
5.1.3	Implementation der Unterrichtseinheit.	78
5.2	Stichprobe	80
5.3	Erhebungsinstrumente und Operationalisierung.	81
5.3.1	Nutzung der Promptkarten	81
5.3.2	Akzeptanz.	82
5.3.3	Individuelle Schüler*innenmerkmale	86
5.4	Statistische Auswertung	92
5.4.1	Berücksichtigung der hierarchischen Datenstruktur	92
5.4.2	Konfirmatorische Faktorenanalysen.	95
5.4.3	Regressionsmodelle	98
5.4.4	Mediationsmodelle	101
5.4.5	Umgang mit fehlenden Werten.	104
5.4.6	Signifikanztestung.	105
6	Ergebnisse	106
6.1	Deskriptive Statistik	106
6.1.1	Nutzung der Promptkarten	106
6.1.2	Akzeptanz.	107

6.1.3	Individuelle Schüler*innenmerkmale	107
6.1.4	Bivariate Zusammenhänge der untersuchten Konstrukte.	109
6.2	Fragestellung 1: Sagen die Facetten des Implementationserfolgs interventionsspezifische Schüler*innenoutcomes vorher?	112
6.2.1	Direkte Vorhersagekraft der Nutzung der Promptkarten und Akzeptanzdimensionen für interventionsspezifische Schüler*innenoutcomes.	112
6.2.2	Vermittelnde Rolle der Akzeptanz	115
6.2.3	Zusammenfassung	117
6.3	Fragestellung 2: Welche individuellen Schüler*innenmerkmale sagen die Facetten des Implementationserfolgs vorher?	117
6.3.1	Individuelle Determinanten der Nutzung der Promptkarten.	118
6.3.2	Individuelle Determinanten der Akzeptanz	120
6.3.3	Zusammenfassung	124
6.4	Fragestellung 3: Sagt die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen interventionsspezifische Schüler*innenoutcomes vorher?	125
6.4.1	Direkte Vorhersagekraft der Nutzung der Promptkarten in einer anderen Sprache für interventionsspezifische Schüler*innenoutcomes.	125
6.4.2	Vermittelnde Rolle der Akzeptanz	128
6.4.3	Zusammenfassung	129
6.5	Fragestellung 4: Welche individuellen Schüler*innenmerkmale sagen die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen vorher?	130
6.5.1	Individuelle Determinanten für die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen	130
6.5.2	Zusammenfassung	131
7	Diskussion	132
7.1	Evaluation des Implementationserfolgs auf Schüler*innenebene	133
7.2	Bedeutung der Mehrsprachigkeit für den Implementationserfolg.	137
7.2.1	Mehrsprachigkeit als Prädiktor für die Facetten des Implementationserfolgs	138
7.2.2	Aktivierung der Mehrsprachigkeit durch die Nutzung der bilingualen Promptkarten in einer anderen Sprache	140
7.3	Praktische Implikationen.	142
7.4	Limitationen und Ausblick	144
7.5	Fazit.	149

Literatur151

Abbildungsverzeichnis171

Tabellenverzeichnis172

Abkürzungen.....174

Anhang175

1 Einleitung

Sprachliche Vielfalt stellt eine der vielen Heterogenitätsdimensionen im Klassenzimmer dar, auf die Lehrkräfte mit ihrem Unterricht eingehen müssen. Nach Henschel et al. (2019) finden sich in Schulklassen in Großstadtbezirken häufig bis zu acht unterschiedliche Familiensprachen neben dem Deutschen. Prediger et al. (2021, S. 167) bezeichnen dies in Anlehnung an das Konzept der Superdiversität nach Vertovec (2007) als „superdiverse[n] Sprachkontext“, in dem unterschiedliche sprachbiografische Kombinationen von (migrationsbedingter) Mehrsprachigkeit zusammenkommen. In der unterrichtlichen Praxis wird diese sprachliche Heterogenität und lebensweltliche Mehrsprachigkeit der Schüler*innen allerdings selten aufgegriffen (Bredthauer & Engfer, 2018; Kropp, 2017; Krumm, 2016). Stattdessen ist die Sprachnutzung im schulischen Kontext vor dem Hintergrund eines *monolingualen Habitus* (Gogolin, 1994) durch Einsprachigkeit gekennzeichnet und mehrsprachige Schüler*innen werden dadurch in ihrer natürlichen Sprachpraxis eingeschränkt. Dies resultiert in einer strukturellen Benachteiligung mehrsprachiger Schüler*innen, da sie ihre sprachlichen Potenziale im Lernprozess nicht einbringen können (Fürstenau, 2017; García & Wei, 2014). Die Berücksichtigung von Familiensprachen im Unterricht kann für Schüler*innen jedoch lernförderlich sein und positiv zur Identitätsentwicklung beitragen (Grosjean, 2013; Honneth, 1992; Krumm, 2016). Um gerechte Bildungschancen für alle Schüler*innen gemäß ihrer individuellen Lernvoraussetzungen und Potenziale zu ermöglichen, wird daher bereits im Grundschulunterricht die Integration von Mehrsprachigkeit gefordert (z. B. Dirim & Wegener, 2016; Fürstenau, 2011; KMK, 2015). Dafür werden innovative und mehrsprachigkeitssensible Unterrichtskonzepte benötigt, mit denen alle Sprachen der Schüler*innen im Lernprozess integriert werden können, ohne dass die Lehrkraft die Sprachen selbst sprechen muss. Kooperative Lernarrangements bieten hierfür einen Rahmen, in dem sich Schüler*innen beim Lernen gegenseitig unterstützen und ihre Sprachen aktiv nutzen können.

Damit sich innovative didaktische Konzepte in der unterrichtlichen Praxis erfolgreich etablieren und langfristig genutzt werden, müssen diese nicht nur eine hohe Wirksamkeit, sondern auch eine gute Umsetzbarkeit aufweisen. Um förderliche und hinderliche Faktoren beim Übergang in das praktische Feld zu identifizieren, ist es daher von Bedeutung, neben der Wirksamkeit auch die Implementation durch die Akteur*innen zu untersuchen. Die Implementation von Innovationen im Schulsystem ist ein komplexer Prozess, dessen Erfolg an verschiedenen Kriterien gemessen werden kann. Über unterschiedliche Forschungsperspektiven hinweg haben sich die Wiedergabetreue und Akzeptanz als zentrale Facetten von Implementationserfolg herausgebildet. Häufig wird dabei die Perspektive der Lehrkräfte als umsetzende Akteur*innen fokussiert. Vor dem Hintergrund konstruktivistischer Lerntheorien und insbesondere im Rahmen kooperativer Lernarrangements stellen Schüler*innen jedoch auch aktive Akteur*innen im Lernprozess dar, die die Implementation mitgestalten können (Schrader et al., 2020; Vieluf et

al., 2020). Die Wirkung und der Erfolg der Maßnahme sind dabei maßgeblich von der Eigenaktivität und Nutzung der Lernenden abhängig. Die Berücksichtigung der Schüler*innenperspektive kann somit ermöglichen, den Übergang in das praktische Feld über die Umsetzung der Lehrkräfte hinaus zu erfassen und Faktoren abzuleiten, die für Prozesse der Adaption und Weiterentwicklung genutzt werden können. Bislang ist jedoch wenig darüber bekannt, wie Schüler*innen mehrsprachigkeitssensible Konzepte wahr- und annehmen (Hilbe & Kutzelmann, 2017). Die vorliegende Dissertation knüpft an diese Forschungslücken an und untersucht Facetten des Implementationserfolgs einer Intervention zum mehrsprachigkeitssensiblen Reziproken Lehren (Decristan et al., 2022) aus Schüler*innenperspektive.

Beim mehrsprachigkeitssensiblen Reziproken Lehren wird das Reziproke Lehren, was sich als wirksame kooperative Instruktionmethode zur Förderung der Lesekompetenz erwiesen hat (z.B. Rosenshine & Meister, 1994), mit mehrsprachigkeitssensiblen Impulsen angereichert, um das gesamte sprachliche Repertoire der Schüler*innen im Lernprozess zu integrieren. Durch die Etablierung eines wertschätzenden Klimas für Mehrsprachigkeit, die Integration von mehrsprachigem Unterrichtsmaterial und die Zuordnung von Sprachpartnerschaften innerhalb des kooperativen Settings, werden mehrsprachigkeitssensible Impulse zur Aktivierung der sprachlichen Ressourcen eingebunden (Decristan et al., 2022).

Ziel der Arbeit ist es, die Implementation des didaktischen Konzepts aus Schüler*innenperspektive zu untersuchen. Hierbei sollen zunächst aus einem breiteren Blick der Implementationsforschung die Wiedergabetreue und Akzeptanz als Facetten des Implementationserfolgs auf Schüler*innenebene untersucht werden. Fokussiert wird hierbei, wie diese Facetten des Implementationserfolgs mit dem Lernerfolg der Schüler*innen zusammenhängen und wie diese durch individuelle Lernvoraussetzungen der Schüler*innen erklärt werden können. Die individuelle Mehrsprachigkeit der Schüler*innen stellt dabei einen besonderen Fokus dar. Zusätzlich werden in einer differenziellen Perspektive die Implementation des mehrsprachigen Unterrichtsmaterials in den Blick genommen und auch hier Zusammenhänge zum Lernerfolg und die Bedeutung individueller Lernvoraussetzungen für mehrsprachige Schüler*innen untersucht. Die Ergebnisse sollen dazu beitragen, die Rolle der Schüler*innen im Implementationsprozess besser zu verstehen und Implikationen für die Implementation mehrsprachigkeitssensibler Unterrichtskonzepte abzuleiten.

In der vorliegenden Untersuchung werden Daten einer quasi-experimentellen Interventionsstudie (Projekt meRLe) quantitativ analysiert, in der das mehrsprachigkeitssensible Reziproke Lehren in vierten Grundschulklassen implementiert wurde. Das didaktische Konzept besteht aus einer Einheit von zwölf Unterrichtsstunden, die je 45 Minuten umfassen. Kern des Konzepts ist das Einüben von Lese- und Lernstrategien im kooperativen Setting. Schüler*innen nutzen dabei sogenannte Promptkarten, die die wichtigsten Schritte zum Ausführen der Lesestrategien „Fragen stellen“ und „Zusammenfassen“ angeben. Als mehrsprachigkeitssensibler Impuls sind die Promptkarten bilingual gestaltet, sodass die

Strategien auf Deutsch und zusätzlich auf einer Familiensprache der Schüler*innen dargestellt sind.

Zur theoretischen Rahmung wird im folgenden Kapitel zunächst der Themenkomplex des mehrsprachigkeitssensiblen Reziproken Lehrens beschrieben (vgl. Kapitel 2). Hierbei wird auf den Umgang mit Mehrsprachigkeit im schulischen Kontext eingegangen, bevor theoretische Ansätze und praktische Impulse für einen mehrsprachigkeitssensiblen Unterricht abgeleitet werden. Anschließend wird das didaktische Konzept des Reziproken Lehrens zur Förderung der Lesekompetenz dargestellt und Möglichkeiten zur Integration von Mehrsprachigkeit beleuchtet. Der zweite Themenkomplex beschäftigt sich mit dem Implementationserfolg von Interventionen im Unterricht (vgl. Kapitel 3). Hierbei werden die Wiedergabetreue und Akzeptanz als zentrale Facetten von Implementationserfolg näher beschrieben und die Erfassung der Konstrukte aus Perspektive der Schüler*innen beleuchtet. Aktuelle Forschungsbefunde zu Zusammenhängen und Determinanten der Facetten des Implementationserfolgs im Kontext von Unterrichtsinterventionen schließen das Kapitel ab. Auf Basis des in den Kapiteln 2 und 3 dargestellten theoretischen Hintergrundes werden die Forschungsfragen und Hypothesen der Arbeit abgeleitet (vgl. Kapitel 4). Anschließend wird das methodische Vorgehen der Arbeit beschrieben, das die Darstellung und Implementation der untersuchten Unterrichtseinheit umfasst sowie die Beschreibung der Stichprobe, der Erhebungsinstrumente und statistischen Verfahren zur Auswertung (vgl. Kapitel 5). Nach Darstellung der Ergebnisse (vgl. Kapitel 6) werden diese in einer zusammenfassenden Diskussion interpretiert und kritisch beleuchtet (vgl. Kapitel 7). Die Arbeit schließt in diesem Rahmen mit der Ableitung praktischer Implikationen und Reflexion von Limitationen, die mit ausblickenden Forschungsfragen einhergehen. Die zentralen Befunde der Arbeit werden im Fazit abschließend zusammengefasst.

2 Mehrsprachigkeit und Reziprokes Lehren

Lesen ist eine Schlüsselkompetenz für bildungsbezogene Aktivitäten und gesellschaftliche Teilhabe (Artelt et al., 2007; Becker-Mrotzek et al., 2019; Lenhard & Artelt, 2009). Die Vermittlung und Förderung der Lesekompetenz ist daher eine zentrale Aufgabe schulischer Bildung. Bildungsstudien im deutschsprachigen Raum haben gezeigt, dass Schüler*innen mit Migrationshintergrund häufig über geringere Kompetenzen in der Unterrichtssprache verfügen und folglich schlechtere Leistungen erzielen als ihre Mitschüler*innen ohne Migrationshintergrund (Edele et al., 2020). Dies ist auch in Bezug auf die Deutsch-Lesekompetenz zu beobachten (Wendt & Schwippert, 2017). Schüler*innen mit Migrationshintergrund sind daher eine besonders benachteiligte Gruppe im Schulsystem (Edele et al., 2020) und der Leseförderung unter Berücksichtigung von Mehrsprachigkeit kommt eine zentrale Rolle zu. Eine wirksame Methode der Leseförderung stellt das Reziproke Lehren nach Palinscar und Brown (1984) dar, bei dem Lesestrategien in kleinen Gruppen kooperativ eingeübt werden. In der vorliegenden Arbeit wird das innovative Konzept des *mehrsprachigkeitssensiblen* Reziproken Lehrens untersucht, welches das Reziproke Lehren mit mehrsprachigkeitssensiblen Impulsen verknüpft, um mehrsprachigen Schüler*innen den Einbezug aller ihrer sprachlichen Potenziale im Lernprozess zu ermöglichen.

Zur Einleitung in die Thematik der vorliegenden Arbeit wird in diesem Kapitel zunächst der Umgang mit Mehrsprachigkeit im schulischen Kontext beschrieben (vgl. Kapitel 2.1). Darauf aufbauend werden theoretische Ansätze sowie praktische Impulse für einen mehrsprachigkeitssensiblen Unterricht, in dem alle sprachlichen Ressourcen der Schüler*innen aufgegriffen werden, dargestellt (vgl. Kapitel 2.2). Abschließend wird die Methode des Reziproken Lehrens zur Förderung der Lesekompetenz eingeführt (vgl. Kapitel 2.3), die Themenfelder werden verknüpft und die Methode des mehrsprachigkeitssensiblen Reziproken Lehrens vorgestellt (vgl. Kapitel 2.4).

2.1 Mehrsprachigkeit im schulischen Kontext

Das folgende Kapitel thematisiert den Umgang mit Mehrsprachigkeit im schulischen Kontext. Hierbei wird zunächst der Begriff *Mehrsprachigkeit* für die vorliegende Arbeit bestimmt und der Wandel von traditionellen Auffassungen über Sprachen hin zu einem ganzheitlichen Blick auf Mehrsprachigkeit beschrieben. Weiterhin werden Einstellungen von Lehrkräften zum Einbezug von Mehrsprachigkeit in den Unterricht angesprochen und abschließend Funktionen und Potenziale aufgeführt, die Mehrsprachigkeit im Lernprozess einnehmen kann.

2.1.1 Begriffsbestimmung Mehrsprachigkeit

Mehrsprachigkeit ist ein komplexes, dynamisches und interdisziplinäres Phänomen, das in unterschiedlichen Forschungsperspektiven untersucht wird. Mit linguistischen, psychologischen, soziologischen oder erziehungswissenschaftlichen Schwerpunkten sind Definitionen und Konzepte entsprechend heterogen (Cenoz, 2013; Jessner & Allgäuer-Hackl, 2015; Lengyel, 2017). Im Allgemeinen können Personen als mehrsprachig definiert werden, die im Alltag mehr als eine Sprache nutzen, um zu kommunizieren und zu interagieren (Grosjean, 2020; Riehl, 2014). In einigen Definitionen wird Mehrsprachigkeit erst ab drei Sprachen definiert (z. B. Ahrenholz, 2020; Keim, 2012; Kemp, 2009) und in diesem Sinne vom Konzept der Bilingualität, welches terminologisch nur zwei Sprachen fokussiert, abgegrenzt. Da Sprachbiografien jedoch vielfältig und komplex sind, wird in der vorliegenden Arbeit bereits ab mehr als einer Sprache von Mehrsprachigkeit gesprochen, um alle Formen sprachlicher Heterogenität zu berücksichtigen.

Frühe Definitionen von Mehrsprachigkeit sind unter dem Begriff von Bilingualität zu finden. Dabei galten Sprachkenntnisse auf einem muttersprachlichen Niveau in beiden Sprachen als zentrales Definitionscharakteristikum (z. B. Bloomfield, 1984). Dieses idealistische Konzept wird als „balanced bilingualism“ (García, 2009, S. 44) beschrieben. In der Realität zeigen sich jedoch vielseitige Ausprägungen von Mehrsprachigkeit und kaum existieren dabei ausgeglichene hohe Kompetenzen in beiden bzw. allen Sprachen. Aktuelle Definitionen betonen dagegen eine funktionale Perspektive von Mehrsprachigkeit (Kropp, 2017), indem Mehrsprachigkeit als Kompetenz definiert wird, mit zwei oder mehr Sprachen zu kommunizieren (Oksaar, 2003), bzw. „bedeutungsvolle Äußerungen zu erzeugen“ (Grosjean, 2020, S. 14). Der Fokus der Definition liegt hier auf dem Sprachgebrauch, wobei ein Mindestmaß an Sprachkompetenz nicht genannt wird. Auch Franceschini (2015, S. 291) sieht das bedeutende Merkmal mehrsprachiger Personen in der „kommunikativen Expertise“, die bei Einsprachigen aus „einem Bündel von Varietäten“ einer gemeinsamen Sprache bestehen kann und sich bei mehrsprachigen Sprecher*innen über unterschiedliche Sprachen und Varietäten erstreckt. Dabei kann es sich sowohl um unterschiedliche Dialekte oder migrationsbedingte Mehrsprachigkeit handeln als auch um in der Schule gelernte Fremdsprachen. Denn Mehrsprachigkeit kann unterschiedliche Gründe haben und zu unterschiedlichen Zeiten im Lebenslauf erworben werden. Begrifflich wird hierbei zwischen *lebensweltlicher* und *fremdsprachlicher* Mehrsprachigkeit differenziert (Gogolin et al., 2020). Im schulischen Kontext wird die fremdsprachliche Mehrsprachigkeit, die durch das Erlernen einer (häufig prestigehohen) Schulfremdsprache entsteht, positiv bewertet, während lebensweltliche Mehrsprachigkeit, die die Sprachkenntnisse von Schüler*innen aus ihrem familiären Gebrauch, der häufig mit Migrationserfahrungen verbunden ist, meist als problematisch erachtet wird (Slembrouck et al., 2018).

Für die unterschiedlichen Sprachen im Leben einer mehrsprachigen Person gibt es in der Literatur und im alltäglichen Gebrauch unterschiedliche Bezeichnungen. Als *Erstsprache* (oder L1) wird im Allgemeinen die Sprache bezeichnet, die zuerst erworben wurde, während die *Zweitsprache* (oder L2) im späteren Verlauf des Lebens, z.B. im Rahmen von Migrationssituationen erworben wird (Khan, 2018). Der Begriff der Zweitsprache wird häufig auch dann verwendet, wenn im Kontext des familiären Spracherwerbs bereits mehr als eine Sprache erlernt wurde. Als *Erstsprache* wird zudem meist die Sprache bezeichnet, die im familiären Kontext genutzt wird (Khan, 2018). Hierfür finden sich oft auch die Begriffe *Muttersprache*, *Herkunftssprache* und *Familiensprache*, die unterschiedlich konnotiert sind. Der Begriff der Muttersprache bezeichnet die Sprache, die als erstes erworben wurde und das höchste Kompetenzniveau aufweist (Fürstenau, 2017). Für Sprachbiografien, in denen simultan mehrere Sprachen im familiären Kontext erworben werden, ist diese Definition jedoch problematisch, ebenso für Spracherwerbsprozesse, an denen (auch) andere Personen im familiären und sozialen Umkreis beteiligt sind (Ahrenholz, 2020; Eisenclas & Schalley, 2020). Die Begriffe der Herkunftssprache und Familiensprache beziehen sich auf die „mitgebrachten Sprachen eingewanderter Minderheiten“ (Fürstenau, 2017, S. 32). Fürstenau (2017) führt an, dass die Verwendung des Begriffs Herkunftssprache zum einen dahingehend problematisch sei, da das Herkunftsland bzw. die Region die Sprache nicht eindeutig definieren kann und es zudem eine diskriminierende Wirkung haben kann, wenn Sprecher*innen eine andere Herkunft zugeschrieben wird. In der vorliegenden Arbeit wird daher der Begriff der *Familiensprachen* als Bezeichnung für weitere Sprachen verwendet, die mehrsprachige Schüler*innen im Kontext lebensweltlicher Mehrsprachigkeit mitbringen, und „lebendige Kommunikationssprachen im familiären Kontext“ (Mehlhorn, 2020, S. 24) darstellen.

2.1.2 Ganzheitliche Perspektive auf Mehrsprachigkeit

Insgesamt ist zu verzeichnen, dass sich das Verständnis von Sprachen im Rahmen eines *multilingual turns* (z.B. Conteh & Meier, 2014) von der Vorstellung über getrennte lineare Systeme hin zu einem dynamischen und integrierten Sprachrepertoire gewandelt hat. Traditionelle Auffassungen von Sprachen und Mehrsprachigkeit sind an einer monolingualen Norm orientiert, in denen Sprachen als unabhängige Systeme betrachtet werden, die parallel zueinander bestehen (García, 2009; García & Sylvan, 2011). Diesen traditionellen Auffassungen zufolge wird eine strikte Trennung der Sprachen als förderlich für die Entwicklung der einzelnen Sprachen angesehen, wohingegen die parallele Nutzung der Sprachen als Beeinträchtigung und hinderlich für die Sprachentwicklung gilt (Cummins, 2000; Creese & Blackledge, 2010). Im Sinne der *Time-on-task*-Hypothese wird angenommen, dass die Sprachen bestmöglich gefördert werden, wenn Personen der Zielsprache möglichst lange ausgesetzt sind und die Lernzeit entsprechend hoch ist.

Unterrichtspraktiken orientierten sich folglich lange an einer strikten Trennung der Sprachen und der Methode der Immersion (García, 2009). Mit der *Interdependenzhypothese* (Cummins, 2000), die mit der Annahme einhergeht, dass Sprachen sich gegenseitig beeinflussen und sprachliche Fähigkeiten über Sprachen hinweg übertragen werden können, wurden traditionelle Sichtweisen herausgefordert.

Neuere Modelle von Mehrsprachigkeit betonen einen integrierten Gebrauch von Sprachen sowie eine flexible Sprachentwicklung, die nicht mehr an einer monolingualen Norm ausgerichtet ist. Mit dem Konzept des *Translanguagings*, welches den statischen und homogenen Annahmen von Sprache gegenübersteht, werden natürliche Sprachpraktiken mehrsprachiger Sprecher*innen fokussiert (García, 2009). Translanguaging wird definiert als Prozess „of making meaning, shaping experiences, gaining understanding and knowledge through the use of two languages“ (Baker, 2011, S. 288). Sprachen werden hierbei nicht als getrennte Systeme angenommen, sondern bilden ein gemeinsames Sprachrepertoire aus allen sprachlichen Ressourcen, in dem die Sprachen co-existieren und dynamisch miteinander interagieren. García und Wei (2014, S. 24) definieren Translanguaging weiterhin als „going between linguistic structures, systems and modalities and going beyond them“. Das Ziel ist hierbei stets, das Verständnis zu erleichtern und Kommunikationssituationen so effektiv wie möglich zu gestalten, indem Sprachen miteinander interagieren und Sprachsysteme simultan verwendet werden können. Es wird angenommen, dass Translanguaging Vorteile für das Lernen mit sich bringt, wie die Förderung eines besseren und tieferen Verständnisses der Unterrichtsinhalte, die Entwicklung der schwächeren Sprache und die Integration der familiären und schulischen Lebenswelten (García & Wei, 2014).

Translanguaging kann dabei unterschiedliche diskursive Praktiken, wie z. B. das Code-Switching beinhalten, ist jedoch nicht mit diesem Konzept gleichzusetzen. Code-Switching zeichnet sich dadurch aus, dass von einer Sprache und deren grammatikalischen System zu einer anderen Sprache gewechselt wird (Cenoz & Gorter, 2021). Die Sprachsysteme werden dabei als autonom und unabhängig voneinander betrachtet, während dem Konzept des Translanguagings die Auffassung zugrunde liegt, dass das sprachliche Repertoire einer Person eine Einheit bildet. Die natürliche Sprachpraxis bildet sich dadurch je nach Interaktionspartner*innen und sozialer Situation und zeichnet sich dadurch aus, dass mehrere Sprachsysteme simultan verwendet werden (García, 2009). Nach Grosjean (2020) verfügen mehrsprachige Sprecher*innen über verschiedene Sprachmodi, welche sich auf einem Kontinuum zwischen einem monolingualen und einem mehrsprachigen Modus bewegen. Je nach individueller Kompetenz, Gesprächspartner*innen und Kontext wird in der Interaktion ein passender Sprachmodus gewählt, der unterschiedliche Abstufungen und Kombinationen annehmen kann. Sprachliche Ressourcen, die im jeweiligen Modus nicht verwendet werden, sind in diesem Moment nach Grosjean (2020) deaktiviert. Dies erfolgt jedoch selten vollständig und „gemischtsprachiges Sprechen“ (Grosjean, 2020, S. 16) kann sich ergeben.

2.1.3 Einstellungen zu Mehrsprachigkeit

Der beschriebene Paradigmenwechsel hin zu einer ressourcenorientierten Sichtweise auf Mehrsprachigkeit spiegelt sich im Umgang mit dieser im Unterricht noch nicht wider (Kropp, 2017). Untersuchungen aus dem deutschsprachigen sowie internationalen Raum zeigen häufig, dass Lehrkräfte sprachliche und kulturelle Vielfalt durchaus als Chance und Gewinn wahrnehmen, Unterrichtspraktiken jedoch trotzdem mehrheitlich monolingual an einer Verkehrssprache orientiert sind (z.B. Lui et al., 2021; Pohlmann-Rother & Lange, 2020; Pulinx et al., 2017; Schnitzer, 2020). In einer qualitativen Interviewstudie mit deutschsprachigen Lehrkräften konstatiert Schnitzer (2020, S. 243) eine „Ambivalenz“ der Einstellungen zu Mehrsprachigkeit. Viele Lehrkräfte in der Untersuchung bewerten Mehrsprachigkeit im Allgemeinen positiv und vorteilhaft für die berufliche Laufbahn, ihr Umgang mit Mehrsprachigkeit im schulischen Kontext ist gleichzeitig jedoch von einer defizitorientierten Haltung geprägt. Lehrkräfte erkennen zwar, dass Mehrsprachigkeit als Ressource für das Lernen genutzt werden kann, konkrete Unterrichtskonzepte sind allerdings weitgehend unbekannt und werden daher im Unterricht nicht implementiert (Bredthauer & Engfer, 2018). Die Angst vor Autoritätsverlust der Lehrkräfte, wenn sie mehrsprachige Aussagen nicht verstehen oder überprüfen können, und mögliche Ausgrenzung monolingualer Schüler*innen sowie Zeitmangel werden dabei häufig als Gründe für eine monolinguale Orientierung des Unterrichts berichtet (Bredthauer & Engfer, 2018; Fürstenau, 2016; Ticheloven et al., 2021). Wenn Schüler*innen sich in anderen Sprachen verständigen, wird dies zudem als Störfaktor wahrgenommen (Camenzuli et al., 2022; Schnitzer, 2020). Die vielen verschiedenen Sprachen, die in einer Klasse zusammenkommen, stellen dabei eine besondere Herausforderung für Lehrkräfte dar und auch der Aufwand in der Unterrichtsvorbereitung wird als Hürde für die Implementation innovativer sprachlicher Konzepte wahrgenommen (Bredthauer & Engfer, 2018).

Weiterhin ist zu beobachten, dass negative Einstellungen mit bestehenden Mythen über Mehrsprachigkeit und einer entsprechenden Defizitorientierung einhergehen. Der Gebrauch von anderen Sprachen im Unterricht wird insbesondere für Schüler*innen mit weniger gut ausgeprägten Deutschkenntnissen als Gefahr für den Deutscherwerb angesehen, wobei kommunikative Funktionen und lernförderliche Potenziale nur für Schüler*innen mit ausgeprägten Kompetenzen in der Familiensprache angenommen werden (Bredthauer & Engfer, 2018). Die Argumentation über die *Time-on-task*-Hypothese, nach der möglichst viel Zeit für das Lernen einer Sprache eingeräumt werden sollte, zeigt sich auch in einer ablehnenden Haltung gegenüber der Förderung von herkunftssprachlichen Kompetenzen im Unterricht, durch die weniger Zeit für die Förderung der Deutschkenntnisse zur Verfügung steht (Schnitzer, 2020). Der Ausschluss von Familiensprachen wird in dem Sinne über einen vermeintlich höheren Bildungserfolg legitimiert (Mecheril & Quehl, 2006). Entsprechend wird der Regelunterricht als monolingualer Raum erachtet und Lehrkräfte sehen Möglichkeiten für den Einbezug von Mehrsprachig-

keit höchstens in Ergänzungsangeboten (z. B. in herkunftssprachlichem Unterricht) (Bredthauer & Engfer, 2018). Im Sinne einer neo-linguistischen Argumentationslinie (Dirim, 2010) wird Mehrsprachigkeit hierbei nur unter bestimmten Voraussetzungen akzeptiert, z. B. wenn Schüler*innen die deutsche Sprache gut beherrschen und ihre Mehrsprachigkeit nicht mehr sichtbar ist oder ihre Kenntnisse in allen ihren Sprachen gut ausgebildet sind (Bredthauer & Engfer, 2018; Schnitzer, 2020). Offenheit gegenüber mehrsprachigkeitsdidaktischen Methoden signalisieren nach Bredthauer und Engfer (2018) insbesondere Lehrkräfte sprachlicher Fächer. Auch hier scheitert die Umsetzung jedoch daran, dass Lehrkräften Aus- und Weiterbildungsmöglichkeiten sowie konkretes Unterrichtsmaterial fehlen.

Insgesamt ist die Umsetzung mehrsprachigkeitssensibler Angebote im Unterricht maßgeblich von den Einstellungen der Lehrkräfte abhängig (Mehlhorn, 2020). Dennoch sind auch positive Einstellungen keine Garantie dafür, dass Mehrsprachigkeit produktiv in den Unterricht integriert wird und der Zusammenhang zwischen Einstellungen und Verhalten ist komplex. Pohlmann-Rother und Lange (2020) zeigen, dass die befragten Grundschullehrkräfte wenig bis keine negativen Folgen für das Klassenklima befürchten, wenn sie andere Sprachen in den Unterricht integrieren. Eine entsprechende Umsetzung erfolgt häufig jedoch nicht. Daher ist anzunehmen, dass weitere Faktoren die Implementation und das Gelingen mehrsprachiger Angebote behindern können, selbst wenn Lehrkräfte positive Einstellungen aufweisen.

Die Einstellungen der Lehrkräfte tragen nicht nur dazu bei, ob mehrsprachige Ressourcen im Unterricht produktiv genutzt werden, sondern können sich ebenfalls auf die Einstellungen und Leistungen der Schüler*innen auswirken (Hu, 2003). Pulinx et al. (2017) zeigen in einer Untersuchung mit niederländischen Lehrkräften, dass eine stärkere monolinguale Orientierung der Lehrkräfte mit weniger Zutrauen in die Kompetenzen mehrsprachiger Schüler*innen einhergeht. Insgesamt finden sich ebenso bei Schüler*innen kontrastierende Einstellungen zu Mehrsprachigkeit. Verschiedene Untersuchungen kommen zu dem Ergebnis, dass mehrsprachige Schüler*innen ihren Familiensprachen grundsätzlich positiv gegenüberstehen und sich mehr Wertschätzung ihrer sprachlichen Ressourcen im unterrichtlichen Kontext wünschen (Binander & Jessen, 2020; Brehmer & Mehlhorn, 2018; Wamhoff et al., 2022). Es ist jedoch ebenso zu beobachten, dass negative Einstellungen zu Mehrsprachigkeit und sprachliche Hierarchisierungen der Lehrkräfte und Gesellschaft von Schüler*innen übernommen werden (Brehmer & Mehlhorn, 2018; Wamhoff et al., 2022). Auch die Legitimierung von Sprachverböten über einen vermeintlich höheren Bildungserfolg durch Monolingualität werden von Schüler*innen verinnerlicht und selbst argumentiert (Wamhoff et al., 2022). Insgesamt ist zu berücksichtigen, dass Einstellungen zur eigenen Mehrsprachigkeit bei Schüler*innen ganz individuell ausgeprägt sein können. Ritterfeld et al. (2014) zeigen, dass bei Schüler*innen auch unabhängig von selbstberichteten Sprachkompetenzen ein differentieller Umgang mit der eigenen Mehrsprachigkeit zu beobachten ist. In einer Untersuchung mit 765 mehrsprachigen Schüler*innen

zwischen 10 und 18 Jahren identifizieren die Autor*innen vier Cluster, die sich zwischen einem monolingualen und multilingualen Typus unterscheiden sowie unterschiedlich erlebten Sprach- und Kommunikationsbarrieren. Somit können sich Schüler*innen als kompetent in verschiedenen Sprachen erleben und gerne mehrsprachig interagieren, mit gleichen sprachlichen Kompetenzen jedoch auch Einsprachigkeit in der Verkehrssprache bevorzugen.

2.1.4 Mehrsprachigkeit als Ressource für das Lernen

Viele Studien bestätigen kognitive Vorteile mehrsprachiger Sprecher*innen. Eine Metaanalyse von Adesope et al. (2010) aus dem amerikanischen Raum zeigt, dass bilinguale Sprecher*innen über ein höheres Maß an metasprachlichem und meta-kognitivem Bewusstsein verfügen und eine bessere Steuerung der Aufmerksamkeit aufweisen können. Auch in europäischen Ländern zeigen Studien positive Effekte bilingualer Unterrichtsprogramme für die Lernwirksamkeit im Lesen (Reljić et al., 2015). Kompetenzen in den Familiensprachen werden für den schulischen Erfolg von Schüler*innen mit Migrationshintergrund jedoch unterschiedliche Bedeutungen beigemessen. Verschiedene theoretische Annahmen sowie widersprüchliche und methodisch eingeschränkte empirische Ergebnisse (z.B. Schüler-Meyer et al., 2019; Reljić et al., 2015) tragen zu einer fortwährenden kontroversen Diskussion über den Nutzen mehrsprachigkeitssensibler Angebote bei (z.B. Esser, 2009). Ausbleibende positive Effekte mehrsprachiger Lernprogramme werden in diesem Kontext bspw. über die *Time-on-task*-Hypothese (z.B. Scheele et al., 2010) oder durch kognitive Kosten aufgrund von kognitiver Überforderung argumentiert (z.B. McElvany et al., 2017), wenn Schüler*innen zusätzlich ihre Familiensprachen beim Lernen verwenden. Unter der Annahme der Interdependenzhypothese, dass Sprachen sich gegenseitig beeinflussen und sich sprachliche Fähigkeiten übertragen lassen (Cummins, 2000), können jedoch positive Effekte durch Translanguaging angenommen werden, da die Kompetenzen unterschiedlicher Sprachen ein gemeinsames Repertoire bilden (García & Lin, 2017).

Die häufige Beschreibung von Mehrsprachigkeit als *Ressource* legt den Fokus auf die Rolle der Mehrsprachigkeit als „Hilfsmittel zur Bewältigung einer besonderen Herausforderung oder schwierigen Situation“ (Kropp, 2017, S. 115). Eine Öffnung für Familiensprachen im Unterricht kann Kommunikationsmöglichkeiten für Schüler*innen erweitern und neue Lerngelegenheiten schaffen (García, 2009; Grosjean, 2013). Die Anerkennung individueller sprachlicher Ressourcen als Potenziale kann dabei zu mehr Bildungsgerechtigkeit beitragen (vgl. Honneth, 1992) sowie die Identitätsentwicklung und das Selbstwertgefühl mehrsprachiger Schüler*innen unterstützen (Krumm, 2016). Aus kognitiver Perspektive können der Nutzung aller sprachlichen Ressourcen kommunikative und lernförderliche didaktische Funktionen zugesprochen werden, da mit der Erweiterung des sprachlichen Repertoires auch die Erweiterung von Ausdrucks- und Kommunikations-

möglichkeiten einhergeht (García, 2009; Grosjean, 2013; Prediger et al., 2021). Sprachliche Barrieren und kommunikative Herausforderungen können durch kurze Übersetzungen oder Verhandlung fachlicher Inhalte abgebaut werden, sodass ein besseres Verständnis der Lerninhalte möglich wird (Moschkovic, 2007; Schüler-Meyer et al., 2019; Planas & Setati, 2009). Die Nutzung des gesamten sprachlichen Repertoires durch mehrsprachige Äußerungen kann dann als Hilfsmittel fungieren, um möglichst effektiv zu kommunizieren und an Vorwissen sowie an individuelle Lernvoraussetzungen anzuknüpfen (Jakisch, 2015). Zudem werden mehrsprachige Interaktionen häufig auch zur Organisation des Arbeitsprozesses genutzt und können den Lernprozess auf diese Weise indirekt unterstützen (Bühlig & Duarte, 2013; Schastak, 2020). Der Einbezug und die Förderung sprachlicher sowie kultureller Ressourcen im Bildungskontext wird daher aus politischer (Council of Europe, 2001) und wissenschaftlicher Perspektive gefordert (z. B. Fürstenau, 2017; García, 2009). Mehrsprachige Schüler*innen haben im Unterricht jedoch selten die Möglichkeit, ihre Familiensprachen einzubringen und diese als Ressource zu erfahren (Schastak et al., 2017).

2.2 Mehrsprachigkeitssensibles Unterrichten

Im folgenden Kapitel werden Möglichkeiten zum mehrsprachigkeitssensiblen Unterrichten in sprachlich heterogenen Klassenräumen beschrieben. Hierbei werden theoretische Ansätze dargestellt, aus denen sich praktische Impulse zum Einbezug von Familiensprachen ableiten lassen. Ein zentraler Baustein sind dabei Strategien des pädagogischen Translanguagings. Zudem werden Forschungsbefunde zur Wirksamkeit mehrsprachigkeitssensiblen Unterrichtens dargestellt, wobei Ergebnisse zur Wahrnehmung der Konzepte und Umsetzung durch die Schüler*innen fokussiert werden.

2.2.1 Theoretische Ansätze und praktische Impulse für einen mehrsprachigkeitssensiblen Unterricht

Für die Berücksichtigung, Wertschätzung und produktive Nutzung sprachlicher Heterogenität im Unterricht sind in der Literatur unterschiedliche Konzepte zu finden, die verschiedene Zielsetzungen und Bezüge aufweisen. Hierbei ist zunächst zwischen Ansätzen und Strategien zu unterscheiden, die auf unterrichtlicher Ebene und fachspezifisch angesiedelt sind (z. B. Prediger et al., 2021; Schnitzer, 2020) oder fachübergreifend im Sinne eines Gesamtsprachencurriculums umgesetzt werden (z. B. Reich & Krumm, 2013). Zudem sind auch institutionsübergreifende Konzepte einer durchgängigen Sprachbildung zu finden (Gogolin & Lange, 2011). In der vorliegenden Arbeit werden pädagogische Konzepte und Strategien fokussiert, die die Integration von Mehrsprachigkeit auf unterrichtlicher Ebene für das

fachliche Lernen thematisieren und konkrete Impulse für die Unterrichtspraxis bieten. Hierbei steht der Einbezug lebensweltlicher Mehrsprachigkeit im Fokus.

Konzepte, die den Einbezug aller sprachlichen Ressourcen der Schüler*innen im Unterricht ermöglichen, haben meist die Ausbildung von Sprachsensibilität, Sprachbewusstheit und Sprachaufmerksamkeit zum Ziel. Dies umfasst didaktische Ansätze aus unterschiedlichen Bereichen, wie der Mehrsprachigkeitsdidaktik und Fremdsprachenförderung (z. B. Hilbe et al., 2017b; Kutzelmann et al., 2017) sowie Strategien, die an pädagogisches Translanguaging angelehnt sind (z. B. Cenoz & Gorter, 2021; Gantefort, 2020; Lucas et al., 2008; Sierens & van Avermaet, 2014). Diese können aufbauend auf dem Konzept des Translanguagings (vgl. Kapitel 2.1.2) von Lehrkräften im Unterricht eingesetzt werden, um unterschiedliche Familiensprachen von Schüler*innen im Lernprozess zu integrieren und zu aktivieren. Dabei soll die natürliche Sprachpraxis mehrsprachiger Sprecher*innen aufgegriffen und das gesamte sprachliche Repertoire eingesetzt werden (vgl. García & Wei, 2014). Hierbei liegt ein ganzheitliches Verständnis von Mehrsprachigkeit zugrunde, in dem verschiedene Sprachen keine getrennten Einheiten darstellen, sondern ein Gesamtrepertoire bilden, das für alle sprachlichen Handlungen zur Verfügung steht (Gantefort, 2020; vgl. Kapitel 2.1.2). Es wird angenommen, dass Praktiken des Translanguagings im Unterricht ein tieferes Verständnis der Lerninhalte anregen und die Entwicklung der schwächeren Sprache(n) durch ihren Einbezug fördern können (García & Wei, 2014). Auch für monolinguale Schüler*innen können sich positive Effekte ergeben. Sprachverbote beschränken und stigmatisieren dagegen den natürlichen Sprachgebrauch (Sierens & van Avermaet, 2014). Insgesamt sollen beim pädagogischen Translanguaging die sprachlichen Ressourcen der Schüler*innen als fester Faktor im Lernprozess etabliert werden, ohne getrennte Lerngelegenheiten schaffen zu müssen, bspw. im Rahmen herkunftssprachlichen Unterrichts oder bilingualer Programme (Slembrouck et al., 2018).

Viele didaktische Ansätze für mehrsprachigkeitssensibles Unterrichten basieren auf Prinzipien pädagogischen Translanguagings, die sich jedoch in der Bezeichnung und konkreten Ausgestaltung unterscheiden. Es können übergreifende Kernelemente identifiziert werden, die in vielen Konzepten zum Tragen kommen und Leitlinien für den Umgang mit Mehrsprachigkeit in einem mehrsprachigkeitssensiblen Unterricht darstellen können (vgl. Lorenz et al., 2021). Grundlegend ist hierbei zunächst eine ganzheitliche Sichtweise auf Mehrsprachigkeit entgegen einer monolingualen Norm. Darauf aufbauende Ziele umfassen die Wertschätzung aller sprachlichen Ressourcen, die Förderung positiver Einstellungen zu sprachlicher Heterogenität, die Entwicklung metasprachlichen Bewusstseins sowie die Nutzung des gesamten sprachlichen Repertoires für das Lernen (Lorenz et al., 2021). Aus diesen Kernelementen lassen sich praktische Impulse für eine mehrsprachigkeitssensible Unterrichtspraxis ableiten, die im Folgenden beschrieben werden.

Eine wichtige Grundlage zur Aktivierung mehrsprachiger Ressourcen stellt die Etablierung eines wertschätzenden Klimas für sprachliche Vielfalt dar. Im deutschsprachigen Raum ist ein monolingualer Habitus (Gogolin, 1994) im Schulkontext

vorherrschend, der die Aktivierung von Mehrsprachigkeit erschwert und negative Einstellungen zu Familiensprachen begünstigt. Der Sensibilisierung für sprachliche Vielfalt kommt dabei eine besonders wichtige und grundlegende Funktion zu, um dem monolingualen Habitus entgegenzuwirken und die Schüler*innen verstärkt zur Nutzung ihrer Familiensprachen einzuladen. Es wird angenommen, dass ein wertschätzendes Klima insbesondere sozio-emotionale Vorteile mit sich bringt, dadurch dass die Identitätsentwicklung durch die Anerkennung der individuellen Ressourcen gestärkt und die Verständigung zwischen Schüler*innen gefördert wird (Sierens & van Avermaet, 2014). Ein wertschätzendes Klima soll zudem bewirken, dass sich Schüler*innen mit der Nutzung ihrer Sprachen wohler fühlen und keine Angst vor Ausgrenzung oder Stigmatisierung haben (Lucas et al., 2008). Für den Grundschulunterricht gibt es unterschiedliche Möglichkeiten, die sprachliche Vielfalt in den Klassen zu thematisieren und für alle sichtbar zu machen und so zu einem wertschätzenden Klima beizutragen. Die Anfertigung von Sprachenportraits (Gogolin & Neumann, 1991), bei denen die Schüler*innen ihre Sprachen mit unterschiedlichen Farben in eine menschliche Silhouette einzeichnen, trägt zur eigenen Reflexion der sprachlichen Biografie sowie zur Wertschätzung dieser in den Klassen bei und unterstützt die identitätsstiftende Funktion von Sprache (Budde, 2012). Die Arbeit mit Sprachenportraits ermöglicht Schüler*innen und Lehrkräften die emotionale Bedeutung der Sprachen im Klassenzimmer zu erforschen, ohne ein bestimmtes Maß an Kompetenzen in den einzelnen Sprachen vorauszusetzen (Schnitzer, 2020). Ein weiterer kreativer Zugang, um Mehrsprachigkeit sichtbar zu machen und die eigene Mehrsprachigkeit zu reflektieren, ist die Anfertigung individueller Sprachenbäume (Wildemann & Fornol, 2016). Zudem können Sprachen sichtbar in die Lernumwelt integriert werden, indem Bücher in den Familiensprachen der Schüler*innen mitgebracht oder mehrsprachige Beschriftungen im Klassenraum für Objekte etabliert werden, bei denen Schüler*innen ihre sprachliche Expertise einbringen können (Wildemann & Fornol, 2016).

Ein weiteres Kernelement pädagogischen Translanguagings stellt die Aktivierung und Nutzung des sprachlichen Repertoires zur Unterstützung von Lernprozessen dar. Fachliches Wissen, das in einer Familiensprache womöglich bereits erworben wurde, soll durch die Nutzung des gesamten sprachlichen Repertoires zugänglich gemacht werden. Auf diese Weise können auch sprachübergreifende Lernstrategien aufgebaut werden (Lengyel, 2017). Dies kann im Unterricht durch unterschiedliche Aktivitäten erzielt werden. Es wird angenommen, dass das Unterrichtsmaterial in den Familiensprachen als *Scaffold* (vgl. Kapitel 2.3.3) und Werkzeug zur Unterstützung genutzt werden kann, um an das Vorwissen anzuknüpfen und den Aufbau neuen Wissens im Lernprozess zu erleichtern (Cenoz & Gorter, 2021). Dies kann bspw. Texte oder Arbeitsanweisungen in unterschiedlichen Sprachen sowie die Integration von Glossaren umfassen (Lucas et al., 2008). Neben solchen mehrsprachigen Unterrichtsmaterialien können darüber hinaus auch weitere Ressourcen, wie z. B. Übersetzungen aus dem Internet oder Erklärvideos in unterschiedlichen Sprachen eingebunden werden, die besonders niedrigschwellig

von Schüler*innen selbst recherchiert und genutzt werden können (Prediger et al., 2021). Dabei ist auch die Nutzung digitaler Endgeräte als Werkzeuge denkbar, wie z. B. Sprachlernstifte oder Apps zur Übersetzung, um das Verständnis zu fördern und bildungssprachliche Fähigkeiten in den Familiensprachen auszubauen (Slembrouck et al., 2018).

Neben der Aktivierung des sprachlichen Repertoires durch die Integration von Unterrichtsmaterial sollen weiterhin auch Möglichkeiten zur mehrsprachigen Interaktion geschaffen werden, mit der die Schüler*innen verstärkt die Verantwortung zur Gestaltung ihrer Lernprozesse übernehmen (Sierens & van Avermaet, 2014; Slembrouck et al., 2018). Eine besondere Herausforderung für die Aktivierung mehrsprachiger Interaktion stellt im deutschsprachigen Kontext die große Vielzahl an unterschiedlichen Familiensprachen in einer Klasse dar, die Lehrkräfte selbst nicht beherrschen (vgl. Henschel et al., 2019). Hierbei kann die Zuordnung von Sprachpartnerschaften (Lucas et al., 2008) und Nutzung von geteilten Familiensprachen in Kleingruppen (z. B. Meyer & Prediger, 2011; Prediger et al., 2021) ein möglicher Impuls zur Aktivierung mehrsprachiger Interaktion sein. Auf diese Weise können Schüler*innen mit geteilter Familiensprache ihr gemeinsames sprachliches Repertoire einsetzen, um Verständigung und Kommunikation zu erleichtern, sich gegenseitig bei der Bearbeitung der Aufgaben zu unterstützen und Arbeitsprozesse effektiv zu organisieren (Bühlig & Duarte, 2013; Gantefort, 2020; Prediger et al., 2021). Es wird angenommen, dass das sprachliche Anforderungsniveau auf diese Art zeitweise entlastet werden kann, was die Aneignung von Wissen erleichtert. Hierbei soll die kommunikative Funktion von Mehrsprachigkeit in sprachheterogenen Klassen bestmöglich ausgeschöpft werden. Da Lehrkräfte jedoch häufig Kontroll- und Autoritätsverlust befürchten, wenn sie andere Sprachen im Lernprozess zulassen, die sie selbst nicht verstehen, kann mehrsprachige Kommunikation zunächst in geplante Phasen verlegt werden, was Lehrkräften Struktur und Sicherheit für den Unterrichtsverlauf verschaffen kann (Bredthauer & Engfer, 2018).

2.2.2 Forschungsergebnisse zu mehrsprachigkeitssensiblen Unterrichtskonzepten

Um Bedingungen erfolgreicher Umsetzung zu identifizieren und Konzepte effektiv weiterentwickeln zu können, werden wissenschaftliche Evaluationen gefordert (Bredthauer & Engfer, 2018; Lengyel, 2017). Zur Umsetzung und Wirksamkeit mehrsprachigkeitssensibler Unterrichtskonzepte gibt es bislang jedoch nur wenig Forschung und die Befunde sind gemischt.

Positive Effekte mehrsprachigkeitssensiblen Unterrichtens konnten bisher für die Entwicklung metalinguistischen Bewusstseins (Leonet et al., 2020), die Erweiterung des sprachlichen Repertoires (Bosma et al., 2022) und die Förderung des Selbstvertrauens in die Nutzung sprachlicher Ressourcen (Cenoz et al., 2022)

nachgewiesen werden. In Interventionsstudien zur Leseförderung (Decristan et al., 2022) und Mathematikkompetenz (Prediger & Neugebauer, 2023) zeigen sich bisher jedoch keine Vorteile des mehrsprachigkeitssensiblen Unterrichts für mehrsprachige Schüler*innen, da auch monolinguale Schüler*innen gleichermaßen von den Unterrichtskonzepten profitieren.

In einer Studie von Schüler-Meyer und Kolleg*innen (2019) zeigt sich, dass der Anteil an gemischtsprachlichen türkisch-deutschen Äußerungen den größten Zusammenhang mit dem Lernzuwachs in einem fünfwöchigen bilingualen Mathematikurs für siebte Klassen aufweist. Aus bisherigen Forschungsergebnissen zur Integration von Mehrsprachigkeit in den Regelunterricht geht jedoch auch hervor, dass mehrsprachige Kommunikation in diesem Kontext schwierig zu aktivieren ist (vgl. Meyer & Prediger, 2011). Hemmende Faktoren können im monolingualen Habitus, in Sprachasymmetrien zugunsten der in der Schule vorherrschenden Sprache (Halai, 2009; Reich, 2009) sowie in negativen Einstellungen auf Seiten der Lehrkräfte und Schüler*innen gegenüber der Nutzung von Familiensprachen für das Lernen gesehen werden (Prediger et al., 2021; Sierens & van Avermaet, 2014). Zudem müssen auch erschwerende Kontextbedingungen beachtet werden, wie das Vorhandensein vieler unterschiedlicher Sprachen in den Klassen. Erste Studien geben Hinweise darauf, dass Sprachpartnerschaften in kooperativen Settings die Aktivierung von Mehrsprachigkeit unterstützen kann (Bührig & Duarte, 2013; Decristan et al., 2021; Meyer & Prediger, 2011; Prediger et al., 2021).

Reitenbach et al. (2023) untersuchen Gründe für die Nutzung und Nicht-Nutzung von Familiensprachen während einer mehrsprachigkeitssensiblen Unterrichtseinheit zum Reziproken Lehren im Deutschunterricht in vierten Klassen.¹ Hierbei geben 63,2% der mehrsprachigen Schüler*innen an, während der Intervention neben dem Deutschen noch eine weitere Sprache gesprochen zu haben und begründen ihre Antwort in einem kurzen und offenen Antwortformat. Die häufigste Begründung für die Nutzung der Familiensprachen während der Unterrichtseinheit bezieht sich auf die individuelle Kompetenz sowie auf das Erleben positiver Emotionen. Weiterhin wird die Nutzung auch über die Kompetenzen und Emotionen auf Ebene der Gruppe und Interaktionspartner*innen argumentiert sowie über kontextuelle Faktoren, wie bspw. die Aufforderung zur Nutzung durch die Lehrkraft. Auch die Ablehnung des Angebots zur Sprachnutzung wird durch die genannten Hauptkategorien von den Schüler*innen begründet, bspw. mit mangelnden Kompetenzen oder einem Sprachverbot der Lehrkraft. Die Ergebnisse stellen die Individualität bei der Nutzung von Familiensprachen im Unterricht heraus sowie wichtige kontextuelle Merkmale, wie die Gruppenzusammensetzung und ein wertschätzendes und aktivierendes Klima zur Sprachnutzung.

Zur Implementation und Wahrnehmung mehrsprachigkeitssensibler Unterrichtskonzepte durch die Schüler*innen gibt es bislang nur wenig Forschung,

1 Die vorliegende Untersuchung basiert auf den gleichen Daten wie die Untersuchung von Reitenbach et al. (2023).

zumeist qualitativ orientierte Studien mit kleinen Stichproben. Schnitzer (2020) untersucht die Wahrnehmung einer mehrsprachigkeitssensiblen Unterrichtsreihe zum *Sprachenfächer* in fünf Klassen der Sekundarstufe 2. Der Sprachenfächer umfasst Unterrichtsmaterial für einen mehrsprachigen Deutschunterricht, bei dem die Förderung von Sprachaufmerksamkeit und Sprachwissen durch lebensweltnahe Themen im Vordergrund steht (Schnitzer, 2020). Die Autorin resümiert auf Basis von qualitativen Beobachtungs- und Befragungsdatendaten, dass sowohl mehrsprachige als auch monolinguale Schüler*innen großes Interesse am Unterricht mit dem Sprachenfächer zeigen und sich beide Schüler*innengruppen im Vergleich zum regulären Unterricht außerordentlich motiviert am Unterricht beteiligen. Hilbe et al. (2017b) implementieren die Methode des mehrsprachigen Vorlesens im Fremdsprachenunterricht zur Sprachsensibilisierung. Auf Basis einer Befragung von 188 Schüler*innen zeigt sich, dass die Akzeptanz für die Maßnahme insgesamt hoch ausfällt. Auf affektiver und motivationaler Ebene wird das mehrsprachige Vorlesen dabei positiver bewertet als der reguläre Unterricht. Die Nützlichkeit auf kognitiver Ebene wird vergleichbar mit dem regulären Fremdsprachenunterricht angegeben, wobei Schüler*innen einen geringeren kognitiven Aufwand beim mehrsprachigen Vorlesen berichten (Hilbe et al., 2017a). Weitere positive Ergebnisse zeigen sich im Forschungsprojekt „Mehrsprachiges Lesetheater zur Förderung der Leseflüssigkeit und Lesemotivation“ (Kutzelmann et al., 2017), bei dem Familiensprachen von Schüler*innen mit Migrationshintergrund in eine Lesefördermaßnahme für fünfte bis achte Klassen integriert werden. Das methodisch-didaktische Konzept wird auf Basis von sechs Interviews mit teilnehmenden Lehrkräften als sehr positiv und im Unterricht gut implementierbar bewertet. Das Projektteam stellt hierbei fest, dass sich die Einstellungen der Lehrkräfte zur Integration von Mehrsprachigkeit im Unterricht im Verlauf des Projekts verbessern und sich anfängliche Bedenken zur Integration von Sprachen, die sie selbst nicht verstehen, auflösen (Götz et al., 2017).

Ticheloven et al. (2021) befragen in einer Interviewstudie 31 Schüler*innen, die mit mehrsprachigkeitssensiblen Konzepten unterrichtet werden. Wichtige Herausforderungen, die Schüler*innen im mehrsprachigkeitssensiblen Unterricht wahrnehmen, beziehen sich auf zusätzlichen kognitiven Aufwand beim Lernen, wenn die Familiensprache nicht gut beherrscht wird. Aber auch Schüler*innen, die ihre Familiensprachen gut beherrschen, geben in diesem Kontext an, dass sie häufiger die Nutzung der Schul- und Verkehrssprache für die Bearbeitung der Unterrichtsinhalte wählen, da sie den Umgang mit dieser im Bildungskontext gewohnt sind (Ticheloven et al., 2021). Mehrsprachigkeitssensible Unterrichtsstrategien werden stattdessen als Möglichkeit gesehen, die Familiensprache zu erhalten (Ticheloven et al., 2021). Auch Ebe und Chapman-Santiago (2016) stellen in einer Beobachtungsstudie heraus, dass Schüler*innen einer achten Klasse im amerikanischen Raum, die im Englischunterricht ihre Familiensprachen in Sprachpartnerschaften nutzen können, dies als positiv bewerten und sich in ihrer Identität bestärkt fühlen.

Obwohl das Forschungsfeld noch dünn besiedelt ist, zeigen die Ergebnisse insgesamt, dass die Integration von Familiensprachen im Rahmen mehrsprachigkeitssensibler Konzepte gelingen und positiv wahrgenommen werden kann. Zu berücksichtigen sind hierbei jedoch Einschränkungen, die sich auf die individuelle Sprachkompetenz und auf den eigenen präferierten Umgang mit Mehrsprachigkeit beziehen. Zudem ist festzuhalten, dass mehrsprachigkeitssensible Konzepte nicht nur für mehrsprachige, sondern auch für monolinguale Schüler*innen mit einer positiven Lernerfahrung einhergehen können. Ein Forschungsdesiderat stellt die quantitative Untersuchung von Implementationsprozessen mehrsprachigkeitssensibler Konzepte an größeren Stichproben dar.

2.3 Förderung der Lesekompetenz durch Reziprokes Lehren

Schwache Leseleistungen stellen einen Risikofaktor für schulischen Erfolg in fast allen Fächern dar (Müller, 2015). Unterschiede in den Leseleistungen sind meist bereits früh in der schulischen Laufbahn zu beobachten und bilden sich im Bildungsverlauf weiter aus (Müller, 2015). Internationale Schulleistungsstudien zeigen, dass ein großer Teil von Schüler*innen in vielen europäischen Ländern schon am Ende der Grundschule unterdurchschnittliche Leseleistungen aufweist (Bremerich-Vos et al., 2017). Verschiedene Merkmale konnten als Risikofaktoren identifiziert werden, die mit schwachen Leseleistungen einhergehen. Besonders gefährdet sind Schüler*innen aus bildungsfernen Elternhäusern sowie mehrsprachige Schüler*innen mit Migrationshintergrund oder einer anderen Erstsprache als Deutsch (Stanat & Schneider, 2004; Wendt & Schwippert, 2017). Daher wird in der vorliegenden Arbeit das Konzept des mehrsprachigkeitssensiblen Reziproken Lehrens untersucht, welches die Leseförderung mit der Integration von Mehrsprachigkeit verknüpft.

In den vorangehenden Kapiteln wurde in den Themenkomplex der Mehrsprachigkeit im schulischen Kontext eingeleitet und Strategien für mehrsprachigkeitssensibles Unterrichten dargestellt. Im folgenden Kapitel soll nun das Reziproke Lehren beschrieben werden, um die beiden Themenbereiche miteinander zu verbinden. Das Reziproke Lehren ist eine kooperative Methode, die definiert wird als „instructional procedure in which small groups of students learn to improve their reading through ‘scaffolded instruction’ of comprehension-fostering and comprehension-monitoring strategies” (Hacker & Tenent, 2002, S. 699). Im Kern der Methode steht das Einüben von Lesestrategien im kooperativen Setting. Im folgenden Kapitel werden einleitend Grundlagen zur Lesekompetenz beschrieben und danach die zentralen Bausteine des Reziproken Lehrens eingeführt: das kooperative Lernen und die Instruktion von Lesestrategien.

2.3.1 Facetten von Lesekompetenz

Lesen ist ein komplexer und interaktiver Prozess, der verschiedene kognitive, motivationale und soziale Faktoren umfasst (Alexander & The Disciplined Reading and Learning Research Laboratory, 2012; McElvany et al., 2008; Munser-Kiefer, 2014). Allgemein wird Lesekompetenz als Fähigkeit verstanden, „verschiedene Arten von Texten [...] zu verstehen und zu nutzen“ (Goy et al., 2017, S. 143). Hierbei sind unterschiedliche kognitive Prozesse involviert (Lenhard, 2019). Hierarchieniedrigere Prozesse des Erkennens und Dekodierens von Buchstaben auf Wortebene werden dabei als Lesefertigkeit beschrieben, die eine grundlegende Voraussetzung für erfolgreiche hierarchiehöhere Prozesse des Leseverstehens auf Satz- und Textebene bilden (Munser-Kiefer, 2014). Beim Leseverstehen greifen unterschiedliche Teilkompetenzen ineinander, wie die „semantische und syntaktische Integration von Sätzen durch Inferenzen, die Anwendung von Vorwissen sowie das Verknüpfen von Informationen aus unterschiedlichen Teilen eines Textes“ (Müller, 2015, S. 4). Leseverständnis wird daher auch als komplexe Fähigkeit und Prozess verstanden, „Textinhalt zu rekonstruieren“ (Lenhard, 2019, S. 47). Weiterhin zeichnen sich kompetente Leser*innen dadurch aus, dass sie (meta)kognitive Lesestrategien anwenden, um ihren Lese- und Verständnisprozess zu überwachen und zu optimieren (Artelt et al., 2002; Donker et al., 2014; Munser-Kiefer, 2014; vgl. Kapitel 2.3.3).

Neben diesen kognitiven Prozessen werden zum kompetenten Lesen auch motivationale Merkmale, wie die Lesemotivation und Selbstwirksamkeitserwartungen, gezählt (Guthrie et al., 2004; Koch & Spörer, 2016). Die Lesemotivation spiegelt die emotional-affektive Bedeutung des Lesens für die Schüler*innen wider (Artelt et al., 2010). Die Lesemotivation beeinflusst das Leseverhalten dahingehend, dass motivierte Schüler*innen häufiger und mehr lesen, was sich wiederum positiv auf das Leseverstehen auswirkt (Goy et al., 2017; McElvany et al., 2008; Unrau et al., 2018). Motivation wird als intentionaler Prozess beschrieben, der auf ein Ziel ausgerichtet ist und mit einer Handlungsabsicht einhergeht (Wigfield & Eccles, 2000; Hasselhorn & Gold, 2017). Im Kontext des Lernens ist Motivation daher zu verstehen als die „Bereitschaft [...], sich aktiv, dauerhaft und wirkungsvoll mit bestimmten Themengebieten auseinanderzusetzen, um neues Wissen zu erwerben bzw. das eigene Fähigkeitsniveau zu verbessern“ (Krapp et al., 2014, S. 194). Nach Schiefele (1996, S. 52) ist Lernen intrinsisch motiviert, wenn „die Handlung selbst als interessant, spannend oder [...] zufriedenstellend“ bewertet wird. Die extrinsische Motivation ist dagegen durch äußere Anreize begründbar. Mit der *Self-Determination Theory* erklären Deci und Ryan (1985) das Auftreten intrinsischer Motivation mit der Befriedigung von Grundbedürfnissen nach Autonomie-, Kompetenzerleben und sozialer Eingebundenheit.

Eng verbunden mit der Lesefreude sind auch lesebezogene Selbstwirksamkeitserwartungen, die das Leseverhalten und den Umgang mit Texten beeinflussen können (Munser-Kiefer, 2014; Unrau et al., 2018). Selbstwirksamkeitserwartun-

gen werden als subjektive Überzeugungen definiert, „Anforderungssituationen auf Grund eigener Kompetenz bewältigen zu können“ (Schwarzer & Jerusalem, 2002, S. 35). Diese werden durch positive Erfahrungen bei der Bearbeitung von Aufgaben und positive Leistungsrückmeldungen geformt (Bandura, 1977). Hohe Selbstwirksamkeitserwartungen im Lesen bringen demnach eine höhere Anstrengungsbereitschaft mit sich, auch schwierige Texte zu lesen (Munser-Kiefer, 2014).

2.3.2 Kooperatives Lernen

Kooperatives Lernen ist ein zentraler Bestandteil des Reziproken Lehrens. Kooperative Lernformen sind in (sozial)konstruktivistischen Lerntheorien fundiert und werden definiert als „Lernarrangements [...], die eine synchrone und koordinierte, ko-konstruktive Aktivität der Teilnehmer/innen verlangen, um eine gemeinsame Lösung eines Problems oder ein gemeinsam geteiltes Verständnis einer Situation zu entwickeln“ (Pauli & Reusser, 2000, S. 421). Ein wichtiges Kernelement kooperativen Lernens ist das Erreichen eines gemeinsamen Lernziels (z. B. Johnson & Johnson, 1998). Johnson und Johnson (1998) konzeptualisieren fünf Basiselemente, die für effektives kooperatives Lernen gegeben sein müssen. Diese umfassen die *positive Interdependenz* zwischen den Gruppenmitgliedern, *gegenseitige Unterstützung*, *individuelle Verantwortlichkeit*, *interpersonale Fähigkeiten* und *reflexive Gruppenprozesse* (vgl. auch Büttner et al., 2012). Positive Interdependenz zwischen den Gruppenmitgliedern kann bspw. über ein gemeinsames Gruppenprodukt (Zielinterdependenz) oder eine gemeinsame Belohnung (Belohnungsinterdependenz) erreicht werden (Büttner et al., 2012). Daran knüpft auch die individuelle Verantwortlichkeit der Teilnehmer*innen an, den eigenen Arbeitsauftrag zu bearbeiten, um damit zum Erreichen des gemeinsamen Gruppenziels beizutragen. Damit die Umsetzung dieser Kernelemente gelingt, müssen Schüler*innen notwendige kognitive und soziale Fähigkeiten mitbringen, um den Lernprozess gemeinsam zu gestalten, sich gegenseitig zu unterstützen und die Gruppenarbeit effektiv zu reflektieren.

Die Wirksamkeit kooperativen Lernens gegenüber anderen Instruktionsformen (z. B. Einzelarbeit) konnte in einer Vielzahl von Studien zu unterschiedlichen Zielvariablen, sowohl kognitiver als auch motivationaler Art, bestätigt werden (z. B. Ginsburg-Block et al., 2006; Johnson et al., 2000; Stevens & Slavin, 1995; van Ryzin & Roseth, 2018). Wirksamkeitsmechanismen werden hierbei über unterschiedliche theoretische Ansätze argumentiert (Slavin et al., 2003). Motivationale Perspektiven zur Wirksamkeit fokussieren bspw. extrinsische Anreize, wie Gruppenbelohnungen, die sich durch die Erreichung eines gemeinsamen Ziels ergeben können. Im Sinne der Perspektive der sozialen Kohäsion können sich motivationale Anreize auch aus der Zusammenarbeit in der Gruppe und der gemeinsamen Identifikation ergeben (Borsch, 2018; Slavin et al., 2003). Die Gruppenkohäsion kann beim kooperativen Lernen über eine gemeinsame Gruppenidentität gestärkt werden.

Aus kognitiver Perspektive wird zur Erklärung der Wirksamkeit kooperativer Methoden Vygotskys (1978) konstruktivistische Lerntheorie herangezogen. Nach dieser ist effektives Lernen in der *Zone der nächsten Entwicklung* anzusiedeln, die zwischen dem aktuellen und dem potenziellen Fähigkeitsniveau von Schüler*innen liegt. Die Zone der nächsten Entwicklung kann nach Vygotsky (1978) in der Interaktion mit anderen Personen erreicht werden, indem gemeinsam Wissen konstruiert wird (Borsch, 2018). Hierzu kann die Methode des *Scaffoldings* eingesetzt werden (Bruner, 1985). Beim Reziproken Lehren werden dabei Unterstützungsmaßnahmen durch die Lehrkraft oder eine*n Tutor*in gegeben, die im Lernprozess schrittweise zurückgenommen werden, bis die Schüler*innen die Aufgaben selbstständig bearbeiten können und die Zone der nächsten Entwicklung erreicht haben (vgl. Kapitel 2.3.3).

Zum Lernerfolg kooperativer Methoden ist zudem aus kognitiver Perspektive die elaborative Verarbeitung der Inhalte von Bedeutung (Slavin et al., 2003). Aus dieser Perspektive entsteht Lernen durch das Lösen kognitiver Konflikte und durch die Integration und Umstrukturierung neuen Wissens. Die elaborative Verarbeitung kann in der Gruppenarbeit durch lernförderliche Aktivitäten gefördert werden. Kaendler et al. (2015) unterscheiden drei Arten von lernförderlichen Aktivitäten: kooperative, kognitive und metakognitive Aktivitäten. Kooperative Aktivitäten beschreiben das gemeinsame Arbeiten und das Einhalten von Regeln für eine effektive Gruppenarbeit, wie z. B. einander zuzuhören und einander ausreden zu lassen (Kaendler et al., 2015; Lipowsky et al., 2022). Kognitive lernförderliche Aktivitäten sind dagegen z. B. in gegenseitigen Erklärungen, Argumentationen, Fragen und Begründungen zu finden, die eine elaborierte Verarbeitung der Inhalte ermöglichen (Kaendler et al., 2015; Lipowsky et al., 2022; Webb, 1989). Metakognitive Aktivitäten beziehen sich schließlich auf die Steuerung und Regulation der Gruppenarbeitsprozesse (Kaendler et al., 2015). Mit den genannten Aktivitäten kann die Interaktionsqualität in den Gruppen gesteigert werden, was zu einem höheren Lernerfolg des kooperativen Lernens führt.

Insgesamt kommt der Lehrkraft beim kooperativen Lernen eine wichtige Rolle zu, da kooperative Lernarrangement angemessen angeleitet und begleitet werden müssen, um die notwendigen strukturellen Bedingungen herzustellen, die für die Wirksamkeit zentral sind (vgl. Büttner et al., 2012; Pauli & Reusser, 2000). Zudem müssen Schüler*innen darin unterstützt werden, qualitativ hochwertige Interaktionen zu vollziehen, die eine elaborierte Verarbeitung der Inhalte ermöglichen (Lipowsky et al., 2022).

2.3.3 Instruktion von Lesestrategien beim Reziproken Lehren

Die kooperative Methode des Reziproken Lehrens stellt eine wirksame Methode der Leseförderung dar, in deren Kern die Instruktion von Lesestrategien steht. Lesestrategien gelten als wichtige Werkzeuge bei der Förderung von Lesekompetenz

(Munser-Kiefer, 2014) und werden definiert als „prinzipiell bewusstseinsfähige, aber [...] oft automatisierte mentale Handlungsfolgen, die gezielt und adaptiv eingesetzt werden, um Textverstehen zu erreichen“ (Rosebrock, 2018, S. 19). Kognitive Lesestrategien werden angewendet, um die Informationsverarbeitung zu unterstützen, z. B. durch das Markieren und Zusammenfassen von wichtigen Inhalten oder die Anknüpfung an Vorwissen (Souvignier, 2014). Dagegen haben metakognitive Strategien zum Ziel, den Lernprozess zu überwachen und zu reflektieren (Souvignier, 2014).

Nach Palinscar & Brown (1984) werden beim Reziproken Lehren vier Lesestrategien eingeübt. Diese umfassen das Klären unbekannter Wörter, Fragen an den Text zu stellen, einen Abschnitt zusammenzufassen und den Inhalt des nächsten Abschnitts vorherzusagen. Das Einüben der Strategien wird in der Gruppenarbeit durch die Interaktion zwischen der Lehrkraft und den Schüler*innen vollzogen. Schüler*innen lesen dabei abwechselnd einen Absatz eines Textes und wenden die Lesestrategien auf den Absatz an. Die Anwendung der Strategien wird jedoch zunächst nach der Methode des Scaffoldings durch die Lehrkraft eingeführt. Scaffolding umfasst Methoden und Prozesse, die den Lernprozess der Schüler*innen unterstützen. Beim Reziproken Lehren ist dies so zu verstehen, dass die Lehrkraft die Anwendung der Strategien zunächst modelliert und dabei ihre kognitiven Prozesse für die Schüler*innen verbalisiert (Munser-Kiefer, 2014). Palinscar und Brown (1984, S. 123) beschreiben dies als „expert scaffolding“, das als Unterstützungsmaßnahme zum Erreichen der Zone der nächsten Entwicklung (Vygotsky, 1978; vgl. Kapitel 2.3.2) dient. Der Dialog wird zunächst von der Lehrkraft geleitet, soll zunehmend jedoch durch die Schüler*innen übernommen werden. Dabei unterstützt die Lehrkraft die Anwendung der Strategien durch Feedback, Rückfragen und Erklärungen.

Die Unterstützung wird im weiteren Verlauf zunehmend zurückgenommen, sodass Schüler*innen die eigenständige Anwendung der Strategien schrittweise erlernen (Palinscar & Brown, 1984; Rosenshine & Meister, 1994). Schüler*innen sollen dabei ihre Strategieranwendung gegenseitig überwachen und unterstützen. Dieser Prozess kann durch strukturierte Arbeitsphasen und die Bereitstellung von Unterrichtsmaterial gefördert werden. In der Kleingruppenarbeit werden in verschiedenen Adaptionen des Reziproken Lehrens sogenannte *Promptkarten* zur Verfügung gestellt, auf denen die wichtigsten Schritte zum Ausführen der Strategien beschrieben sind (vgl. Decristan et al., 2021; Decristan et al., 2022; Spörer et al., 2016). Mithilfe der Promptkarten können die Lesestrategien im kooperativen Setting systematisch eingeübt und der Dialog zwischen den Schüler*innen strukturiert werden. Die Darstellung der einzelnen Schritte unterstützt dabei die kognitiven Prozesse und kann als Basis für gemeinsame Erklärungen, Argumentationen und Begründungen fungieren, die die elaborative Verarbeitung fördern.

Die Wirksamkeit des Reziproken Lehrens konnte in vielen Studien und Metaanalysen bestätigt werden (z. B. Fung et al., 2003; Rosenshine & Meister, 1994; Spörer et al., 2009). Schüler*innen, die die hierarchieniedrigeren Prozesse des

Dekodierens schon beherrschen, jedoch im Textverstehen noch Defizite aufweisen, profitieren dabei besonders von der strukturierten Instruktionmethode (Rosenshine & Meister, 1994). Im unterrichtlichen Setting hat es sich zudem als besonders wirksam erwiesen, die Strategien zunächst explizit zu instruieren, bevor diese im kooperativen Setting weiter eingeübt werden (*explicit teaching before reciprocal teaching*; Rosenshine & Meister, 1994). Nach Fung et al. (2003) können Schüler*innen in der Gruppenarbeit dadurch an bereits bestehendes Vorwissen besonders gut anknüpfen. Zudem nehmen die Autor*innen an, dass dies für Schüler*innen mit geringeren sprachlichen Kompetenzen in der Verkehrssprache eine kognitive Entlastung darstellen kann. Klinger und Vaughn (2000) stellen heraus, dass auch bilinguale Schüler*innen von der kooperativen Strategieinstruktion profitieren und sich im Tandem auf unterschiedlichen Sprachen unterstützen, um den Lernprozess zu organisieren, aber auch Hürden im Wortschatz und in der Strategieanwendung überwinden. Die Qualität der Interaktion in der Kleingruppe ist insbesondere bedeutsam für die Wirksamkeit des Reziproken Lehrens (vgl. Palinscar & Brown, 1984; Hacker & Tenent, 2002). Diese kann unzureichend ausfallen, wenn Schüler*innen geringe kooperative Kompetenzen mitbringen und von der Lehrkraft nicht angemessen angeleitet werden (Hacker & Tenent, 2002).

Für den Grundschulunterricht haben sich die Strategien „Fragen stellen“ und „Zusammenfassen“ als besonders wirksam erwiesen (Hacker & Tenent, 2002; Koch & Spörer, 2016). Die Empfehlung lautet hierbei, eine überschaubare Anzahl an kognitiven Lesestrategien mit der Instruktion von metakognitiven Strategien zur Regulation des Lernprozesses anzureichern (Koch & Spörer, 2016; Munser-Kiefer, 2014).

2.4 Mehrsprachigkeitssensibles Reziprokes Lehren

Die Methode des Reziproken Lehrens birgt viele Potenziale, schwächere Schüler*innen im Leseverstehen zu fördern und mehrsprachigkeitssensible Elemente in den Lernprozess zu integrieren. Der kooperative Rahmen des Reziproken Lehrens kann dabei genutzt werden, um mehrsprachige Interaktionen in den Kleingruppen zu ermöglichen, sodass sich Schüler*innen gleicher Familiensprachen im Lernprozess gegenseitig unterstützen können. Decristan et al. (2022) stellen eine quasi-experimentelle Interventionsstudie vor, in der das Reziproke Lehren in vierten Klassen mit mehrsprachigkeitssensiblen Impulsen angereichert wird (mehrsprachigkeitssensibles Reziprokes Lehren). Hierbei werden Strategien des pädagogischen Translanguagings und einer mehrsprachigkeitssensiblen Unterrichtspraxis herangezogen und ein wertschätzendes Klima für die sprachliche Heterogenität der Schüler*innen geschaffen. Zur Unterstützung der Strategieinstruktion werden zudem bilinguale Promptkarten auf Deutsch und in den Familiensprachen der Schüler*innen bereitgestellt. Die Nutzung der Familiensprachen im Lernprozess stellt dabei ein freiwilliges Angebot dar, das durch die Erweiterung

des sprachlichen Repertoires und der Möglichkeit zur Anknüpfung an Vorwissen in den Familiensprachen kommunikative und kognitive Funktionen im Lernprozess einnehmen kann (vgl. Kapitel 2.1.4). Die Ergebnisse der Interventionsstudie zeigen, dass das mehrsprachigkeitssensible Reziproke Lehren geeignet ist, die Lesestrategieanwendung bei Schüler*innen zu fördern. Zudem steht die Aktivierung des mehrsprachigen Repertoires im Zusammenhang mit höheren Lernzuwächsen. Mehrsprachige Schüler*innen, die angeben, während der Intervention auch eine andere Sprache als Deutsch genutzt zu haben, weisen höhere Zuwächse auf als mehrsprachige Schüler*innen, die lediglich Deutsch gesprochen haben. Dagegen zeigt der Vergleich zwischen mehrsprachigen und monolingualen Schüler*innen, dass beide Gruppen in gleichem Maße von der Interventionsstudie profitieren. Die Ergebnisse verdeutlichen, dass die Aktivierung mehrsprachiger Interaktion und aktive Sprachnutzung für mehrsprachige Schüler*innen hilfreich sein kann, Facetten der Lesekompetenz zu fördern.

In der vorliegenden Arbeit wird der Implementationserfolg der Intervention zum mehrsprachigkeitssensiblen Reziproken Lehren nach Decristan et al. (2022) aus Schüler*innenperspektive untersucht. Im folgenden Kapitel wird dazu der Themenkomplex „Implementationserfolg von Interventionen im Unterricht“ dargestellt, um geeignete Facetten von Implementationserfolg für die empirische Untersuchung abzuleiten und die theoretischen Grundlagen zur Untersuchung der Zusammenhänge aufzuzeigen.

3 Implementationserfolg von Interventionen im Unterricht

Innovative didaktische Konzepte müssen sich nicht nur als wirksam erweisen, sondern auch praxiskompatibel und gut umsetzbar sein, um langfristig Anwendung zu finden. Wie dies gelingt, wird im Rahmen der Implementationsforschung untersucht. Im folgenden Kapitel werden zunächst die Begriffe der Implementation und Innovation beschrieben (vgl. Kapitel 3.1). Weiterhin werden die zwei für die vorliegende Arbeit zentralen Facetten von Implementationserfolg, die Wiedergabebetreue und Akzeptanz, herausgearbeitet und definiert (vgl. Kapitel 3.2 –3.4). Für den spezifischen Kontext der unterrichtlichen Interventionen werden abschließend aktuelle Forschungsbefunde zu den Facetten des Implementationserfolgs im unterrichtlichen Kontext zusammengefasst (vgl. Kapitel 3.5).

3.1 Implementation von Innovationen

Innovationen treiben gesellschaftliche Weiterentwicklung an und sind in den verschiedensten Forschungsdisziplinen von Relevanz (Goldenbaum, 2012; Schmieder, 2019). Sie werden als Lösung für bestehende Probleme herangezogen und sollen zu positiven gesellschaftlichen Veränderungen führen (Altrichter & Wiesinger, 2004; Hunneshagen, 2005; John, 2013). Eine einheitliche Definition für Innovationen ist durch eine Vielzahl von theoretischen Zugängen und unterschiedlichen Anwendungsbereichen nicht zu finden. Zentrale Charakteristika über verschiedene Forschungsdisziplinen hinweg stellen jedoch die Aspekte der Neuheit und Verbesserung in den Vordergrund (Goldenbaum, 2012; Gröschner, 2011). Sie konstituieren sich nach der subjektiven Wahrnehmung der handelnden Akteur*innen und können nicht objektiv bestimmt werden (Rogers, 2003). Vielmehr werden Innovationen in Auseinandersetzung mit ihrer Umwelt konstruiert und auch als soziales Phänomen definiert, das stets in einen sozialen Kontext eingebettet ist (John, 2013).

Eine weit verbreitete Differenzierung verschiedener Arten von Innovationen wird über Produkt-, Prozess- und Sozialinnovationen getroffen (Goldenbaum, 2012). Produktinnovationen werden dabei als „Produkte und/oder Dienstleistungen“ (Franken & Franken, 2020, S. 250) definiert und sind häufig Formen neuer Technologien. Sozialwissenschaftliche Betrachtungsweisen ergänzen dieses technisch-ökonomische Verständnis von Innovationen als Gegenstände materieller Natur durch die Definition sozialer Innovationen (Braun-Thürmann, 2005; John, 2013). Soziale Innovationen beschreiben Prozesse „abstrakte[r]“ Art, die sich auf die Veränderung „sozial-gesellschaftlicher Bedingungen“ beziehen (Goldenbaum, 2012, S. 78) und können auch als „symbolische Artefakte“ (Braun-Thürmann, 2005, S. 6) bestehen. Demnach sind soziale Innovationen als „theoretische Konstrukte“ (Gillwald, 2000, S. 8) zu verstehen, die sich als „gesellschaftliche Errun-

genschaften“ (Gillwald, 2000, S. 14) darstellen lassen und im Kern Verhaltensänderungen sowie Veränderungen menschlicher Beziehungen widerspiegeln. Gillwald (2000, S. 4) fasst soziale Innovationen schließlich als „Teil und Ergebnis von sozialem Wandel“ zusammen, die zur Lösung spezifischer Probleme eingesetzt werden und sich von bekannten Praktiken abgrenzen lassen müssen, um dem Merkmal der Neuartigkeit gerecht zu werden. Anhand der genannten Differenzierung sind verschiedene Arten von Innovationen jedoch meist nicht trennscharf einer einzigen Kategorie zuzuordnen. Vielmehr wird zwischen den unterschiedlichen Typen von Innovationen von einem „komplexe[n] Wirkungsgeflecht“ (Franken & Franken, 2020, S. 256) ausgegangen. Eine Produktinnovation kann demnach auch die Prozesse einer Organisation beeinflussen, die wiederum Veränderungen auf technischer sowie sozialer Ebene hervorbringen, woran sich die Veränderung von Praktiken in Form von sozialen Innovationen anschließt. Im Kontext bildungswissenschaftlicher Perspektiven steht jedoch meist der soziale Charakter von Innovationen im Vordergrund, die unterschiedliche Gestalt annehmen können. Century und Cassata (2016) arbeiten aus einer Vielzahl bildungswissenschaftlicher und pädagogischer Forschungsbereiche mögliche Arten von Innovationen heraus und definieren diese als „programs, technologies, processes, approaches, methods, strategies or polices that involve change“ (Century & Cassata, 2016, S. 170). Zusammenfassend werden Innovationen hierbei als das Objekt der Veränderung betrachtet.

Im schulischen Kontext argumentiert Gröschner (2011) für ein konstruktivistisch geprägtes Verständnis von Innovationen, nach welchem individuelle Kompetenzen der Akteur*innen die Grundlage für einen Innovationsprozess darstellen, da diese die Umsetzung der Innovation maßgeblich beeinflussen. Nach dieser Sichtweise sind individuelle Kompetenzen und Merkmale sowohl Voraussetzung für Innovationsprozesse sowie Resultate der Veränderung. Gröschner (2011, S. 46–47) beschreibt, dass in diesem „individuumzentrierte[n] und kompetenzorientierte[n] Ansatz“ der Lernprozess als „Veränderung durch nachhaltige Entwicklung“ ein „Hauptmerkmal von Innovationsprozessen“ darstellt. Nach Gröschner (2011, S. 55) ist eine Innovation in diesem Kontext als „aktiver Entwicklungs- und Veränderungsprozess“ zu definieren, der auf einem individuellen Lernprozess der Akteur*innen beruht. Gröschner (2011, S. 54–55) resümiert, dass „dem lernenden Individuum [...] dabei eine Schlüsselrolle [zukomme]“ und im konstruktivistischen Verständnis individuelle Kompetenzen und Entwicklungen „zu einem wesentlichen Bestandteil von Innovationsprozessen“ werden. Vor dem Hintergrund dieser konstruktivistischen Sichtweise auf Innovationen und ihre Umsetzung in der Praxis sind die Akteur*innen und ihre Voraussetzungen und Kompetenzen daher besonders relevant für erfolgreiche Implementationsprozesse. Hasselhorn et al. (2014, S. 148) plädieren daher für eine verstärkte Einbindung „psychologischer Expertise“ in die Implementationsforschung, um individuelle Prozesse zu untersuchen.

Zur Veränderung und Neugestaltung von Lern- und Bildungsprozessen können didaktische Innovationen zum Einsatz kommen (Lauer, 2006). Diese werden „mit dem Ziel der Verbesserung und Optimierung von Lehr, Lern- und Entwicklungsprozessen von Lernenden wie Lehrenden implementiert“ (Ziegelbauer & Gläser-Zikuda, 2016, S. 9). Didaktische Innovationen können sich nach Reinmann-Rothmeier (2003, S. 11) auf „Inhalte und/oder Methoden des Lehrens [beziehen], die den vorangegangenen Zustand der Wissensvermittlung merklich verändern und als Konsequenz auch einen Wandel der intendierten Bildungs- und Lernprozesse bewirken“. Curriculare Innovationen können bspw. Reformen der Lehrpläne beinhalten. Organisationale Innovationen beziehen sich dagegen auf die Rahmenbedingungen der Bildungsprozesse in den Schulen, wie z. B. die Einführung von Ganztagschulen (Reinmann-Rothmeier, 2003). Instruktionale Innovationen umfassen schließlich die „Einführung neuer Lehr-Lernmethoden“ (Reinmann-Rothmeier, 2003, S. 11), wie z. B. kooperative Methoden oder den Einsatz digitaler Medien im Unterricht. Didaktische Innovationen beziehen sich auf „Bildungs- oder Lehr-Lernproblem[e]“ (Reinmann-Rothmeier, 2003, S. 12) und sind in diesem Sinne sozialen Innovationen zuzuordnen. Didaktische Innovationen können jedoch eine Vielzahl von Neuerungen umfassen, die auch andere Arten von Innovationen nicht ausschließen. Der Einsatz digitaler Medien stellt bspw. eine Produktinnovation dar. Diese können allerdings auch die Prozesse des Lernens maßgeblich beeinflussen und knüpfen demnach an Prozessinnovationen an. Didaktische Innovationen werden im Unterrichtskontext häufig im Rahmen von Interventionen eingeführt. Leutner (2010, S. 18) definiert den Begriff Intervention als „jede gezielte Veränderung [und] jeder gezielte Eingriff in ein System“, der eine Verbesserung der aktuellen unterrichtlichen Praxis erzielen soll. Eine Intervention ist demnach eng mit dem Verständnis von Innovationen verknüpft. Die Begriffe werden in der Literatur daher häufig synonym verwendet (Durlak & DuPre, 2008).

Ein weiteres wichtiges Kriterium von Innovationen wird im Prozess ihrer Umsetzung in die Praxis gesehen (Franken & Franken, 2020; Hunneshagen, 2005). Eine Neuerung stellt demnach erst dann eine Innovation dar, wenn diese in der Praxis zum Einsatz kommt. Dieser Prozess der Umsetzung wird auch als Implementation beschrieben. Mit der Implementation einer Innovation soll eine positive Veränderung einer bisherigen Praxis erzielt werden (Goldenbaum, 2012). Auch der Begriff der Implementation hat im wissenschaftlichen Sprachgebrauch keine einheitliche Bedeutung, weist jedoch viele Überschneidungen zum Verständnis von Innovation auf. In bildungswissenschaftlicher Literatur wird Implementation als die „Umsetzung“ bzw. „Überführung“ von Innovationen in die Schulpraxis definiert (Goldenbaum, 2012, S. 83). Spezifischer betont eine häufig rezipierte Definition von Euler und Sloane (1998, S. 312) die empirische Komponente eines Forschungsprozesses mit der „Umsetzung von wissenschaftlichen Erkenntnissen in die gesellschaftliche Praxis“. Worin über verschiedene Forschungsperspektiven hinweg Einigkeit besteht, ist, dass Implementation als Prozess verstanden wird, der mit der Einführung einer Innovation in das praktische Feld beginnt und erst dann

als beendet gelten kann, wenn die Innovation effektiv und routiniert als institutionalisierte Praxis eingesetzt wird (Altrichter & Wiesinger, 2004; Fixsen & Blase, 2009; Hasselhorn et al., 2014). Dieser Verlauf zwischen erstmaligem Einsatz und routinierter Praxis wird einem engeren Verständnis von Implementation zugeordnet (Goldenbaum, 2012). Der Implementationsprozess kann jedoch auch in einem weiter gefassten Verständnis zusätzliche Stufen der Umsetzung umfassen, wie die Verbreitung der Innovation im breiteren kontextualen Feld, z. B. die Ausweitung auf mehrere Schulen (Goldenbaum, 2012). Bei einem solchen weiter gefassten Verständnis sind zentrale Überschneidungen zu verwandten Perspektiven der Implementationsforschung, wie der Transfer- und Diffusionsforschung, zu erkennen. Die Transferforschung befasst sich nach erfolgreicher Umsetzung einer Innovation mit Fragen der Verbreitung in andere Kontexte und Felder. Ebenso die Diffusionsforschung, die insbesondere den zeitlichen Verlauf und Prozess der Weitergabe zwischen den handelnden Akteur*innen in den Fokus nimmt (Rogers, 2003; Roth et al., 2021). In aktuellen internationalen Beiträgen zu Implementationsfragen aus dem Bildungsbereich ist ein engeres Verständnis von Implementationsforschung meist vorherrschend, das mit der Umsetzung einer Innovation in einem spezifischen Kontext endet (z. B. im Unterricht, in einer Schule) und Bedingungen des Erfolgs in genau diesem Kontext untersucht (Schrader & Hasselhorn, 2020).

Um Unterschiede in den verwandten Konstrukten der Interventions-, Implementations-, Transfer- und Diffusionsforschung besser aufzuzeigen, können diese hierarchisch organisiert und in ein zeitlich gestaffeltes Verhältnis gebracht werden. Idealtypisch setzt die Implementationsforschung nach der Interventionsforschung an, in deren Rahmen die Wirksamkeit einer Maßnahme unter kontrollierten Bedingungen bestätigt wurde, bevor ihre Umsetzung in der Praxis untersucht wird (Leutner, 2010; Roth et al., 2021; Souvignier, 2020). Im Vordergrund der Interventionsforschung steht die „systematisch[e] Evaluation“ (Souvignier, 2020, S. 1) der eingesetzten Maßnahme unter Verwendung wissenschaftlicher Methoden und, soweit möglich, unter Einhaltung kontrollierter Bedingungen zur Untersuchung kausaler Zusammenhänge (Roth et al., 2021). Auf diese Weise sollen Unterrichtskonzepte identifiziert werden, die eine höhere Wirksamkeit aufweisen als bisherige Unterrichtsmethoden (Souvignier, 2020). Da in der Forschungspraxis häufig auch die Umsetzung von Maßnahmen ohne vorherige Wirksamkeitsstudie untersucht wird, differenzieren Schrader et al. (2020) zwischen einem engeren Verständnis von Implementationsforschung und einem breiteren Verständnis. Letzteres schließt auch Implementationsstudien ein, die der idealtypischen Abfolge einer vorangehenden Interventionsstudie nicht entsprechen. Die Transfer- und Diffusionsforschung beschäftigt sich darauf folgend mit der Verbreitung und Überführung der Innovation in die breitere Praxis. Die Diffusionsforschung fokussiert dabei weniger das Individuum selbst als den „Verlauf einer Verbreitung von Innovationen“ (Gräsel, 2010, S. 9) und wie dieser begünstigt werden kann.

Schulische Implementationsstudien werden mit dem Ziel eingesetzt, „(pädagogische) Interventionen im Feld zu erproben und Bedingungen, Prozesse bzw.

Wirkungen ihrer Implementierung zu untersuchen“ (Schrader et al., 2020, S. 12). Hierbei soll insbesondere die Umsetzbarkeit in das praktische Feld in den Blick genommen werden. In den vergangenen Jahren ist eine deutliche Zunahme an Implementationsstudien im schulischen Kontext zu verzeichnen, die die Umsetzung innovativer Ansätze auf verschiedenen Ebenen des Bildungssystems untersuchen (Schrader & Hasselhorn, 2020). In schulischen Implementationsstudien wird die „Diskrepanz zwischen Planung und Verwirklichung von Konzepten“ (Petermann, 2014, S. 122) sowie die Operationalisierung von Kriterien erfolgreicher Umsetzung analysiert. Zudem steht die „systematische Bearbeitung der Frage im Vordergrund, unter welchen Bedingungen die Umsetzung von Neuerungen in die Praxis erfolgt“ (Souvignier & Philipp, 2016, S. 11). In der schulischen und unterrichtlichen Praxis zeigen die implementierten Konzepte häufig eine geringere Wirksamkeit, da es über die verschiedenen Implementationsebenen hinweg zu zahlreichen Adaptationen und individuellen Schwierigkeiten und Barrieren im Implementationsprozess kommen kann und die Interventionen nicht wie geplant umgesetzt werden (Holtappels, 2013; Lendrum & Humphrey, 2012). Daher ist die Identifikation von „institutionellen, organisationalen und personalen“ (Schrader et al., 2020, S. 15) Merkmalen, die die Umsetzung fördern bzw. hemmen können, eine zentrale Forschungsaufgabe.

In der schulischen Implementationsforschung werden verschiedene Strategien unterschieden, mit denen die Umsetzung von Innovationen in die Praxis gelingt. Eine weit verbreitete Einteilung in drei Arten von Strategien geht auf Gräsel und Parchmann (2004) zurück. Die Autorinnen unterscheiden zwischen sogenannten Top-Down-Strategien, Bottom-Up-Strategien und symbiotischen Strategien, die in ihren Kernelementen im Folgenden skizziert werden. Implementationen nach dem Prinzip der Top-Down-Strategien werden vor dem Hintergrund der *Fidelity*-Perspektive beschrieben, mit der eine möglichst treue Umsetzung einer Innovation vorgesehen wird (Gräsel & Parchmann, 2004). Die Innovation, z. B. die Einführung eines neuen Lehrplans, wird hierbei von einer externen Instanz, bspw. einer wissenschaftlichen oder politischen Institution, in den Schul- oder Unterrichtskontext herangetragen (Gräsel & Parchmann, 2004). Extern und vorab wird festgelegt und dokumentiert, bspw. mittels eines Manuals, wie die Umsetzung erfolgen soll und an welchen Kriterien der Erfolg der Innovation gemessen wird (Goldbaum, 2012). Nach Hasselhorn et al. (2014, S. 142) liegt Top-Down-Strategien ein „technokratisches Implementationsverständnis mit hierarchischem Akteursgefüge“ zugrunde. Top-Down-Strategien werden demnach von zwei Gruppen von Akteur*innen definiert (Hasselhorn et al., 2014). Die externe Instanz, die die Innovation an die Zielorganisation heranträgt, sowie die handelnden Akteur*innen in der Organisation, in der die Innovation umgesetzt werden soll. Hierarchisch ist das Gefüge deshalb, weil eine „lineare Übernahme“ (Hasselhorn et al., 2014, S. 142) angestrebt wird und den umsetzenden Akteur*innen (Schulleitungen oder Lehrkräfte) kein bis wenig Gestaltungsspielraum zugeschrieben wird. Diese Trennung der Akteur*innengruppen kann jedoch kritisch betrachtet werden, was

Gräsel und Parchmann (2004, S. 199) auch als „personelle und zeitliche Trennung zwischen Konzeption und Umsetzung“ beschreiben. Dadurch besteht die Gefahr, dass wichtige Bedürfnisse der Praxis nicht berücksichtigt werden und die Implementation als Konsequenz keine langfristige Einbettung in den Schulkontext findet (Gräsel & Parchmann, 2004). Als Vorteil der Top-Down-Strategie gilt stattdessen die Klarheit der vorab festgelegten Ziele. Konträr zu Top-Down-Strategien werden bei Bottom-Up-Strategien die Innovationen intern in der handelnden Organisation selbst entwickelt und erprobt (Goldenbaum, 2012; Gräsel & Parchmann, 2004). Ziele der Umsetzung werden hierbei von der Organisation selbst festgelegt und orientieren sich an zuvor identifizierten Bedarfen der Veränderung und Optimierung im Sinne eines Entwicklungsprozesses (Goldenbaum, 2012). Eine Trennung von Akteur*innen ist hierbei nicht gegeben, woraus sich der Vorteil einer stärkeren Praxisintegration ergibt. Als Beispiel für Bottom-Up-Strategien nennen Gräsel und Parchmann (2004) Schulentwicklungsprozesse. Symbiotische Strategien verbinden schließlich Aspekte der Top-Down- und Bottom-Up-Strategien, indem Veränderungsprozesse von unterschiedlichen Akteur*innen gemeinsam gestaltet und bearbeitet werden. Eine besondere Stärke stellt dabei dar, dass unterschiedliche pädagogische Expertisen von Lehrkräften und Bildungsforscher*innen und politischen Akteur*innen zusammenkommen, um gemeinsam identifizierte Probleme durch Innovationen zu bewältigen (Goldenbaum, 2012; Gräsel & Parchmann, 2004). Konzeption, Umsetzung sowie die anschließende Evaluation und Adaption des Implementationsprozesses wird gemeinsam gestaltet und hat daher „partizipativen“ Charakter (Hasselhorn et al., 2014, S. 143).

3.2 Definitionen von Implementationserfolg

Um Veränderungen in der Praxis erzielen zu können, muss die Umsetzung der Innovation gelingen und der Implementationsprozess erfolgreich verlaufen. Schuli-sche Implementationsforschung beschäftigt sich zu einem großen Teil mit der Frage, welche Bedingungen eine erfolgreiche Implementation fördern oder hemmen können. Hierzu muss jedoch vorab definiert werden, wann eine Implementation als erfolgreich gelten kann. Für die Beschreibung und Messung von Implementationserfolg gibt es unterschiedliche Herangehensweisen und variierende Definitionen stellen eine Herausforderung für die Systematisierung und Vergleichbarkeit von Ergebnissen dar. Zunächst ist festzuhalten, dass sich der Erfolg einer Implementation maßgeblich an der zugrundeliegenden Strategie und den vorher festgelegten Zielen der Implementation orientiert (Hasselhorn et al., 2014). Wie beschrieben können für eine Top-Down-Strategie demnach andere Erfolgskriterien angemessen sein als für Implementationsprozesse nach einem Bottom-Up-Verfahren (vgl. Kapitel 3.1). Für Strategien, die nach dem Top-Down-Prinzip von wissenschaftlicher Seite in Bildungseinrichtungen eingeführt werden, steht eine möglichst getreue Umsetzung der Intervention im Fokus. Doch auch in diesem

Kontext variiert die empirische Umsetzung und Evaluation von Erfolgskriterien in Forschungsarbeiten.

Die Arbeiten von Proctor et al. (2011) sowie Petermann (2014) haben mit dem Entwurf einer Taxonomie von Implementationsoutcomes auch für die bildungsbezogene Implementationsforschung einen wichtigen Beitrag geleistet, indem sie die Untersuchung von Implementationsprozessen und -erfolgen systematisieren, die zu einer Vergleichbarkeit zwischen Studien beitragen sollen. Implementationsoutcomes werden dabei definiert als „effects of deliberate and purposive actions to implement new treatments, practices and services“ (Proctor et al., 2011, S. 65) und dienen als Indikatoren für (1) den Implementationserfolg, (2) den Implementationsprozess und (3) die Wirksamkeit einer Intervention (Proctor et al., 2011). Die Implementationsoutcomes nach Proctor et al. (2011) sowie Petermann (2014) umfassen die Konstrukte Wiedergabetreue, Akzeptanz, Angemessenheit, Machbarkeit, Kosten, Übernahme, Durchdringung und Nachhaltigkeit. Diese werden wie folgt definiert:

(1) Die *Wiedergabetreue* stellt insbesondere für Implementationsstudien aus Top-Down-Perspektive einen der zentralsten Erfolgsindikatoren dar, der in der bildungsbezogenen Interventions- und Implementationsforschung bislang die meiste Aufmerksamkeit erlangt hat. Wiedergabetreue beschreibt das Maß, in dem eine Innovation wie vorgesehen umgesetzt wird (Dane & Schneider, 1998; Petermann, 2014; Proctor et al., 2011) (vgl. Kapitel 3.3). (2) Auch die *Akzeptanz* stellt einen der „aussagekräftigsten Parameter zur Beurteilung des Behandlungsangebotes“ (Petermann, 2014, S. 123) dar und wird häufig mit der Zufriedenheit der Nutzer*innen mit einem Angebot definiert. Akzeptanz bildet eine „grundlegende Voraussetzung dafür, dass ein Programm oder Prinzip sich überhaupt in der (schulischen) Praxis etablieren kann“ (Souvignier & Philipp, 2016, 10f.) (vgl. Kapitel 3.4). (3) Die *Angemessenheit* bezieht sich auf die „wahrgenommene Passung“ (Petermann, 2014, S. 124) der Innovation in den anvisierten Kontext. (4) Das Konstrukt der *Machbarkeit* beschreibt, inwiefern eine Umsetzung als möglich angesehen wird, z. B. in Abhängigkeit der verfügbaren und benötigten Ressourcen. (5) Die *Implementationskosten* werden in der schulischen Implementationsforschung vergleichsweise selten untersucht oder thematisiert und sollen negative Aspekte des Umfangs, der Komplexität oder des Aufwands (auch ökonomischer Natur) erfassen (Petermann, 2014). (6) Die *Übernahme* der Implementation wird auch als Adoption bezeichnet und beschreibt die Entscheidung, eine Maßnahme langfristig in der Praxis umzusetzen (Petermann, 2014). Die Outcomes (7) *Durchdringung* und (8) *Nachhaltigkeit* sind schließlich neben der Übernahme einem weiter gefassten Verständnis von Implementation zuzuordnen, welches sich auch auf Fragen der Transfer- und Diffusionsforschung bezieht (vgl. Kapitel 3.1). Hierbei wird über die Umsetzung einer Innovation in einer Organisation (z. B. Schule) hinaus die weitere Verbreitung in das praktische Feld durch weitere Akteur*innen fokussiert (Petermann, 2014; Proctor et al., 2011). Konstrukte wie die Übernahme, Nachhaltigkeit und Durchdringung können erst in einem fortgeschrittenen Verlauf der Implementation

erfasst werden, wenn Innovationen in der Praxis eingesetzt werden. Es sei daher angemerkt, dass die genannten Implementationsoutcomes nicht als abgrenzbare Merkmale nebeneinander stehen, sondern theoretisch in einem komplexen Wechselspiel gegenseitiger Beeinflussung anzunehmen sind sowie in einem zeitlichen Verlauf, der sich von ersten Reaktionen auf die Einführung der Innovation bis hin zu einer flächendeckenden Anwendung im praktischen Feld erstreckt (Souvignier et al., 2021). Der Begriff der *Implementationsoutcomes* bildet dieses komplexe Wechselspiel jedoch nur unzureichend ab, da sich die genannten Konstrukte auch gegenseitig bedingen können. In der vorliegenden Arbeit wird daher von Facetten des Implementationserfolgs gesprochen, um zu verdeutlichen, dass es sich um unterschiedliche Teilaspekte eines Prozesses handelt, die auch Voraussetzungen füreinander darstellen können.

In der vorliegenden Arbeit soll die Implementation einer Unterrichtsintervention durch die Schüler*innen untersucht werden. Hierbei steht zunächst die erste Reaktion auf die Einführung eines innovativen didaktischen Konzepts im Fokus. Zentrale Facetten von Implementationserfolg stellen in diesem engeren Verständnis von Implementationsforschung die Wiedergabetreue und Akzeptanz dar, die auch auf Schüler*innenebene konzeptualisiert werden können. Im Folgenden werden diese für die Arbeit zentralen Facetten des Implementationserfolgs genauer beschrieben.

3.3 Wiedergabetreue

Mit dem Konstrukt der Wiedergabetreue wird erfasst, ob eine Intervention im praktischen Setting so umgesetzt wurde, wie theoretisch geplant und vorgegeben (Dane & Schneider, 1998; Petermann, 2014; Proctor et al., 2011). Für die Untersuchung der Wirksamkeit einer Intervention ist die Untersuchung der Wiedergabetreue von zentraler Bedeutung. Nur unter Kontrolle der Wiedergabetreue kann erfasst werden, ob die Effekte einer Intervention der Konzeption der Maßnahme zuzuschreiben sind oder ob durch eine mangelhafte Umsetzung in der Praxis keine Wirksamkeit beobachtet werden konnte (Dane & Schneider, 1998; Dumas et al., 2001; Lendrum & Humphrey, 2012). Potenziell wirksame Maßnahmen, die wegen mangelnder Wiedergabetreue nicht die gewünschten Effekte zeigen, können so fälschlicherweise verworfen und die Interpretation der Ergebnisse damit verzerrt werden (O'Donnell, 2008). Die Beurteilung der Wiedergabetreue kann zudem Hinweise auf die Anwendbarkeit des Konzepts in der Praxis liefern und ist für den erfolgreichen Übergang einer Intervention in die Praxis zentral (Proctor et al., 2011). Auf diese Weise verknüpft die Wiedergabetreue die Perspektiven der Interventions- und Implementationsforschung und stellt eine wichtige Brücke zwischen Wissenschaft und Praxis dar. Verschiedene Literaturreviews zeigen jedoch auf, dass Informationen zur Wiedergabetreue in Wirksamkeitsstudien bislang selten erfasst werden (Dane & Schneider, 1998; Gresham, 2009). In aktuelleren

Studien zeichnet sich jedoch ein zunehmender Trend ab und Informationen zur Implementation werden verstärkt gefordert (Rojas-Andrade & Bahamondes, 2019).

Das Konstrukt der Wiedergabetreue wurde aus unterschiedlichen Perspektiven psychologischer, medizinischer und gesundheitswissenschaftlicher Forschung beeinflusst und geformt, weist über unterschiedliche Disziplinen hinweg jedoch keine einheitliche Definition auf (Sanetti & Kratochwill, 2009). Einigkeit besteht insofern, dass die Wiedergabetreue als multidimensionales und komplexes Konstrukt mit verschiedenen Subdimensionen gefasst wird. Im Folgenden werden die unterschiedlichen Dimensionen der Wiedergabetreue beschrieben sowie anschließend der aktuelle Forschungsstand zur Erfassung der Wiedergabetreue aus Perspektive der Schüler*innen dargestellt.

3.3.1 Dimensionen der Wiedergabetreue

Es bestehen unterschiedliche Konzeptualisierungen der Wiedergabetreue, die sich in Anzahl und Inhalt einzelner Subdimensionen unterscheiden (Dane & Schneider, 1998; Gresham, 2009; Sanetti & Kratochwill, 2009). Eine der am häufigsten rezipierten Konzeptualisierung stellt die Differenzierung nach Dane und Schneider (1998) dar, die auch in der unterrichtlichen Implementationsforschung als Grundlage zur Untersuchung des Konstrukts herangezogen wird (z. B. Adl-Amini, 2018; Brisson et al., 2020; Decristan et al., 2021). Dane und Schneider (1998) tragen fünf Dimensionen der Wiedergabetreue aus bestehender Literatur zusammen. Diese umfassen die Adhärenz (*adherence*), die Dosierung bzw. Menge der erhaltenen Maßnahme (*exposure*), die Qualität der Umsetzung (*quality of delivery*), die Responsivität der Teilnehmenden (*participant responsiveness*) und die Abgrenzbarkeit der Inhalte (*program content differentiation*). Die Dimensionen werden im folgenden Abschnitt definiert.

Die *Adhärenz* beschreibt das Ausmaß, in dem wesentliche Kernelemente einer Intervention von der durchführenden Person wie geplant, bzw. wie in begleitenden Manualen beschrieben, umgesetzt werden (Dane & Schneider, 1998; Sanetti & Kratochwill, 2009). Hierbei wird die Umsetzung einzelner relevanter Elemente quantitativ erfasst, häufig anhand von Beobachtungen und Checklisten (z. B. Crawford et al., 2012; Furtak et al., 2008). In der unterrichtlichen Implementationsforschung kann dabei z. B. die Behandlung zentraler Inhalte während der Intervention oder der Einsatz didaktischer Methoden und struktureller Organisationsformen, wie z. B. kooperatives Lernen, auf ihre präzise Umsetzung überprüft werden (Adl-Amini, 2018; Decristan et al., 2022; Guo et al., 2016). Da jedoch bei den meisten Interventionen nicht alle Elemente die gleiche Relevanz und Wirksamkeit aufweisen, müssen sogenannte Kernelemente (*core components*) identifiziert werden, die als Wirkfaktoren fungieren und im Fokus der Untersuchung der Wiedergabetreue stehen (Abry et al., 2015; Blase & Fixsen, 2013; Century & Cassata, 2016; Furtak et al., 2008; Murrain et al., 2016). Damit sind interventionsspezifische Elemente

gemeint, die umgesetzt werden müssen, um die gewünschten Veränderungen zu erzielen (Blase & Fixsen, 2013). Diese Elemente werden in spezifischen Aktivitäten als „active ingredients“ (Blase & Fixsen, 2013, S. 3) in einer Intervention integriert. Die Identifizierung dieser Kernelemente erfolgt theoretisch oder empirisch und ermöglicht es, Wirkmechanismen besser zu verstehen und Elemente zu identifizieren, die von besonderer Bedeutung für die Wirksamkeit sind, sowie solche, bei denen Adaptionen vorgenommen werden können, ohne die Wirksamkeit der Intervention zu gefährden (Blase & Fixsen, 2013). Für Interventionen im unterrichtlichen Setting ist ein gewisses Maß an Adaptionen notwendig, um eine optimale Passung zu den individuell vorliegenden Lernvoraussetzungen der Klasse oder den organisatorischen Rahmenbedingungen zu ermöglichen (Sanetti & Kratochwill, 2009). Eine Adhärenz von mindestens 70 % der identifizierten Kernelemente wird hierbei als Cut-Off-Wert angesehen, um positive Effekte durch die Intervention erzielen zu können (Adl-Amini, 2018; Reinecker, 2009).

Die *Dosierung* beschreibt als weitere Dimension der Wiedergabetreue das Ausmaß, in dem Teilnehmende einer Intervention und ihren Inhalten ausgesetzt sind (Dane & Schneider, 1998). In unterrichtlichen Implementationsstudien kann dies z. B. die Anzahl, Häufigkeit oder die Länge der Unterrichtsstunden umfassen, in denen die Intervention umgesetzt wird. Die *Qualität der Umsetzung* spiegelt eine qualitative Dimension der Wiedergabetreue wider und bezieht sich nach Dane und Schneider (1998) auf die Art und Weise der Vermittlung der Intervention, die an die Fähigkeiten und Einstellungen der durchführenden Person geknüpft ist. Mit Blick auf unterrichtliche Interventionen können hierbei auch Aspekte der Unterrichtsqualität von Relevanz sein (Adl-Amini, 2018). Ein weiterer qualitativer Aspekt der Wiedergabetreue in der Konzeptualisierung nach Dane und Schneider (1998) ist die *Responsivität der Teilnehmenden*, die das Engagement und die Begeisterung der Teilnehmenden während der Intervention beschreibt. Diese Dimension ist auf der Seite der Empfänger*innen der Intervention anzusiedeln, deren Reaktionen demnach auch Teil der Wiedergabetreue darstellen. Dies kann sich im schulischen Kontext auf die Reaktionen der Schüler*innen beziehen, die an der Intervention teilnehmen. Die *Abgrenzbarkeit der Inhalte* stellt schließlich die inhaltliche Dimension der Wiedergabetreue dar, mit der die Intervention von anderen Methoden (z. B. Unterrichtsformen) differenziert werden kann. Dies wird erreicht, indem z. B. die Daten einer Kontrollgruppe, in der die Elemente der Intervention nicht umgesetzt wurden, zusätzlich analysiert werden (Dane & Schneider, 1998; Gresham, 2009).

Empirische Ergebnisse zur mehrdimensionalen Struktur der Wiedergabetreue liefern bislang gemischte Befunde. Faktoranalytische Untersuchungen zeigen bspw. eindimensionale (Fogarty et al., 2014), zweidimensionale (Humphrey et al., 2018) oder dreidimensionale (Guo et al., 2016) Modellergebnisse. In Interventions- und Implementationsstudien werden jedoch meist nur einzelne Dimensionen der Wiedergabetreue erfasst, wobei sich am häufigsten Angaben zur Adhärenz und Dosierung finden (Dane & Schneider, 1998; Durlak & DuPre, 2008; Sanetti et al.,

2020). Die Adhärenz wird dabei meist mit nur einem Indikator als Summenwert oder Index erfasst (Sanetti et al., 2020). Abry und Kolleg*innen (2015) haben unterschiedliche Methoden zur Erfassung der Wiedergabetreue verglichen und aufgezeigt, dass unterschiedliche Schlussfolgerungen gezogen werden, je nachdem wie Indizes zur Umsetzung von Kernelementen gebildet werden. Dabei wird die gängige Methode zur Erfassung eines Summenwertes durch neuere Methoden, in denen Kernelemente differenziert und einzeln betrachtet werden, übertrifft. Herausforderungen für die Untersuchung der Wiedergabetreue ergeben sich weiterhin auch dadurch, dass Instrumente häufig interventionsspezifisch angepasst bzw. erstellt werden. Dies schränkt die Vergleichbarkeit der Ergebnisse und Interpretation der Schlussfolgerungen ein (Sanetti & Kratochwill, 2009). Für zukünftige Studien wird es wichtig sein, psychometrisch vergleichbare und valide Instrumente zu entwickeln (Durlak & DuPre, 2008).

3.3.2 Untersuchung der Wiedergabetreue durch Schüler*innen

Die Wiedergabetreue wird in unterrichtlichen Implementationsstudien oftmals über die Adhärenz und Umsetzung von Kernelementen operationalisiert und ist häufig primär auf Ebene der durchführenden Lehrkraft angesiedelt. Die Seite der Empfänger*innen einer Intervention, die im Kontext unterrichtlicher Interventionen die Schüler*innen darstellen, wird in der Konzeptualisierung von Dane und Schneider (1998) nur über die Responsivität der Teilnehmenden als qualitative Dimension abgedeckt. In Implementationsstudien wird dieses Konstrukt jedoch selten aufgegriffen (Low et al., 2014). Neuere Konzeptualisierungen haben den Aspekt der Teilnehmenden weiter ausdifferenziert. Nach Jones et al. (2008; zit. nach Sanetti & Kratochwill, 2009) kann die Adhärenz der Umsetzung auch auf Ebene der Empfänger*innen angesiedelt werden als *participant adherence*. Hiermit wird die Umsetzung von Kernelementen der Intervention durch die Schüler*innen beschrieben.

Dies birgt das Potenzial, Herausforderungen der Implementation auch über die Umsetzung der Lehrkräfte hinaus auf Ebene der Schüler*innen zu identifizieren. Vor dem Hintergrund konstruktivistischer Lerntheorien kann davon ausgegangen werden, dass Interventionen auf diese Weise von Schüler*innen mitgestaltet werden. Werden Schüler*innen in die Untersuchung der Adhärenz einbezogen, erfolgt dies allerdings meist in kombinierter Betrachtung mit den Aktivitäten der Lehrkräfte auf Klassenebene (z. B. Adl-Amini, 2018; Crawford et al., 2012). Die Instruktionen der Lehrkraft und Lernaktivitäten der Schüler*innen werden dabei in einem gemeinsamen Index zusammengefasst. Decristan et al. (2021) führen jedoch aus, dass es insbesondere für kooperative Lernarrangements, in denen Schüler*innen Inhalte selbstständig einsetzen müssen, sinnvoll sein kann, die Wiedergabetreue der Lehrkräfte und Schüler*innen getrennt zu erfassen. In der Untersuchung einer außerunterrichtlichen Intervention zum tutoriellen Lernen werden

von den Autor*innen entsprechend Kernelemente auf Schüler*innenebene erfasst, wie z. B. die Anwendung von Lernstrategien beim kooperativen Lernen und das Arbeiten in vorgesehenen Rollen (Decristan et al., 2021).

3.4 Akzeptanz

Die Untersuchung von Akzeptanzurteilen wird in vielen verschiedenen Forschungsfeldern aufgegriffen. Akzeptanzforschung ist dabei keine abgrenzbare wissenschaftliche Disziplin für sich, sondern eine Teildisziplin unterschiedlicher Forschungsperspektiven (Simon & Bernd, 2001). Untersuchungsgegenstand jeglicher Akzeptanzforschungen ist die Reaktion auf eine Integration von Neuerungen in bestehende Strukturen und Praktiken. Aus sozialwissenschaftlicher und soziologischer Perspektive werden bspw. politische Veränderungen, die Einführung von Reformen oder die Veränderung gesellschaftlicher Verhältnisse untersucht (Kleidat, 2011; Lucke, 1995). Eine eher betriebswirtschaftlich eingebettete Akzeptanzforschung beschäftigt sich dagegen mit der Einführung von Technologien und Produkten in Organisationen sowie der Untersuchung von Gründen und Erklärung von Verhalten, warum eine Innovation von Anwender*innen angenommen oder abgelehnt wird (Quiring, 2006; Simon & Bernd, 2001; Venkatesh et al., 2003). Im Rahmen der Implementationsforschung wird Akzeptanz definiert als „die Auffassung der an der Implementation Beteiligten, dass, oder in welchem Maße, eine bestimmte Intervention, Behandlung etc. zufriedenstellend ist“ (Petermann, 2014, S. 123). Diese Definition ist angelehnt an das Konstrukt *Treatment Acceptability*, welches die Bewertung einer Intervention zum Ziel hat und als individuelles Gesamturteil der Zufriedenheit und Angemessenheit ausgelegt wird (Kazdin, 1980).

Im folgenden Kapitel wird Akzeptanz als einstellungsähnliches Konstrukt im Rahmen der Implementationsforschung und als Facette von Implementationserfolg genauer beschrieben. Hierbei werden zunächst zentrale Definitionscharakteristika abgeleitet (vgl. Kapitel 3.4.1) und anschließend der theoretische Rahmen des Konstrukts *Treatment Acceptability* dargestellt (vgl. Kapitel 3.4.2). Die Untersuchung von Akzeptanzurteilen aus Schüler*innenperspektive wird abschließend beschrieben (vgl. Kapitel 3.4.3).

3.4.1 Akzeptanz als einstellungsähnliches Konstrukt

Obwohl theoretische Zugänge zur Untersuchung von Akzeptanzurteilen variieren, können zentrale Definitionscharakteristika perspektivübergreifend abgeleitet werden, die das Verständnis des Konstrukts schärfen und eine Abgrenzung zu anderen Konstrukten ermöglichen. Untersuchungen von Akzeptanzurteilen sind eng mit der Einstellungsforschung verwandt, weswegen Akzeptanz auch als einstellungsähnliches Konstrukt gefasst wird (Lucke, 1995). Akzeptanz wird dabei als

eine „im Prinzip affirmative, jedoch nach Kontext, Situation und Bezugsobjekt verschiedene Einstellung“ (Lucke, 1995, S. 103) definiert und stellt in Anlehnung an eine allgemeine Definition von Einstellungen eine positive Gesamtbewertung eines Einstellungsobjekts dar (Haddock & Maio, 2014). Abgrenzungen zu allgemeineren Einstellungen sind dahingehend möglich, dass Akzeptanz Betroffenheit von Personen voraussetzt (Ullrich, 2008). Lucke (1995) beschreibt Akzeptanz in einem dreifach bezogenen Spannungsfeld. Eine Innovation wird als Akzeptanzobjekt durch die teilnehmenden Personen (die Akzeptanzsubjekte) in einem bestimmten Akzeptanzkontext bewertet. Akzeptanzsubjekte können dadurch näher definiert werden, dass „Entscheidungen und Handlungen Auswirkungen auf Akteure haben, die am Zustandekommen dieser Entscheidungen nicht (maßgeblich) beteiligt waren“ (Ullrich, 2008, S. 22). Diese Betroffenheit zeigt sich vornehmlich in hierarchischen Beziehungen, in denen Akzeptanzsubjekte Entscheidungen anderer Akteur*innen ausgesetzt sind (Ullrich, 2008). Akzeptanz ist demnach als Reaktion auf die Einführung einer Innovation als etwas Nachgelagertes zu verstehen. Einstellungen können sich dagegen auf beliebige Gegenstände, Erfahrungen oder Ideen beziehen, die auch zukünftige oder hypothetische Situationen miteinschließen (Haddock & Maio, 2014).

In der psychologischen Literatur werden Einstellungen klassischerweise als dreidimensionale Konstrukte konzeptualisiert, die sich aus affektiven, kognitiven und verhaltensbezogenen Komponenten zusammensetzen (Rosenberg & Hovland, 1960). Die affektive Dimension umfasst dabei Gefühle und Emotionen, die mit dem Einstellungsobjekt verbunden werden, während die kognitive Dimension Überzeugungen, Gedanken und Wissen über das Einstellungsobjekt widerspiegelt (Haddock & Maio, 2019). Mit der behavioralen Komponente werden Erfahrungen oder Verhaltensweisen gemeint, die sich auf das Einstellungsobjekt beziehen (Haddock & Maio, 2019). Definitionen von Akzeptanz orientieren sich auch an dieser Mehrdimensionalität, indem bspw. in soziologischer und technologischer Akzeptanzforschung eine latente Einstellungsakzeptanz sowie eine beobachtbare Verhaltensakzeptanz differenziert wird (Lucke, 1995; Müller-Böling & Müller, 1986; Venkatesh et al., 2003). Die Einstellungsebene umfasst dabei die affektive und kognitive Dimension und spiegelt eine positive Bewertung der Innovation wider, die als latente Bereitschaft, eine Innovation anzunehmen, verstanden wird. Die Verhaltensakzeptanz stellt dagegen das beobachtbare Verhalten im Sinne einer Annahmemeinung dar, z. B. wenn technologische Innovationen im praktischen Feld von den Nutzer*innen eingesetzt werden. Auch in der Implementationsforschung ist diese Trennung einer latenten Einstellungsebene und manifesten Verhaltensebene zu beobachten, die sich jedoch auf unterschiedliche Facetten von Implementationserfolg beziehen. Die Akzeptanz ist dabei, angelehnt an das *Treatment-Acceptability*-Konstrukt, als Gesamturteil einer Intervention zu verstehen (Kazdin, 1980; Petermann, 2014). Davon abzugrenzen ist die beobachtbare Annahme einer Innovation, die in der Implementationsforschung mit dem Konstrukt der Übernahme (Adoption) beschrieben wird und eine eigene Facette von Implementationserfolg

darstellt. Die Übernahme wird als Entscheidung gefasst, eine Maßnahme langfristig in der Praxis umzusetzen und kann erst in einem fortgeschrittenen Verlauf der Implementation untersucht werden (vgl. Petermann, 2014; Proctor et al., 2011; vgl. Kapitel 3.2.). Das beobachtbare Verhalten, das auf die Einstellungsebene folgt, wird im Rahmen der Implementationsforschung demnach nicht zum Konstrukt der Akzeptanz gezählt und stellt eine wichtige Unterscheidung zu soziologischen oder technologischen Akzeptanzurteilen dar. Da in der vorliegenden Arbeit die erste Reaktion auf eine Innovation durch die Schüler*innen untersucht wird und nicht der Prozess der Verbreitung in das praktische Feld, wird eine behaviorale Kompetente nicht beleuchtet. Stattdessen wird in Anlehnung an das *Treatment-Acceptability*-Konstrukt eine affektive und kognitive Dimension der Akzeptanz fokussiert, was im Folgenden näher ausgeführt wird.

3.4.2 *Treatment Acceptability* zur Untersuchung von Akzeptanzurteilen im Unterricht

Die Untersuchung von Akzeptanzurteilen im Rahmen von Interventionsstudien erfolgt im schulischen Kontext häufig in Anlehnung an die *Treatment-Acceptability*-Forschung (z. B. Eckert et al., 2021; Henninger, 2010; Mautone et al., 2009). *Treatment Acceptability* wurde im Rahmen verhaltenspsychologischer Behandlungen und Interventionen konzeptualisiert (Kazdin, 1980) und dabei als Teil der sozialen Validität einer Intervention untersucht. Soziale Validität wird erreicht, wenn die Ziele, Prozeduren und Effekte einer Intervention als gesellschaftlich angemessen bewertet und akzeptiert werden (Wolf, 1978). Die Anfänge der theoretischen Auseinandersetzung zur sozialen Validität gehen auf Kazdin (1977) und Wolf (1978) und ihre Untersuchungen zu Verhaltensinterventionen im therapeutischen Bereich zurück. Wolf (1978) argumentiert, dass die Effektivität und Effizienz einer Intervention nicht ausreicht, um auch die erwünschte Nutzung durch die Zielgruppe zu garantieren: „If the participants don't like the treatment then they may avoid it, or run away, or complain loudly. And thus, society will be less likely to use our technology” (Wolf, 1978, S. 206). Wolf (1978) bezieht sich hierbei auf die Untersuchung von Interventionsmaßnahmen zum Zweck der Verhaltensänderung. Die Ziele, Prozeduren und Wirkungen müssen nicht nur klinisch sinnvoll und nachweisbar effektiv, sondern auch gesellschaftlich angemessen sein, damit diese in der Praxis implementiert werden und die gewünschten Effekte erzielen (Witt & Elliott, 1985). Zudem besteht die Annahme, dass Behandlungen bzw. Interventionen, die hierbei als *Treatment* bezeichnet werden, mit einer höheren Wahrscheinlichkeit genutzt werden, wenn sie zuvor als angemessen bewertet wurden (Kazdin, 1980; Witt & Elliott, 1985). Die Identifizierung bestimmter Merkmale von *Treatments*, die Nutzer*innen positiv oder negativ bewerten, wird außerdem als wichtige Voraussetzung für den Transfer von Wissenschaft zu Praxis angesehen (Eckert & Hintze, 2000; Schwartz, 1996). Vor diesem Hintergrund wird *Treatment*

Acceptability von Kazdin (1980), aufbauend auf Wolfs (1978) Arbeiten wie folgt definiert:

„Acceptability refers to the judgements about treatment procedures by nonprofessionals, lay persons, clients, and other potential consumers of treatment. Judgements of acceptability are likely to embrace evaluation of whether treatment is appropriate for the problem, whether treatment is fair, reasonable, and intrusive, and whether treatment meets with conventional notions about what treatment should be. In general, acceptability refers to the overall evaluation of the procedures.“ (Kazdin, 1980, S. 259)

Diese Definition ist für die Untersuchung von Akzeptanzurteilen in der Interventionsforschung zentral und wird zur theoretischen Einbettung auch im schulischen Kontext vielfach rezipiert (z.B. Eckert et al., 2021; Hebbecke et al., 2020; Meudt et al., 2020). *Treatment Acceptability* beschreibt dabei, inwiefern eine Person eine Intervention als angemessen, fair, zumutbar oder intrusiv empfindet und ob diese mit gesellschaftlichen und sozialen Normen vereinbar ist. Zusammenfassend wird das Konstrukt häufig auch als „consumers’ judgements about treatment procedures“ (Eckert & Hintze, 2000, S. 124) wiedergegeben und fokussiert damit eine globale Bewertung der Methoden der Intervention durch die (potentiellen) Nutzer*innen. Die normative Bewertung in Kazdins (1980) Definition ist jedoch von besonderer Bedeutung und geht auf die Entstehung des Konstrukts im Kontext der Untersuchung von Interventionsmaßnahmen zur Änderung devianten Verhaltens zurück, von der häufig Personengruppen betroffen waren, die in ihren Fähigkeiten und Möglichkeiten zur Äußerung ihrer Rechte eingeschränkt, den Behandlungen jedoch ausgesetzt waren, wie z.B. psychiatrische Patient*innen, geistig beeinträchtigte oder inhaftierte Personen (Kazdin, 1980). Inzwischen hat sich die Anwendung des Konstrukts in weitere Forschungskontexte ausgeweitet und kommt so auch bei der Untersuchung von Interventionen zur Förderung fachlicher Fähigkeiten für den Regelunterricht (z.B. Eckert et al., 2017; Eckert et al., 2021; Henninger, 2010) und für Kinder mit sonderpädagogischem Förderbedarf (z.B. Mautone et al., 2009) zum Einsatz. Das Verständnis des Konstrukts hat sich dabei verändert von der Erfassung der Angemessenheit der Prozeduren hin zur Fokussierung der Zufriedenheit und Motivation, eine Intervention durchzuführen bzw. daran teilzunehmen.

Die Erfassung von *Treatment Acceptability* beruhte lange Zeit auf der Bewertung hypothetischer Interventionen mit Fallvignetten, die eine sparsame Methode der Erhebung darstellen, jedoch nur eingeschränkt auf reale Settings übertragbar sind (Eckert & Hintze, 2000). Zudem wird angenommen, dass Proband*innen eine Intervention erst dann richtig bewerten können, wenn sie direkte Erfahrung mit dieser gemacht und diese durchgeführt bzw. daran teilgenommen haben (Gresham, 2009). *Treatment Acceptability* wird somit als dynamisches Konstrukt konzeptualisiert, das sich im Verlauf einer Intervention verändern kann (Silva et al., 2019). In aktuellen Studien aus dem schulischen Kontext wird die Akzeptanz

daher erst während der Durchführung einer Intervention (z. B. Hebbecke et al., 2020) oder am Ende der Intervention erhoben (z. B. Eckert et al., 2017; Mautone et al., 2009), um dieser Kritik zu begegnen. Um Ergebnisse zwischen Interventionen valide vergleichbar zu machen, werden zudem zunehmend einheitliche Richtlinien und Erhebungsmethoden zur Erfassung von *Treatment Acceptability* gefordert. Dies gestaltet sich jedoch schwierig, da Instrumente häufig interventionsspezifisch adaptiert werden, um die Inhalte der Intervention zu erfassen. Zudem ist zu beachten, dass Bewertungsurteile einer Intervention auch als allgemeine Einstellungen (z. B. Halliday et al., 2019) erhoben werden und somit zwar Rückmeldungen zu Treatments und Interventionen erfasst werden, theoretische Fundierungen jedoch variieren. In unterrichtsbezogenen Implementationsstudien stellt die Wahrnehmung einer Intervention durch die Schüler*innen aktuell ein zentrales Forschungsdesiderat dar (Schrader et al., 2020; Decristan et al., 2021; Goldenbaum, 2012)

3.4.3 Erfassung der Akzeptanz von Interventionen im Unterricht aus Schüler*innenperspektive

In der vorliegenden Arbeit wird auf Basis der theoretischen Grundlagen die Akzeptanz als einstellungsähnliches Konstrukt definiert, welches in Anlehnung an das Konstrukt *Treatment Acceptability* als Gesamturteil der Intervention zu verstehen ist, das sich aus affektiven und kognitiven Informationen zusammensetzt und auf einer direkten Erfahrung mit der Intervention beruht. Die Akzeptanz ist hierbei von der langfristigen Übernahme (Adoption) als verhaltensbasierte Facette von Implementationserfolg abzugrenzen. Das spezifische Merkmal der Betroffenheit, mit dem Akzeptanz von allgemeineren Einstellungen abgegrenzt werden kann (vgl. Kapitel 3.4.1), zeigt sich schließlich darin, dass Schüler*innen an der Intervention teilnehmen, die Lehrkräfte im Unterricht einsetzen.

Ein häufig genutztes Instrument zur Erfassung von Akzeptanzurteilen bei Lehrkräften, die Interventionen im schulischen Kontext umsetzen, ist das *Intervention Rating Profile* (IRP-20) (Witt & Martens, 1983). Die Skala wurde zur Beurteilung von Verhaltensinterventionen im Unterricht konzipiert und umfasst 20 Items, die neben der Akzeptanz auch die Machbarkeit als weitere Facette von Implementationserfolg fokussiert. Die Skala wurde als *Childrens' Intervention Rating Profile* (CIRP) für Schüler*innen adaptiert, die fiktive Verhaltensinterventionen mittels Fallvignetten bewerten sollen (Witt & Elliott, 1985). Die Items dieser Skala bilden nur die Bewertung der Intervention auf affektiver und kognitiver Ebene ab, z. B. durch die Fairness und Wirksamkeitserwartung. Facetten der Umsetzbarkeit, die im Instrument für Lehrkräfte enthalten sind, werden von den Schüler*innen nicht erfragt. Die CIRP-Skala wurde für die Evaluation von unterrichtlichen Interventionen zur Förderung fachlicher Fähigkeiten von Eckert et al. (2017) als *Kids Intervention Profile* (KIP) adaptiert. Die Skala erfasst die Akzeptanz von

Schüler*innen am Ende einer Intervention zur Förderung der Schreibfähigkeiten und bezieht sich demnach nicht auf eine fiktive Situation, sondern auf die direkte Erfahrung der Schüler*innen mit der Intervention. Die Items und Antwortkategorien sind so konzipiert, dass sie für Schüler*innen ab der dritten Klasse einsetzbar sind. In einer explorativen Faktorenanalyse identifizieren die Autor*innen zwei Faktoren, die eine allgemeine Akzeptanz („General Acceptability“) und die Einschätzung der Fähigkeitsverbesserung („Skill Improvement“) bilden (Eckert et al., 2017, S. 268). Der Faktor zur allgemeinen Akzeptanz enthält fünf Items dazu, wie gut den Schüler*innen einzelne Aspekte der Intervention gefallen bzw. nicht gefallen haben und stellt damit die affektive Dimension der Akzeptanz dar. Der Faktor zur Einschätzung der Fähigkeitsverbesserung umfasst zwei Items zur wahrgenommenen Verbesserung bzw. Verschlechterung der eigenen Schreibfähigkeiten durch die Intervention. Dies wird als kognitive Akzeptanz interpretiert, da hierbei die Überzeugungen und das Wissen, dass die Intervention wirksam und nützlich für den eigenen Lernzuwachs ist, fokussiert werden. Weitere Erfassungen von Akzeptanzurteilen aus Schüler*innenperspektive zeigen ebenfalls diese Trennung in affektive und kognitive Komponenten. Hebbecker und Souvignier (2018) erheben die Akzeptanz einer Leseförderintervention aus Schüler*innenperspektive und differenzieren dabei eine Akzeptanzdimension, die die Freude während der Intervention widerspiegelt und eine Akzeptanzdimension, die sie als Wirksamkeitserwartung beschreiben und die entsprechend der Ausführungen als kognitive Facette von Akzeptanz verstanden werden kann.

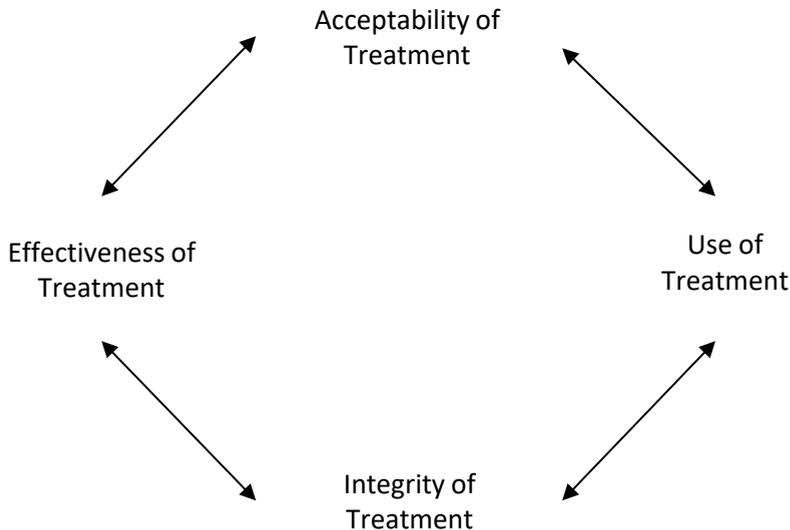
3.5 Facetten des Implementationserfolgs in Unterrichtsinterventionen: Zusammenhänge und Determinanten

Wiedergabetreue und Akzeptanz wurden als zentrale Facetten von Implementationserfolg abgeleitet, mit denen eine erste Reaktion auf die Einführung einer Innovation auch auf Schüler*innenebene untersucht werden kann. Das folgende Kapitel beschreibt Zusammenhänge und Determinanten der Wiedergabetreue und Akzeptanz im Kontext von Unterrichtsinterventionen. Hierzu werden zunächst angenommene Zusammenhänge auf theoretischer Ebene dargestellt, die aus Modellen der Implementations- und Unterrichtsforschung abgeleitet werden können. Anschließend werden der Forschungsstand zu Zusammenhängen der Facetten untereinander sowie zum Lernerfolg beschrieben und Einflussfaktoren auf Implementationsprozesse abgeleitet.

3.5.1 Konzeptuelle Modelle der Implementations- und Unterrichtsforschung als theoretischer Rahmen

Facetten des Implementationserfolgs stehen in komplexem Verhältnis zueinander. Ein konzeptuelles Modell nach Witt und Elliott (1985), das im Rahmen der *Treatment-Acceptability*-Forschung entwickelt wurde, zeigt das Wechselspiel zwischen Konstrukten, die für eine erstmalige Umsetzung einer Innovation in der Praxis relevant sind. Das Modell dient als Heuristik zur Untersuchung der Zusammenhänge von Konstrukten im Implementationsprozess (Witt & Elliott, 1985; vgl. Abbildung 1). Hierbei werden Akzeptanz und Wiedergabetreue sowie die Nutzung und Wirksamkeit einer Intervention in ein Wirkungsfüge gebracht. Die Beziehungen zwischen den Konstrukten sind dabei als aufeinanderfolgend sowie reziprok anzunehmen (Witt & Elliott, 1985).

Abbildung 1: *Model of Treatment Acceptability* nach Witt & Elliott (1985)



Die Akzeptanz (*Acceptability of Treatment*) kann hierbei den Ausgangspunkt darstellen und die Auswahl und tatsächliche Anwendung (*Use of Treatment*) einer Intervention beeinflussen (Witt & Elliott, 1985). Die Anwendung ist hierbei nicht als langfristige Übernahme zu verstehen, sondern führt auf die Anfänge der *Treatment-Acceptability*-Forschung zurück, in der die Akzeptanz von Interventionen hypothetisch erfasst und eine Intervention auch vor ihrer erstmaligen Anwendung bewertet wurde. Nach Witt und Elliott (1985) steigt die Wahrscheinlichkeit, dass eine Intervention eingesetzt wird, wenn dieses zuvor als akzeptabel eingestuft wurde. Die Wiedergabetreue (*Integrity of Treatment*) kann weiterhin den Zusammenhang zwischen der Anwendung und Wirksamkeit (*Effectiveness of Treatment*) einer Intervention vermitteln. Demnach steigt die Wahrscheinlichkeit, dass sich eine

Intervention als wirksam erweist und wünschenswerte Veränderungen verursacht, wenn diese mit hoher Wiedergabetreue implementiert wurde. Im Modell werden die Zusammenhänge reziprok dargestellt. Entsprechend ist es auch möglich, dass eine wirksame Intervention zu mehr Akzeptanz führt, da ein effektives Treatment mit höherer Wahrscheinlichkeit positiv bewertet wird und entsprechend auch mit höherer Wiedergabetreue implementiert wird. Ebenfalls kann eine hohe Wiedergabetreue eine höhere Akzeptanz einer Intervention bedingen (Witt & Elliott, 1985). Je nach Zeitpunkt der Untersuchung und Umsetzung kann Akzeptanz demnach als Bedingung oder Ergebnis von Implementationsprozessen gefasst werden.

Das Modell stellt einen Ausgangspunkt für zahlreiche Untersuchungen zu Zusammenhängen von Akzeptanz und Wiedergabetreue dar. Während in verhaltenspsychologischen Studien und Untersuchungen zur Gesundheitsförderung in der Schule ein positiver Zusammenhang zwischen dem Implementationserfolg und wünschenswerten Veränderungen der Teilnehmer*innen etabliert ist (Dane & Schneider, 1998; Durlak & DuPre, 2008), gibt es wenige Untersuchungen im Kontext unterrichtlicher Interventionen. Die Wirksamkeit von Interventionen im Unterricht wird meist an den Lernerträgen und der Verbesserung der Leistungen von Schüler*innen gemessen. Im Gegensatz zu Interventionen in kontrollierten Settings ist die Untersuchung von Implementationserfolg bei Unterrichtsinterventionen zusätzlich in das Wirkungsgefüge des Unterrichts eingebettet und muss vor diesem Hintergrund analysiert werden.

Wirkungszusammenhänge in Lehr-Lernsettings können mithilfe von Angebot-Nutzungsmodellen beschrieben werden, auf die auch zur Untersuchung von Einflussfaktoren von Implementationsprozessen im unterrichtlichen Kontext Bezug genommen wird (Schrader et al., 2020). Angebot-Nutzungsmodelle bieten einen konzeptuellen Rahmen zur Wirkweise des Unterrichts, indem das Zusammenspiel zentraler Merkmale und Faktoren des Unterrichtsgeschehens systematisiert und auf unterschiedlichen Ebenen schematisch abgebildet wird (Kohler & Wacker, 2013; Seidel, 2014; Vieluf et al., 2020). Interdisziplinär wird ihnen eine „verständigungsstiftende Funktion“ (Kohler & Wacker, 2013, S. 242) zugeschrieben, da sie das komplexe Zusammenspiel unterschiedlicher Determinanten übersichtlich bündeln und als „Metamodell[e]“ (Lipowsky, 2015, S. 75) Forschungserkenntnisse über verschiedene Facetten des Unterrichtsgeschehens in einen Wirkungszusammenhang bringen. Hierbei finden schulische sowie außerschulische Merkmale auf unterschiedlichen Ebenen Beachtung und werden miteinander in Beziehung gesetzt. In Angebot-Nutzungsmodellen ist die konstruktivistische Grundannahme des Zusammenspiels von Angebot und Nutzung zentral, das die Wechselseitigkeit und Ko-Konstruktion zur Erzeugung von Wissen im Unterricht in den Vordergrund stellt. Demnach werden positive Lernergebnisse theoretisch durch eine optimale Nutzung eines bestmöglichen Angebots erklärt (Fend, 1998; Kohler & Wacker, 2013).

Das Angebot wird häufig auf Seiten der Lehrkräfte konzeptualisiert, die Lerngelegenheiten für Schüler*innen gestalten (Brühwiler & Blatchford, 2011; Lipowsky, 2015). Die Gestaltung des Angebots wird dabei durch individuelle Merkmale der Lehrkraft beeinflusst (Helmke, 2012; Seidel, 2014). Wird allerdings das Angebot mit einer stärkeren Fokussierung auf die wechselseitige Ko-Konstruktion durch alle beteiligten Akteur*innen – Lehrkräfte sowie Schüler*innen – konzeptualisiert, stehen auch Schüler*innen in der Verantwortung zur Mitgestaltung des Angebots, bspw. im Rahmen kooperativer Lernarrangements. Vor diesem Hintergrund definieren Vieluf et al. (2020, S. 75) Unterricht im Rahmen des „integrierten Angebot-Nutzungsmodells“ als „fachspezifische Ko-Konstruktion“, die aus der Interaktion von Schüler*innen, Lehrkräften und einem Unterrichtsgegenstand hervorgeht. Das Angebot wird auf diese Weise gemeinsam gestaltet. Vor dem Hintergrund kooperativer Lernmethoden ist dies insbesondere relevant und eine Fokussierung auf das Handeln der Lehrkraft nicht ausreichend, da diese in kooperativen Lernphasen in den Hintergrund rückt und die Schüler*innen verstärkt Verantwortung für die Lernprozesse übernehmen (Pauli & Reusser, 2000).

Zwischen Angebot und Nutzung wird eine reziproke Beziehung angenommen, die über Wahrnehmungs- und Interpretationsprozesse vermittelt wird (Vieluf et al., 2020). Beobachtbare Arbeitsweisen und Aktivitäten, wie z. B. die aktive Lernzeit, Mitarbeit im Unterricht oder an Gesprächen werden als äußere Nutzung ebenso auf Ebene des Angebots verortet, da sie zur Gestaltung des Angebots für andere Schüler*innen beitragen (Vieluf et al., 2020). Vieluf et al. (2020, S. 75) definieren Nutzung im integrierten Angebot-Nutzungsmodell als „komplexen Prozess, in dem sich Kognitionen, emotionales und motivationales Erleben ständig gegenseitig beeinflussen“. Die inneren, mental ablaufenden Prozesse kognitiver Informationsverarbeitung und Elaboration werden durch motivational-affektive Prozesse gesteuert und vermittelt sowie durch individuelle Lernvoraussetzungen beeinflusst (Helmke, 2012; Kunter & Trautwein, 2013; Seidel, 2014). Ein Lernangebot ist demnach nicht für alle Schüler*innen gleich wirksam, sondern wird durch individuelle Verarbeitungsprozesse mediiert, die sich auch in Abhängigkeit individueller Lernvoraussetzungen ergeben. Die Schüler*innen nehmen dabei eine zentrale Rolle für den Erfolg von Unterricht ein. Für distale Schüler*innenmerkmale, wie das Geschlecht, der Migrationshintergrund oder sozioökonomische Status, wird angenommen, dass Zusammenhänge zum Lernerfolg über proximale Merkmale vermittelt werden. Diese weisen eine größere Nähe zum Lerngegenstand auf, wie kognitive und affektiv-motivationale Variablen, die als besonders starke Prädiktoren für Leistungsergebnisse angenommen werden (Helmke & Weinert, 1997). Kognitive Voraussetzungen beeinflussen kognitive Verarbeitungsprozesse, wie eine schnellere Verarbeitung und bessere Anknüpfung an bereits bekanntes Vorwissen (Chi & Wylie, 2014). Motivationale und affektive Voraussetzungen der Schüler*innen beeinflussen in Interaktion mit dem Unterrichtsangebot das motivationale und emotionale Erleben als Teil von Nutzungsprozessen (Vieluf et al., 2020). In der pädagogisch-psychologischen Lehr-Lernforschung spielt das affektive

Erleben eine bedeutsame Rolle und motivationale Voraussetzungen der Lernenden zählen zu den bedeutsamsten Faktoren erfolgreicher Lernprozesse (Spinath, 2011). In Angebot-Nutzungsmodellen findet sich zudem schon früh eine multi-kriteriale Betrachtung von Lernergebnissen in fachlichen und überfachlichen Zielen (Klime et al., 2006; Seidel, 2014). So stellt auch die Entwicklung und Förderung von Lernmotivation, neben fachlichem Wissenszuwachs, ein zentrales Bildungsziel dar (Hasselhorn & Gold, 2017; Spinath, 2011).

Zur Systematisierung von Ergebnissen der Implementationsforschung im Bildungsbereich übertragen Schrader et al. (2020) Facetten des Implementationserfolgs in die Heuristik der Angebot-Nutzungsmodelle. Auf dieser Basis sollen das Zusammenspiel unterschiedlicher Merkmale im Implementationsprozess sowie Einflussfaktoren, die auf Implementationsfacetten wirken, dargestellt und strukturiert werden. Facetten des Implementationserfolgs werden damit in das Wirkungsgefüge des Unterrichts eingebettet und Implementations- und Unterrichts-forschung miteinander verknüpft. Im Modell nach Schrader et al. (2020) werden Lehrkräfte als Nutzende und Implementierende beschrieben. Vor dem Hintergrund konstruktivistischer Lerntheorien kann diese doppelte Rolle für Schüler*innen ebenfalls angenommen werden, da sie die Angebotsstrukturen im Unterricht nicht nur passiv nutzen, sondern aktiv mitgestalten. Facetten der Wiedergabetreue als Gestaltung des Lernangebots können demnach sowohl auf Ebene der Lehrkraft als auch der Schüler*innen zugeordnet werden. Hierbei ist allerdings zu beachten, dass eine trennscharfe Einteilung der Facetten des Implementationserfolgs in die Konstrukte der Angebot-Nutzungsmodelle durch die Reziprozität und ständige wechselseitige Beeinflussung von Angebot und Nutzung im integrierten Angebot-Nutzungsmodell (Vieluf et al., 2020) kaum möglich ist. Mit den gängigen Methoden der Implementationsforschung zur Erfassung der Wiedergabetreue (z. B. Checklisten) werden nur beobachtbare Prozesse abgebildet. Die für den Lernprozess zentralen inneren Nutzungsprozesse können nicht beobachtet werden. Die Erfassung der Wiedergabetreue auf Schüler*innenebene kann jedoch als Indikator für die Gestaltung des Lernangebots dienen, da abzuleiten ist, wie und in welchem Ausmaß eine Intervention umgesetzt wurde. Anknüpfend an das Modell nach Witt und Elliott (1985) kann auch vor dem Hintergrund der Angebot-Nutzungsmodelle angenommen werden, dass eine Intervention eine höhere Wirksamkeit zeigt, wenn sie mit hoher Wiedergabetreue umgesetzt wurde, da die zu überprüfenden Kernelemente zentrale Wirkfaktoren darstellen. Die Berücksichtigung individueller Lernvoraussetzungen und innerer Nutzungsprozesse, die den Zusammenhang zwischen Angebot und Lernerfolg mediierten, sind jedoch im Kontext von Unterrichtsinterventionen zusätzlich zu berücksichtigen. Akzeptanzurteile können dabei als Gesamtbewertung der Intervention Anhaltspunkte für evaluierende affektive und kognitive Wahrnehmungsprozesse der Schüler*innen sein (vgl. Kapitel 3.4). Es wurde dargestellt, dass Angebot-Nutzungsmodelle zur Präzisierung der angenommenen Zusammenhänge der Implementationsforschung zwischen zentralen

Facetten des Implementationserfolgs unter Berücksichtigung der Wirkweise des Unterrichts herangezogen werden können.

3.5.2 Zusammenhänge zwischen Wiedergabetreue, Akzeptanz und Schüler*innenoutcomes

Es gibt bislang wenige Studien, die Facetten des Implementationserfolgs von unterrichtlichen Interventionen untersuchen. In verhaltenspsychologischen Untersuchungen ist dagegen ein positiver Zusammenhang zwischen dem Implementationserfolg und wünschenswerten Veränderungen der Teilnehmer*innen bereits etabliert (Dane & Schneider, 1998; Durlak & DuPre, 2008). Im Folgenden wird die aktuelle Befundlage zu Zusammenhängen von Wiedergabetreue, Akzeptanz und Schüler*innenoutcomes im Kontext *unterrichtlicher Interventionen* beschrieben.

Auf Ebene der Lehrkraft zeigt sich in einigen Studien ein positiver Zusammenhang zwischen selbstberichteter Akzeptanz und selbstberichteter Wiedergabetreue. Meudt et al. (2020) beschreiben in einer korrelativen Untersuchung, dass die selbstberichtete Akzeptanz und selbstberichtete Wiedergabetreue von Leseförderkonzepten in der Sekundarstufe I bei elf Lehrkräften positiv zusammenhängt. Auch in einer Untersuchung von Hebbecker et al. (2020), in der 73 Grundschullehrkräfte formatives Assessment im Leseunterricht einsetzen, zeigt sich ein positiver Zusammenhang zwischen der selbstberichteten Akzeptanz zu Beginn und während der Intervention und der selbstberichteten Nutzung von Feedbackmaterialien am Ende der Intervention. Bei Tanol (2010) konnte weiterhin auch mit der objektiven Erfassung der Wiedergabetreue über Beobachtungen ein positiver Zusammenhang zu den Akzeptanzurteilen von 44 Lehrkräften, die eine Intervention zur Verbesserung des Classroom-Managements einsetzen, gezeigt werden (Tanol, 2010).

In Untersuchungen, die die Bedeutung der Wiedergabetreue für die Entwicklung von Schüler*innenoutcomes untersucht haben, sind Befunde gemischt. Auch hier variiert die Erfassung der Wiedergabetreue zwischen Selbstberichten und objektiven Verfahren der Beobachtung. Kuhn (2018) untersucht den Einfluss verschiedener Facetten der Wiedergabetreue von Lehrkräften im Rahmen einer Intervention zur Förderung der Lernmotivation in achten Klassen auf die Lernfreude und Leistungsmaße zu Lese- und Schreibkompetenzen. In einer kleinen Stichprobe von neun Lehrkräften zeigen sich positive Zusammenhänge der selbstberichteten Adhärenz und Dosierung der Intervention zu den fokussierten Schüler*innenoutcomes. Die selbstberichtete Qualität der Umsetzung und Responsivität der Teilnehmenden hängt mit diesen dagegen nicht zusammen. Crawford et al. (2012) erfassen die Wiedergabetreue einer computerbasierten Intervention zur Förderung mathematischer Fähigkeiten bei 485 Schüler*innen in der Mittelstufe über die Beobachtung der Aktivitäten der Lehrkräfte und der Schüler*innen in einem gemeinsamen Index. In dieser Untersuchung sagen die Adhärenz und

Dosierung der Intervention die Entwicklung der mathematischen Fähigkeiten auf Schüler*innenebene positiv vorher. Dagegen bleiben angenommene Zusammenhänge in einer Untersuchung von Adl-Amini (2018) aus. In einer Peer-Tutoring-Intervention im Sachunterricht der Grundschule konnte kein Zusammenhang zwischen der Adhärenz auf Klassenebene und der Entwicklung des konzeptuellen Verständnisses der Schüler*innen bestätigt werden. Die Wiedergabetreue wurde dabei in 14 Klassen durch Beobachtung spezifischer Kernelemente erfasst, in denen sowohl Aktivitäten der Lehrkräfte als auch der Schüler*innen in einem gemeinsamen Index zusammengefasst wurden.

Für den Bereich der Leseförderung zeigt Henniger (2010) in einer quantitativen Untersuchung, dass die beobachtete Adhärenz der Umsetzung einer Leseintervention durch die Lehrkraft positiv mit ihren Akzeptanzurteilen am Ende der Intervention zusammenhängt. Zusätzlich zeigt sich in der Untersuchung ein positiver Zusammenhang zwischen der Adhärenz durch die Lehrkräfte und der Wirksamkeit der Intervention, gemessen an der Leseflüssigkeit der Schüler*innen. Eine differenzierte Erhebung der Wiedergabetreue findet sich auch bei Fogarty et al. (2014), die in einer Leseintervention in 6. und 8. Klassen mit 14 Lehrkräften und 859 Schüler*innen alle fünf Dimensionen der Wiedergabetreue über Beobachtungen, Audioaufnahmen und Befragungen der Lehrkräfte erfassen. Den Ergebnissen einer konfirmatorischen Faktorenanalyse folgend werden die Dimensionen jedoch zu einem gemeinsamen Faktor zusammengefasst. Der Faktor hängt in der Untersuchung signifikant positiv mit dem Leseverstehen und den Schreibfähigkeiten der Schüler*innen am Ende der Intervention zusammen.

Es konnte nur eine Studie identifiziert werden, in der Akzeptanz oder Wiedergabetreue im unterrichtlichen Kontext im Zusammenhang mit der Wirksamkeit einer Intervention aus Perspektive der Schüler*innen erfasst wurde. Eckert et al. (2021) untersuchen in einer Intervention zur Förderung der Schreibfähigkeiten die Akzeptanz von 70 Schüler*innen in der Grundschule. Dabei zeigt sich eine signifikant positive Korrelation der affektiven Akzeptanz am Ende der Intervention mit der Verbesserung der Schreibfähigkeiten durch die Intervention. Der Zusammenhang zur kognitiven Akzeptanz war dabei nicht signifikant.

Die Darstellung der Studien macht deutlich, dass die Befundlage im Kontext unterrichtlicher Interventionen insgesamt heterogen ist, wobei die Vergleichbarkeit der Studien eingeschränkt ist. Meist werden unterschiedliche Facetten der Wiedergabetreue erhoben, die zwischen Selbstberichten und objektiven Verfahren variieren. Zudem müssen die Erhebungsinstrumente und Outcomevariablen jeweils spezifisch für den Kontext und thematischen Rahmen der Intervention angepasst werden. Da Implementationsstudien häufig mit großem Aufwand einhergehen, finden sich auch viele Untersuchungen, die kleine Stichproben und nur eine geringe Anzahl an teilnehmenden Klassen aufweisen (z. B. Adl-Amini, 2018; Henninger, 2010; Kuhn, 2018; Tanol, 2010), wodurch die statistische Power eingeschränkt ist. Zudem gibt es Unterschiede in den Erhebungszeitpunkten. Insbesondere die Erfassung der Akzeptanz variiert zwischen der Erhebung zu Beginn oder nach Ab-

schluss der Intervention. Daraus ergeben sich auch unterschiedliche Wirkrichtungen, die im Modell nach Witt und Elliott (1985) bereits als reziprok angenommen werden. Die Akzeptanz kann demnach sowohl Voraussetzung als auch Ergebnis einer hohen Wiedergabetreue darstellen.

3.5.3 Determinanten für Implementationserfolg im Unterricht

In verschiedenen Perspektiven der Implementationsforschung haben sich Merkmale auf unterschiedlichen Ebenen als einflussreich für Implementationserfolg erwiesen, wie Merkmale der Innovation, individuelle Merkmale der handelnden Akteur*innen oder Merkmale der Organisation und des Kontextes (Fullan, 2007; Goldenbaum, 2012; Rogers, 2003; Sanetti & Kratochwill, 2009). Merkmale der Innovation und Merkmale der handelnden Akteur*innen stehen dabei in engem Zusammenhang, da die subjektive Wahrnehmung innovationsspezifischer Merkmale von individuellen Merkmalen beeinflusst wird (Hartmann et al., 2020; Rogers, 2003). Aus Befunden der Implementationsforschung ist bereits bekannt, dass die Passung zum wahrgenommenen Bedarf, die Klarheit der Innovation sowie die Komplexität und die Umsetzbarkeit relevante Merkmale der Innovation darstellen, die die Umsetzung beeinflussen (Fullan, 2007; Henninger, 2010; Perepletchikova & Kazdin, 2005; Rogers, 2003). Ein hohes Maß an Kosten, benötigten Materialien und (zeitlichen) Ressourcen kann dagegen hemmend auf eine hohe Wiedergabetreue wirken (Perepletchikova & Kazdin, 2005). Insgesamt wird die Passung zwischen den individuellen Merkmalen der implementierenden Person sowie implementationsspezifischen Merkmalen betont und entsprechend die Notwendigkeit einer adressatengerechten Gestaltung von Implementationsmaßnahmen abgeleitet (Hartmann et al., 2020; Souvignier, 2020). Dabei ist es insbesondere förderlich, wenn Möglichkeiten zur Adaption an individuelle Rahmenbedingungen gegeben sind (Gräsel & Parchmann, 2004; Schrader et al., 2020). Im unterrichtlichen Kontext ist von Bedeutung, dass sich die Umsetzung der Innovation in die gegebenen Unterrichtsabläufe integrieren lässt und an die Fähigkeiten der Lehrkräfte sowie Schüler*innen in angemessenem Maße anknüpft, damit eine bestmögliche Passung gegeben ist (Gräsel & Parchmann, 2004). Förderliche Einflussfaktoren auf Ebene der Lehrkräfte beziehen sich auf fachliche Kompetenzen und Qualifikationen, die für den Implementationsgegenstand relevant sind, sowie positive Einstellungen und Selbstwirksamkeitserwartungen bezüglich der Umsetzung der Intervention (vgl. Goldenbaum, 2012).

Merkmale der Innovation können in der Angebot-Nutzungs-Heuristik dem Implementationsgegenstand und Lernangebot zugeordnet werden, die sich auf die Wahrnehmung und Nutzungsprozesse auswirken (Schrader et al., 2020). Diese werden im Unterricht wiederum beeinflusst durch distale sowie proximale individuelle Lernvoraussetzungen der Schüler*innen. Kognitive und affektiv-motivationale Merkmale beeinflussen dabei maßgeblich Prozesse der Informa-

tionsverarbeitung, die sich auf den Lernerfolg auswirken (vgl. Kapitel 3.5.1). Einige wenige Untersuchungsergebnisse zeigen bereits, dass distale und proximale Schüler*innenmerkmale mit dem Implementationserfolg im Unterricht zusammenhängen. In einer Systematisierung von Forschungsergebnissen nach Schrader et al. (2020) werden das Geschlecht, positive lerngegenstandsbezogene Motivation und die Passung zu Lernbedürfnissen beschrieben, die als förderliche Faktoren für Wahrnehmungs- und Nutzungsprozesse sowie die Wirksamkeit der Interventionen identifiziert wurden. Eine Studie von Brisson et al. (2020) kann zusätzlich ergänzt werden, in der die Responsivität von Schüler*innen in einer Unterrichtsintervention zur Steigerung der Motivation im Mathematikunterricht durch das Geschlecht der Schüler*innen, die lerngegenstandsbezogene Motivation und die Gewissenhaftigkeit der Schüler*innen positiv vorhergesagt wird. Mädchen zeigen in der Untersuchung ein höheres Maß an Responsivität, das über die Bearbeitung schriftlicher Aufgaben zu den Inhalten der Intervention erfasst wurde.

Auch im Rahmen der *Treatment-Acceptability*-Forschung wurden Merkmale der teilnehmenden Schüler*innen in verhaltenspsychologischen Untersuchungen als einflussreich für Facetten des Implementationserfolgs identifiziert. Zusammenhänge zur Wiedergabetreue der Lehrkraft zeigen sich z. B. mit dem affektiven Erleben der Schüler*innen und der Schwere des zu behandelnden Problems (Perepletchikova & Kazdin, 2005). Als Determinanten von Akzeptanzurteilen der durchführenden Fach- und Lehrkräfte haben sich ebenfalls die benötigte Zeit, Kosten und mögliche Nebenwirkungen der Implementation als einflussreich erwiesen (Allinder & Oats, 1997). Aus Modellen der technologischen Akzeptanzforschung, die an psychologische Handlungstheorien angelehnt sind, geht zudem hervor, dass individuelle Merkmale wichtige Determinanten von Akzeptanzurteilen darstellen. Die Einstellungen, Erfahrungen und Kompetenzen sowie Kosten-Nutzen-Abwägungen stellen hierbei Prädiktoren für die Akzeptanz von Technologien dar (z. B. Venkatesh et al., 2003).

Zu Kontextmerkmalen auf unterschiedlichen Ebenen, wie die Schule und ressourcenbezogene oder kollegiale Unterstützungsstrukturen, liegen schließlich eine Vielzahl von Ergebnissen vor (Schrader et al., 2020). In Implementationsstudien aus dem Bildungsbereich gilt auf Ebene der Organisation (Schule) insbesondere die Schulleitung als einflussreicher Faktor (Goldenbaum, 2012). Die Schulleitung kann die Bedingungen für erfolgreiche Veränderungen schaffen, indem zeitliche oder materielle Ressourcen bereitgestellt sowie Lehrkräfte ermuntert, unterstützt und fortgebildet werden, um als handelnde Akteur*innen den Implementationsprozess wiederum selbst positiv beeinflussen zu können (Fullan, 2007; Goldenbaum, 2012; Gräsel & Parchmann, 2004; Hunneshagen, 2005). Weiterhin erstreckt sich der Einfluss der Schulleitung auch auf Eltern und das schulische Umfeld, die zur Umsetzung von Innovationen ebenfalls einen bedeutsamen Beitrag leisten können (Gräsel, 2010), bspw. durch eine positive Haltung oder die Erteilung von Einwilligungserklärungen für eine wissenschaftliche Begleitung. Die Akzeptanz der Schulleitung gegenüber der zu implementierenden Maßnahme und die Identi-

fikation mit den Inhalten stellen weitere relevante Indikatoren für den Erfolg von Reformprojekten dar (Böse et al., 2019). Auch kooperative Strukturen im Kollegium innerhalb der Schule sind für den Implementationsprozess bedeutsam und können die Umsetzung von Maßnahmen unterstützen (Gräsel, 2010). Auf Ebene der Organisation sind Unterstützungssysteme und Normen der Organisation angesiedelt und auf der obersten Ebene externe Kontextfaktoren, die der Mesoebene des Systems zuzuschreiben sind, wie z. B. politische Rahmenbedingungen und bürokratische Verfahren (Sanetti & Kratochwill, 2009).

4 Fragestellungen

Im schulischen Kontext orientieren sich Konzeptualisierungen von Implementationserfolg maßgeblich an der Perspektive der Lehrkräfte, die eine Innovation z. B. im unterrichtlichen Rahmen umsetzen. Vor dem Hintergrund konstruktivistischer Lerntheorien kann jedoch angenommen werden, dass das Unterrichtsangebot auch von Schüler*innen mitgestaltet wird und die Implementation einer Innovation in den Unterricht nicht mit dem Handeln der Lehrkraft abgeschlossen ist (z. B. Helmke, 2012; Schrader et al., 2020; Vieluf et al., 2020). Die Berücksichtigung der Prozesse auf Schüler*innenebene ist daher auch in der Implementationsforschung von Bedeutung. Bislang werden diese in Implementationsstudien allerdings kaum eingebunden und insbesondere die Bedeutung der Wahrnehmung einer Intervention durch die Schüler*innen stellt ein aktuelles Forschungsdesiderat unterrichtsbezogener Implementationsstudien dar (Decristan et al., 2021; Goldenbaum, 2012; Schrader et al., 2020). Bei der Integration von Mehrsprachigkeit in den Unterricht ist die Perspektive der Schüler*innen jedoch besonders relevant, da mehrsprachigkeitssensible Konzepte bislang wenig untersucht sind und es kaum Forschung dazu gibt, wie Schüler*innen diese im Unterricht annehmen (Hilbe & Kutzelmann, 2017). Gleichzeitig weisen Lehrkräfte Ängste und Unsicherheiten in Bezug auf die Integration von Mehrsprachigkeit auf, die eine Etablierung mehrsprachigkeitssensibler Konzepte erschweren. Die vorliegende Dissertation knüpft an diese Forschungslücken an und untersucht Facetten des Implementationserfolgs einer Intervention zum mehrsprachigkeitssensiblen Reziproken Lehren aus Schüler*innenperspektive.

Das mehrsprachigkeitssensible Reziproke Lehren hat die Förderung der Deutsch-Lesekompetenz zum Ziel (Decristan et al., 2022; vgl. Kapitel 2.4). Kern der Methode ist das Einüben von Lese- und Lernstrategien im kooperativen Setting und mit reziproken Rollen. In einer Unterrichtseinheit mit zwölf Unterrichtsstunden von je 45 Minuten bearbeiten Schüler*innen Lesetexte und wenden die Lesestrategien „Fragen stellen“ und „Zusammenfassen“ an. Die Strategieverwendung wird durch sogenannte Promptkarten unterstützt, auf denen die wichtigsten Schritte zum Ausführen der Strategien beschrieben sind. Die Promptkarten dienen in der Gruppenarbeit dazu, die Strategien strukturiert einzuüben und anhand der genannten Kriterien die Ergebnisse gemeinsam zu diskutieren und zu prüfen. Die Promptkarten sind beim mehrsprachigkeitssensiblen Reziproken Lehren bilingual gestaltet, sodass die Strategien auf Deutsch und auf jeweils einer der Familiensprachen der Schüler*innen angegeben sind. Neben diesem mehrsprachigkeitssensiblen Impuls wird die Intervention zusätzlich angereichert durch ein wertschätzendes Klima für Mehrsprachigkeit und die Bildung von Sprachpartnerschaften innerhalb der Kleingruppen, um auch mehrsprachige Interaktionen zu ermöglichen (vgl. Kapitel 5.1.1).

Als Facetten von Implementationserfolg werden in der vorliegenden Arbeit die Wiedergabetreue und Akzeptanz fokussiert, die sich über unterschiedliche For-

schungsperspektiven hinweg als wichtige Indikatoren für den Implementationserfolg einer Intervention erwiesen haben und auch auf Ebene der Schüler*innen relevant sind (vgl. Kapitel 3). Für die zu untersuchende Intervention stellt das Einüben der Lesestrategien ein Kernelement dar, das durch die Promptkarten in der kooperativen Arbeit unterstützt wird. Die Promptkarten stellen ein besonders innovatives Element des mehrsprachigkeitssensiblen Reziproken Lehrens dar, mit dem das Üben der Lesestrategien mit dem Element der Mehrsprachigkeit verknüpft wird. Mithilfe der Promptkarten soll der für das Reziproke Lehren besonders wichtige Dialog zwischen den Schüler*innen und das Auftreten lernförderlicher kognitiver Aktivitäten in der Gruppenarbeit für eine elaborierte Verarbeitung (z. B. durch Rückfragen, Erklären) gefördert werden (vgl. Kapitel 2.3.2). Die Nutzung der Promptkarten wird daher als „active ingredient“ (Blase & Fixsen, 2013, S. 1; vgl. Kapitel 3.3.1) der Intervention angenommen und die Wiedergabetreue der Schüler*innen als *participant adherence* (Jones et al., 2008, zit. nach Sanetti & Kratochwill, 2009) untersucht. Als weitere Facette des Implementationserfolgs wird in Anlehnung an das *Treatment-Acceptability*-Konstrukt die Akzeptanz der Schüler*innen als affektive und kognitive Gesamtbewertung am Ende der Intervention erfasst (vgl. Kapitel 3.4).

Zunächst soll aus einem breiteren Blick der Implementationsforschung untersucht werden, wie die genannten Facetten des Implementationserfolgs mit dem Lernerfolg der Schüler*innen in Zusammenhang stehen (Fragestellung 1) und wie Unterschiede im Implementationserfolg auf individueller Ebene durch Lernvoraussetzungen der Schüler*innen erklärt werden können (Fragestellung 2). Dabei findet die Mehrsprachigkeit als individuelles Merkmal besondere Betrachtung. Anschließend werden aus einer differenziellen Perspektive die Implementation des mehrsprachigen Unterrichtsmaterials fokussiert und auch hier Zusammenhänge zum Lernerfolg (Fragestellung 3) und die Bedeutung individueller Lernvoraussetzungen mehrsprachiger Schüler*innen für diese untersucht (Fragestellung 4). Im Folgenden werden die Fragestellungen auf Basis des theoretischen Hintergrunds und der beschriebenen Forschungsergebnisse abgeleitet.

4.1 Fragestellung 1: Sagen die Facetten des Implementationserfolgs interventionsspezifische Schüler*innenoutcomes vorher?

Aus den bisherigen Befunden der Implementationsforschung geht hervor, dass die Konstrukte Wiedergabetreue und Akzeptanz in positivem Zusammenhang mit der Wirksamkeit einer Intervention stehen (vgl. Kapitel 3.5.2). Aktuell liegen jedoch nur wenige Studien vor, die die Wiedergabetreue und bzw. oder Akzeptanz von unterrichtlichen Interventionen aus Schüler*innenperspektive erheben und im Zusammenhang zum Lernerfolg der Schüler*innen untersuchen. Dies wird im Rahmen der ersten Forschungsfrage adressiert und die Vorhersagekraft der Nutzung

der Promptkarten und Akzeptanz des mehrsprachigkeitssensiblen Reziproken Lehrens für Schüler*innenoutcomes untersucht, die mit der Intervention adressiert werden. Neben direkten Zusammenhängen (vgl. Kapitel 4.1.1) wird zusätzlich die vermittelnde Rolle der Akzeptanz (vgl. Kapitel 4.1.2) analysiert. Als interventionspezifische Schüler*innenoutcomes werden die Ausprägungen des Leseverstehens, der Lesefreude, der Selbstwirksamkeit der Lesestrategieanwendung und der Einstellungen zu Mehrsprachigkeit am Ende der Intervention untersucht. Die ausgewählten Outcomes spiegeln die zwei zentralen Bausteine der Intervention wider – die Förderung der Lesekompetenz und die Integration von Mehrsprachigkeit.

4.1.1 Direkte Vorhersagekraft der Nutzung der Promptkarten und Akzeptanz für interventionsspezifische Schüler*innenoutcomes

Es wird angenommen, dass durch die Verwendung der Promptkarten in der Kleingruppe und die damit zusammenhängende Gestaltung des Lernangebots lernförderliche innere Nutzungsprozesse auf Seiten der Schüler*innen angestoßen werden (Vieluf et al., 2020). Höhere Ausprägung der Facetten des Implementationserfolgs sollten demnach höhere Ausprägungen interventionsspezifischer Schüler*innenoutcomes am Ende der Intervention vorhersagen. Dabei wird vermutet, dass die Lesestrategien durch die Nutzung der Promptkarten systematisch eingeübt werden und dies zur Verbesserung der Lesekompetenz beiträgt (vgl. Kapitel 2.3). Weiterhin wird vermutet, dass die Nutzung der Promptkarten in der Gruppenarbeit lernförderliche kognitive Aktivitäten begünstigt (z. B. gegenseitiges Erklären, Argumentieren, Elaborieren, Vergleichen), die die Interaktionsqualität erhöhen und mit einem höheren Lernzuwachs beim kooperativen Lernen in Zusammenhang stehen (Kaendler et al., 2015; Lipowsky et al., 2022). Die Nutzung der Promptkarten während der Gruppenarbeit sollte demnach mit höheren Ausprägungen interventionsspezifischer Schüler*innenoutcomes einhergehen. In Anlehnung an Modelle der Implementationsforschung wird weiterhin angenommen, dass eine hohe Akzeptanz im Sinne einer positiven affektiven und kognitiven Gesamtbewertung mit der Wirksamkeit einer Maßnahme in Zusammenhang steht (Witt & Elliott, 1985). Auch aus unterrichtspsychologischer Perspektive kann dies argumentiert werden, da die Wahrnehmung und Interpretation von Unterricht kognitive und motivational-affektive Lernprozesse beeinflussen kann, die sich im Lernzuwachs widerspiegeln können (Vieluf et al., 2020). Durch die bilinguale Gestaltung der Promptkarten, die auch eine Wertschätzung für sprachliche Heterogenität mit sich bringt, wird weiterhin angenommen, dass die Facetten des Implementationserfolgs mit den Einstellungen zu Mehrsprachigkeit positiv zusammenhängen. Vor diesem Hintergrund werden folgende Hypothesen spezifiziert:

(H1.1) (a) Die Nutzung der Promptkarten während der Intervention und (b) die Akzeptanz sagen das *Leseverstehen* am Ende der Intervention positiv vorher.

(H1.2) (a) Die Nutzung der Promptkarten während der Intervention und (b) die Akzeptanz sagen die *Lesefreude* am Ende der Intervention positiv vorher.

(H1.3) (a) Die Nutzung der Promptkarten während der Intervention und (b) die Akzeptanz sagen die *Selbstwirksamkeit der Lesestrategieanwendung* am Ende der Intervention positiv vorher.

(H1.4) (a) Die Nutzung der Promptkarten während der Intervention und (b) die Akzeptanz sagen die *Einstellungen zu Mehrsprachigkeit* am Ende der Intervention positiv vorher.

4.1.2 Vermittelnde Rolle der Akzeptanz

Modelle der Implementationsforschung nehmen reziproke Zusammenhänge zwischen der Wiedergabetreue und Akzeptanz sowie der Wirksamkeit einer Intervention an (Witt & Elliott, 1985). In der vorliegenden Arbeit wird vermutet, dass die Nutzung der Promptkarten während der Intervention ein positives Erleben auf affektiver und kognitiver Ebene begünstigt, welches sich in einer höheren Ausprägung an Akzeptanz am Ende der Intervention zeigt. Auch in Angebot-Nutzungsmodellen der Unterrichtsforschung wird angenommen, dass Nutzungsprozesse der Schüler*innen über ihre individuelle Wahrnehmung des Lernangebots vermittelt werden (Vieluf et al., 2020). Es wird vermutet, dass die Nutzung der Promptkarten während der Intervention lernförderliche Nutzungsprozesse begünstigt, die über die Akzeptanz als Wahrnehmungsprozesse vermittelt werden und mit Outcomes auf kognitiver und motivational-affektiver Ebene zusammenhängen. Im Rahmen von Implementationsstudien wurde diese Annahme mit Konstrukten, die ausschließlich auf Ebene der Schüler*innen angesiedelt sind, noch nicht überprüft.

(H1.5) Es wird angenommen, dass die Akzeptanz den positiven Zusammenhang zwischen der Nutzung der Promptkarten und (a) *dem Leseverstehen*, (b) *der Lesefreude*, (c) *der Selbstwirksamkeit der Lesestrategieanwendung* und (d) *den Einstellungen zu Mehrsprachigkeit* am Ende der Intervention vermittelt.

4.2 Fragestellung 2: Welche individuellen Schüler*innenmerkmale sagen die Facetten des Implementationserfolgs vorher?

In einem konstruktivistischen Verständnis von Unterricht werden Lernangebote von Schüler*innen in Abhängigkeit ihrer Lernvoraussetzungen unterschiedlich wahrgenommen und genutzt (z. B. Helmke, 2012; Vieluf et al., 2020). Nach einem konstruktivistischen Verständnis von Innovationen kann zudem angenommen werden, dass Schüler*innen als Akteur*innen den Implementationserfolg in Abhängigkeit ihrer individuellen Merkmale und Kompetenzen beeinflussen (Gröschner, 2011). Im Rahmen der zweiten Forschungsfrage soll untersucht werden, wie individuelle Unterschiede in den Facetten des Implementationserfolgs durch Lernvoraussetzungen der Schüler*innen erklärt werden können.

Für den spezifischen Kontext einer mehrsprachigkeitssensiblen Intervention ist in der vorliegenden Arbeit zunächst interessant, ob die Mehrsprachigkeit der Schüler*innen als individuelles Merkmal mit den Facetten des Implementationserfolgs in Zusammenhang steht. Aus bisherigen Untersuchungen geht hervor, dass mehrsprachige Schüler*innen sich zwar wünschen, dass ihre Sprachen im Unterricht zum Einsatz kommen und wertgeschätzt werden, die Aktivierung mehrsprachiger Kommunikation sich im Regelunterricht jedoch häufig schwierig gestaltet und Schüler*innen unterschiedlich mit ihrer eigenen Mehrsprachigkeit umgehen (vgl. Kapitel 2.2.2). Erste qualitative Untersuchungen zur Umsetzung mehrsprachigkeitssensibler Konzepte zeigen zudem, dass sowohl mehrsprachige als auch monolinguale Schüler*innen Unterricht mit mehrsprachigkeitssensiblen Elementen positiv bewerten (Götz et al., 2017; Hilbe et al., 2017a; Schnitzer, 2020). Das Merkmal der Mehrsprachigkeit könnte entsprechend als distales Merkmal die Facetten des Implementationserfolgs vorhersagen.

Für distale Merkmale, wie das Geschlecht, die Mehrsprachigkeit oder der sozioökonomische Status, wird jedoch angenommen, dass Zusammenhänge zum Lernerfolg über proximale Merkmale vermittelt werden, z. B. kognitive und affektiv-motivationale Variablen, die eine größere Nähe zum Lerngegenstand aufweisen (Helmke & Weinert, 1997). Diese beeinflussen die kognitive Informationsverarbeitung und werden in der Lehr-Lernforschung daher als besonders einflussreich für den Lernerfolg angesehen. Erste Untersuchungen zu individuellen Determinanten des Implementationserfolgs auf Schüler*innenebene zeigen ebenfalls, dass lerngegenstandsbezogene Voraussetzungen die Wiedergabetreue (Brisson et al., 2020) und Akzeptanz (Eckert et al., 2021) von Schüler*innen vorhersagen können.

Im Rahmen der zweiten Fragestellung wird zunächst die Mehrsprachigkeit der Schüler*innen als distales Merkmal in den Blick genommen. Aufgrund der wenigen und uneinheitlichen Forschungsbefunde wird explorativ untersucht, ob die Mehrsprachigkeit die Facetten des Implementationserfolgs vorhersagen kann. Als proximale Merkmale werden weiterhin das Leseverstehen, die Lesefreude, die Selbstwirksamkeit der Lesestrategianwendung und die Einstellungen zu Mehrsprachigkeit zu Beginn der Intervention untersucht. Die Konstrukte wurden

analog zu den interventionsspezifischen Schüler*innenoutcomes (Ausprägung der Konstrukte am Ende der Intervention) gewählt, um unterschiedliche Facetten der Lesekompetenz abzubilden, die mit der Intervention angesprochen werden (vgl. Kapitel 2.3.1), sowie die Integration von Mehrsprachigkeit als innovatives Element zu berücksichtigen.

4.2.1 Individuelle Determinanten der Nutzung der Promptkarten

Das Merkmal der Mehrsprachigkeit könnte die Nutzung der Promptkarten vorhersagen, da diese bilingual gestaltet sind und für mehrsprachige Schüler*innen demnach besonders ansprechend sein könnten. Die Promptkarten sind jedoch für alle Schüler*innen zentraler Bestandteil der Leseintervention, um die Aufgaben effektiv zu bearbeiten, weswegen die Zusammenhänge explorativ untersucht werden sollen.

(F2.1) Sagt die *Mehrsprachigkeit* der Schüler*innen die Nutzung der Promptkarten während der Intervention vorher?

Für die lerngegenstandsbezogenen Voraussetzungen der Lesekompetenz kann angenommen werden, dass das Leseverstehen zu Beginn der Intervention die Nutzung der Promptkarten positiv vorhersagt, da die Nutzung von Lernangeboten in besonderem Maße vom Vorwissen abhängig ist (Chi & Wylie, 2014). Kompetente Leser*innen können zudem Lesestrategien besser anwenden (z.B. Artelt et al., 2002) und die Promptkarten daher eher nutzen, weil ihnen die Anwendung der Lesestrategien leichtfällt. Auch das fachspezifische Interesse und affektiv-motivationale Voraussetzungen wie die Lesefreude und Selbstwirksamkeitserwartungen beeinflussen leistungsbezogene Entscheidungen und Handlungen (Guthrie et al., 2004; Schiefele, 1996; Wigfield & Eccles, 2000) und werden daher in positivem Zusammenhang mit der Nutzung der Promptkarten vermutet. Analog zum Merkmal der Mehrsprachigkeit kann auch für die Einstellungen zu Mehrsprachigkeit vermutet werden, dass Schüler*innen mit positiven Einstellungen die Promptkarten besonders ansprechend finden und eher verwenden als Schüler*innen mit negativen Einstellungen. Da die Promptkarten jedoch für alle Schüler*innen zentraler Bestandteil der Intervention sind, werden die Zusammenhänge der Einstellungen zu Mehrsprachigkeit analog zu Fragestellung F2.1 explorativ untersucht. Folgende Hypothesen und Forschungsfragen werden formuliert:

(H2.1) Das *Leseverstehen* zu Beginn der Intervention sagt die Nutzung der Promptkarten während der Intervention positiv vorher.

(H2.2) Die *Lesefreude* zu Beginn der Intervention sagt die Nutzung der Promptkarten während der Intervention positiv vorher.

(H2.3) Die *Selbstwirksamkeit der Lesestrategieanwendung* zu Beginn der Intervention sagt die Nutzung der Promptkarten während der Intervention positiv vorher.

(F2.2) Sagen die *Einstellungen zu Mehrsprachigkeit* zu Beginn der Intervention die Nutzung der Promptkarten während der Intervention vorher?

Neben den direkten Zusammenhängen soll anschließend ebenfalls explorativ untersucht werden, ob die lerngegenstandsbezogenen Voraussetzungen als proximale Merkmale einen möglichen Zusammenhang zwischen dem Merkmal Mehrsprachigkeit und der Nutzung der Promptkarten medieren.

(F2.3) Wird der Zusammenhang zwischen der Mehrsprachigkeit und der Nutzung der Promptkarten über lerngegenstandsbezogene Voraussetzungen vermittelt?

4.2.2 Individuelle Determinanten der Akzeptanz

Es wird erwartet, dass die Akzeptanz als einstellungsähnliches Konstrukt, das auf der Erfahrung mit einem Akzeptanzobjekt und resultierenden affektiven und kognitiven Bewertungsprozessen beruht (vgl. Kapitel 3.4), von affektiven und kognitiven Lernvoraussetzungen der Schüler*innen beeinflusst wird. Im Vergleich zur Nutzung der Promptkarten muss jedoch unterschieden werden, dass die Akzeptanz als Gesamturteil am Ende der Intervention erhoben wurde und damit alle Bestandteile der Intervention in die Bewertung einfließen können.

Für die individuelle Mehrsprachigkeit als distales Merkmal soll ebenfalls explorativ untersucht werden, ob diese die Ausprägung der Akzeptanz vorhersagt. Ähnlich wie für die Nutzung der Promptkarten kann auf der einen Seite argumentiert werden, dass die Intervention für mehrsprachige Schüler*innen besonders ansprechend ist, allerdings ist die Befundlage hierzu uneinheitlich (vgl. Kapitel 4.2).

(F2.4) Sagt die *Mehrsprachigkeit* der Schüler*innen die Akzeptanz am Ende der Intervention vorher?

Für die individuellen Voraussetzungen der Lesekompetenz können ähnliche Wirkungszusammenhänge wie zur Nutzung der Promptkarten angenommen werden. Kompetente und motivierte Leser*innen könnten ein höheres Maß an Akzeptanz zeigen, da sie die Lesetexte besser verstehen, die Strategien besser beherrschen und insgesamt ein größeres Interesse am Unterrichtsgegenstand aufweisen als weniger kompetente Leser*innen. Der Einfluss der Einstellungen zu Mehrsprachigkeit wird jedoch explorativ untersucht, da hier unterschiedliche Erklärungsansätze möglich sind (vgl. Kapitel 4.2.1).

(H2.4) Das *Leseverstehen* zu Beginn der Intervention sagt die Akzeptanz am Ende der Intervention positiv vorher.

(H2.5) Die *Lese Freude* zu Beginn der Intervention sagt die Akzeptanz am Ende der Intervention positiv vorher.

(H2.6) Die *Selbstwirksamkeit der Lesestrategieanwendung* zu Beginn der Intervention sagt die Akzeptanz am Ende der Intervention positiv vorher.

(F2.5) Sagen die *Einstellungen zu Mehrsprachigkeit* zu Beginn der Intervention die Akzeptanz am Ende der Intervention positiv vorher?

Anschließend soll auch hier explorativ untersucht werden, ob die lerngegenstandsbezogenen Voraussetzungen als proximale Merkmale einen möglichen Zusammenhang zwischen der Mehrsprachigkeit und der Akzeptanz medieren.

(F2.6) Wird der Zusammenhang zwischen der Mehrsprachigkeit und der Akzeptanz über lerngegenstandsbezogene Voraussetzungen vermittelt?

4.3 Fragestellung 3: Sagt die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen interventionsspezifische Schüler*innenoutcomes vorher?

Die Integration von mehrsprachigem Unterrichtsmaterial stellt eine pädagogische Strategie und didaktische Innovation zum mehrsprachigkeitssensiblen Unterrichten dar (vgl. Kapitel 2.2; vgl. Kapitel 3.1). Bislang ist wenig darüber bekannt, wie mehrsprachige Schüler*innen diese Impulse annehmen. Im Rahmen der dritten Fragestellung soll daher die Nutzung der Promptkarten in einer anderen Sprache fokussiert werden. Analog zur ersten Fragestellung wird zunächst der Frage nachgegangen, ob die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen mit höheren Ausprägungen in interventionsspezifischen Outcomes einhergeht.

4.3.1 Direkte Vorhersagekraft der Nutzung der Promptkarten in einer anderen Sprache für interventionsspezifische Schüler*innenoutcomes

Aus theoretischer Perspektive kann angenommen werden, dass die Nutzung des gesamten sprachlichen Repertoires Lernprozesse bei mehrsprachigen Schüler*innen erleichtern kann, indem Kommunikationsmöglichkeiten erweitert und sprachliche Barrieren abgebaut werden (García 2009; Grosjean 2013; vgl. Kapitel 2.1.4). Daher wird angenommen, dass mehrsprachige Schüler*innen höhere Ausprägungen in den ausgewählten interventionsspezifischen Schüler*innen-

outcomes aufweisen, wenn sie die Promptkarten während der Intervention auch in einer anderen Sprache nutzen und somit ihr gesamtes sprachliches Repertoire im Lernprozess aktivieren (vgl. Kapitel 4.1.1). Analog zu Fragestellung 1 wird die Akzeptanz der Intervention als Facette des Implementationserfolgs ebenfalls in die Analysen der Teilstichprobe eingeschlossen.

(H3.1) (a) Die Nutzung der Promptkarten in einer anderen Sprache und (b) die Akzeptanz sagen bei mehrsprachigen Schüler*innen das *Leseverstehen* am Ende der Intervention positiv vorher.

(H3.2) (a) Die Nutzung der Promptkarten in einer anderen Sprache und (b) die Akzeptanz sagen bei mehrsprachigen Schüler*innen die *Lesefreude* am Ende der Intervention positiv vorher.

(H3.3) (a) Die Nutzung der Promptkarten in einer anderen Sprache und (b) die Akzeptanz sagen bei mehrsprachigen Schüler*innen die *Selbstwirksamkeit der Lesestrategieanwendung* am Ende der Intervention positiv vorher.

(H3.4) (a) Die Nutzung der Promptkarten in einer anderen Sprache und (b) die Akzeptanz sagen bei mehrsprachigen Schüler*innen die *Einstellungen zu Mehrsprachigkeit* am Ende der Intervention positiv vorher.

4.3.2 Vermittelnde Rolle der Akzeptanz

Analog zu Kapitel 4.1.2 wird angenommen, dass die Nutzung der Promptkarten in einer anderen Sprache positiv mit den Akzeptanzurteilen der Schüler*innen zusammenhängt. Durch die Erweiterung des sprachlichen Repertoires können Lernprozesse positiv beeinflusst werden und dies könnte sich in höheren Ausprägungen der Akzeptanzurteile zeigen, welche wiederum interventionsspezifische Schüler*innenoutcomes positiv beeinflussen.

(H3.5) Es wird angenommen, dass die Akzeptanz den positiven Zusammenhang zwischen der Nutzung der Promptkarten in einer anderen Sprache und (a) *dem Leseverstehen*, (b) *der Lesefreude*, (c) *der Selbstwirksamkeit der Lesestrategieanwendung* und (d) *den Einstellungen zu Mehrsprachigkeit* am Ende der Intervention vermittelt.

4.4 Fragestellung 4: Welche individuellen Schüler*innenmerkmale sagen die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen vorher?

Um einen näheren Einblick zu erhalten, welche Schüler*innen vom innovativen Angebot der bilingualen Promptkarten Gebrauch machen und Prozesse ableiten zu können, die mit dem Erfolg mehrsprachigkeitssensibler Konzepte in Zusammenhang stehen, werden schließlich individuelle Determinanten der Nutzung der Promptkarten in einer anderen Sprache untersucht. Wie auch in Fragestellung 2 wird vermutet, dass individuelle Lernvoraussetzungen die Umsetzung des Lernangebots beeinflussen (vgl. Kapitel 4.2). Da Schüler*innen individuell mit ihrer eigenen Mehrsprachigkeit umgehen und spezifisch die Aktivierung der Mehrsprachigkeit fokussiert wird, wird hier jedoch vermutet, dass die Einstellungen zu Mehrsprachigkeit einen positiven Zusammenhang zur Nutzung der Promptkarten in einer anderen Sprache aufweisen. Folgende Hypothesen werden untersucht:

(H4.1) Das *Leseverstehen* zu Beginn der Intervention sagt die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen positiv vorher.

(H4.2) Die *Lese Freude* zu Beginn der Intervention sagt die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen positiv vorher.

(H4.3) Die *Selbstwirksamkeit zur Lesestrategieanwendung* zu Beginn der Intervention sagt die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen positiv vorher.

(H4.4) Die *Einstellungen zu Mehrsprachigkeit* zu Beginn der Intervention sagen die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen positiv vorher.

5 Methodisches Vorgehen

Im folgenden Kapitel wird das methodische Vorgehen zur Analyse der Fragestellungen erläutert. Hierzu werden die Datengrundlage (vgl. Kapitel 5.1), die Stichprobe (vgl. Kapitel 5.2) und die verwendeten Erhebungsinstrumente (vgl. Kapitel 5.3) vorgestellt sowie anschließend die Durchführung der statistischen Analysen beschrieben (vgl. Kapitel 5.4).

5.1 Datengrundlage

Die Untersuchung basiert auf den Daten des vom Bundesministerium für Bildung und Forschung geförderten Projekts „Förderung der Deutsch-Lesekompetenz durch mehrsprachigkeitssensibles Reziprokes Lehren im Grundschulunterricht“ (meRLe). Das Verbundprojekt meRLe wurde von Oktober 2017 bis Dezember 2021 umgesetzt, unter der Leitung von Prof. Dr. Jasmin Decristan an der Bergischen Universität Wuppertal und Prof. Dr. Dominique Rauch am DIPF, dem Leibniz-Institut für Bildungsforschung und Bildungsinformation in Frankfurt am Main und der Pädagogischen Hochschule Ludwigsburg. Ziel des meRLe-Projekts war es, die Deutsch-Lesekompetenz von Grundschulkindern unter Einbezug der lebensweltlichen Mehrsprachigkeit in den Klassen zu fördern. Hierzu wurde eine Unterrichtseinheit für vierte Klassen entwickelt, in der die Methode des Reziproken Lehrens mit mehrsprachigkeitssensiblen Impulsen verknüpft wurde (Decristan et al., 2022). Im Folgenden wird die Unterrichtseinheit sowie das Design der Intervention und das Vorgehen bei der Datenerhebung beschrieben.

5.1.1 Intervention zum mehrsprachigkeitssensiblen Reziproken Lehren

Die im Projekt meRLe entwickelte Unterrichtseinheit wurde auf Basis bestehender Programme zur Leseförderung durch Reziprokes Lehren konzipiert (Decristan et al., 2021; Munser-Kiefer, 2014; Spörer et al., 2016) und mit mehrsprachigkeitssensiblen Elementen angereichert. Die Inhalte der Unterrichtseinheit wurden vom Projektteam in Zusammenarbeit mit drei Lehrkräften aus der schulischen Praxis entwickelt. Die Lehrkräfte unterstützten die didaktische Ausarbeitung der einzelnen Stunden und Aufbereitung der Materialien durch ihre praktische Expertise in den Bereichen Grundschul- und Förderschullehramt sowie der Förderung von Deutsch als Zweitsprache. Im Folgenden werden Ablauf und Inhalte der Unterrichtseinheit sowie die begleitenden Fortbildungen für teilnehmende Lehrkräfte beschrieben.

5.1.1.1 Unterrichtseinheit

Die Unterrichtseinheit zum mehrsprachigkeitssensiblen Reziproken Lehren umfasst zwölf Unterrichtsstunden von jeweils 45 Minuten. Im Kern der Einheit steht die Methode des Reziproken Lehrens, in der Schüler*innen in Kleingruppen und reziproker Rollenverteilung die Lesestrategien „Fragen stellen“ und „Zusammenfassen“ einüben. Zusätzlich werden mehrsprachigkeitssensible Elemente integriert, um alle sprachlichen Ressourcen der Schüler*innen einzubeziehen.

Der Ablauf der Unterrichtseinheit kann in drei Phasen unterteilt werden. In der „Einführungsphase“ (1.–3. Unterrichtsstunde) wird zunächst die Rahmenstory vorgestellt. Die Schüler*innen begleiten den Hauptcharakter „Merle“ bei einer Reise um die Welt und lernen verschiedene Kontinente und Länder kennen. Die Rahmenstory wird zu Beginn jeder Unterrichtsstunde fortgeführt. Die erste Phase dient zusätzlich zur Etablierung eines wertschätzenden Klimas für Mehrsprachigkeit, das die gesamte Unterrichtseinheit begleitet. In der anschließenden „Einübungsphase“ (4.–7. Unterrichtsstunde) führt die Lehrkraft die Lesestrategien „Fragen stellen“ und „Zusammenfassen“ erstmals ein. Die Strategien werden nach dem Prinzip *explicit teaching before reciprocal teaching* (Rosenshine & Meister, 1994; vgl. Kapitel 2.3.3) von der Lehrkraft zunächst erklärt und die Ausführung durch Verbalisierung der kognitiven Prozesse modelliert. Dies wird unterstützt durch die Verwendung von Promptkarten, auf denen die wichtigsten Schritte zum Ausführen der Strategien beschrieben sind. Die Promptkarten sind in der meRLe-Intervention bilingual gestaltet und die Strategien jeweils auf Deutsch und einer der Familiensprachen der Schüler*innen angegeben (vgl. Abbildung 2).

In der „Festigungsphase“ (8.–11. Unterrichtsstunde) werden die Lesestrategien in Kleingruppen nach der Methode des Reziproken Lehrens von den Schüler*innen weiter eingeübt. Die Unterstützung der Lehrkraft nach der Methode des *Scaffoldings* (vgl. Kapitel 2.3.3) wird dabei schrittweise zurückgenommen, sodass die Schüler*innen die Strategien zunehmend selbstständig anwenden. In jeder Stunde der Festigungsphase werden zwei Lesetexte in deutscher Sprache zur Auswahl gestellt, die ein Erlebnis von Merle auf ihrer Weltreise beschreiben. Die Lesetexte bestehen aus vier Abschnitten. Im ersten Schritt der Gruppenarbeit werden die Abschnitte des Lesetextes von den Schüler*innen abwechselnd vorgelesen. Anschließend werden die Lesestrategien „Fragen stellen“ und „Zusammenfassen“ auf alle Abschnitte des Lesetextes angewendet. Die Schüler*innen arbeiten hierbei zunächst im Tandem und verwenden die Promptkarten, um die Anwendung zu unterstützen und ihre Ergebnisse zu überprüfen. Anschließend werden die Ergebnisse der Tandems in der Vierergruppe vorgestellt und gegenseitig mithilfe der Promptkarten überprüft. Die Promptkarten stellen dabei ein wichtiges Kernelement der Intervention dar, um lernförderliche kognitive Aktivitäten in der Gruppenarbeit (z. B. durch gegenseitiges Erklären, Argumentieren und Begründen) (vgl. Kapitel 2.3.2) und die Qualität der Interaktion zu unterstützen. Am Ende der Stunde wird ein gemeinsames Gruppenprodukt erstellt, indem die Zusam-

Abbildung 2: Bilinguale Promptkarten zu den Lesestrategien „Fragen stellen“ und „Zusammenfassen“ (Beispiel: Deutsch-Türkisch)

Zusammenfassen		Fragen stellen	
!	!	?	?
1. Nenne das WICHTIGSTE WER (Hauptperson) oder WAS (Hauptsache)!	1. Ana KİM (karakter) veya ana NEY (eşyayı) söyle!	1. Ist es eine W-FRAGE? (Wer? Wie? Wo? Was? Wann? Warum?)	1. Konuyu aydınlatacak bir SORU mu? (Kim? Nasıl? Nerede? Ne? Ne zaman? Neden?)
			
2. Nenne das WICHTIGSTE ÜBER das WER (Hauptperson) oder WAS (Hauptsache)!	2. Ana KİM (karakter) veya ana NEY (eşya) hakkındaki EN ÖNEMLİ bilgiyi ver!	2. Ist es eine W-Frage zu einem WICHTIGEN Inhalt?	2. Metnin içeriğiyle ilgili ÖNEMLİ bir soru mu?
			
3. Verwende EIGENE WORTE!	3. KENDİ CÜMLELERİNLE ifade et!	3. Lässt sich die Frage mit dem TEXT BEANTWORTEN?	3. Sorunun cevabı METİNDE BULUNUYOR MU?
			
4. Fasse dich KURZ! Schreibe deine Zusammenfassung in ein bis zwei Sätzen auf!	4. Özeti KISA olmalıdır. Özeti 1-2 cümlede çıkar.	4. Schreibe deine Frage auf!	4. Sorunu kağıda yaz.

menfassungen für jeden Textabschnitt in der richtigen Reihenfolge auf ein buntes Papier aufgeklebt werden. Die Gruppenprodukte werden anschließend im Klassenverband vorgestellt und gemeinsam besprochen. Mithilfe eines Lernportfolios wird der Lernprozess durch die Lernstrategien „Ziele setzen“ und „Reflektieren“ unterstützt. Dieser Ablauf des Reziproken Lehrens wird in den vier Sitzungen der Festigungsphase jeweils wiederholt. Die einzelnen Arbeitsphasen sind dabei zeitlich strukturiert und werden in der Klasse mithilfe eines Tafeltimers visualisiert. Die letzte Sitzung (12. Unterrichtsstunde) dient schließlich der Zusammenfassung und gemeinsamen Reflexion der Reise um die Welt mit Merle. Alle Unterrichtsstunden werden mit motivationalen Elementen gerahmt. Als Einstieg wird jeweils ein neuer Teil der Rahmengeschichte vorgelesen und ein neues Land bereist. Am Ende der Stunde erhalten die Schüler*innen einen Belohnungsticker für ihr Lernportfolio, das in Form eines Reisepasses gestaltet ist.

Die Unterrichtseinheit wird durch mehrsprachigkeitssensible Impulse angereichert. Wie bereits beschrieben, dient die Einführungsphase der Etablierung eines wertschätzenden Klimas für Mehrsprachigkeit. Dazu wird die Mehrsprachigkeit der Schüler*innen zu Beginn der Einheit thematisiert und im Klassenraum sichtbar gemacht. Hierbei wurde ein kreativer Zugang gewählt, bei dem alle Schüler*innen ein individuelles Sprachenportrait (Gogolin & Neumann, 1991) anfertigen. Schüler*innen zeichnen hierbei ihre Sprachen mit unterschiedlichen Farben in eine menschliche Silhouette ein, um die eigene Sprachbiografie zu re-

flektieren und die emotionale Bedeutung der Sprachen sichtbar zu machen (vgl. Kapitel 2.2.1). Auch die thematische Einbettung der Rahmengeschichte dient der Etablierung eines wertschätzenden Klimas, da Merle auf ihrer Reise unterschiedliche Kontinente, Kulturen und Sprachen kennenlernt. Weiterhin sollen im gesamten Verlauf der Unterrichtseinheit mehrsprachige Äußerungen zugelassen und die Schüler*innen zur Nutzung ihres gesamten sprachlichen Repertoires angeregt werden. Um mehrsprachige Kommunikation in allen Phasen des Reziproken Lehrens zu ermöglichen, erfolgte die Gruppenzusammensetzung nach Möglichkeit so, dass innerhalb einer Kleingruppe Sprachpartner*innen in Tandems zusammenarbeiten können, die eine gemeinsame Sprache neben dem Deutschen sprechen. Ideal ist hierbei eine Zusammensetzung von zwei mehrsprachigen Kindern mit möglichst gleicher Familiensprache und zwei monolingual-deutschen Kindern, die gemeinsam eine Vierergruppe bilden. Zusätzlich wird die Aktivierung der Mehrsprachigkeit durch die bilingualen Promptkarten und die Nutzung eines Sprachlernstifts unterstützt. Selbst wenn Schüler*innen kein*e Sprachpartner*in zugeordnet werden kann, können die Familiensprachen über die Nutzung der Promptkarten in den Lernprozess integriert werden. Mit dem Sprachlernstift können zusätzlich die Angaben in den Familiensprachen vorgelesen werden, falls die Schüler*innen in den Sprachen nicht alphabetisiert sind.

Die Promptkarten wurden im Projekt meRLe von sprachkompetenten Personen in 33 Familiensprachen² übersetzt und den Schüler*innen entsprechend ihrer Sprachen ausgeteilt. Zusätzlich zu den schriftlichen Übersetzungen wurden Audioaufnahmen in den Familiensprachen für die Sprachlernstifte erstellt. Zum Anhören dieser Übersetzungen wurden die Promptkarten in der Familiensprache mit speziellen Aufklebern versehen, mit denen die Audiodateien auf dem Stift abgespielt werden können. Für alle 33 schriftlichen Übersetzungen lagen Audiodateien in den Sprachen vor. Für die Sprachen Tigrinisch und Mazedonisch konnten keine schriftlichen Übersetzungen erstellt werden, weswegen hier lediglich Audiodateien vorlagen. Die Sticker für diese beiden Sprachen wurden auf monolingual-deutsche Promptkarten ohne schriftliche Übersetzung aufgeklebt.

5.1.1.2 Fortbildungen für Lehrkräfte

Bevor teilnehmende Lehrkräfte die Unterrichtseinheit in ihren Klassen umsetzten, nahmen sie an einer dreitägigen Fortbildung teil, in der die Unterrichtseinheit vorgestellt und gemeinsam diskutiert wurde. Die Fortbildungen fanden an drei

2 Die Promptkarten wurden in folgende 33 Sprachen übersetzt: Albanisch, Arabisch, Armenisch, Berberisch, Bulgarisch, Chinesisch, Dari, Englisch, Französisch, Griechisch, Italienisch, Japanisch, Kiswahili, Kroatisch, Kurdisch, Mongolisch, Niederländisch, Paschto, Persisch, Polnisch, Portugiesisch, Punjabi, Rumänisch, Russisch, Serbisch, Spanisch, Tamilisch, Tschechisch, Türkisch, Ukrainisch, Ungarisch, Urdu, Vietnamesisch.

Nachmittagen in aufeinander folgenden Wochen für jeweils vier Stunden statt und wurden von einem interdisziplinären Team, bestehend aus dem wissenschaftlichen Projektteam und Lehrkräften aus der Schulpraxis, die die Konzeption der Unterrichtseinheit begleitet haben, geleitet. Zur Begleitung der Unterrichtseinheit erhielten Lehrkräfte in der Fortbildung ein Manual mit allen Stundenverlaufsplänen, der Rahmengeschichte, Lesetexten und Arbeitsblättern sowie die bilingualen Promptkarten zu den Lesestrategien, die jeweils auf die sprachlichen Ressourcen der Schüler*innen ihrer Klasse abgestimmt waren. Jede Lehrkraft erhielt zudem zwei Sprachlernstifte, die die Lesestrategien auf den Promptkarten in den jeweiligen Sprachen abspielen konnten.

Am ersten Tag der Fortbildung wurde die Einführung in die Unterrichtseinheit und die Mehrsprachigkeit in den Klassen der Lehrkräfte thematisiert. Gemeinsam mit dem Projektteam erarbeiteten die Lehrkräfte die mögliche Gruppenzusammensetzung nach Sprachpartner*innen. Am zweiten Tag wurden die Einführung der Lesestrategien und die Nutzung der Promptkarten behandelt. Die Lehrkräfte wurden hierbei angeleitet, die Lesestrategien und den Dialog in der Kleingruppe mithilfe der Promptkarten zu modellieren. Am dritten Tag wurde der strukturierte Ablauf des Reziproken Lehrens in der Festigungsphase besprochen. Notwendige kleine Adaptionen für den individuellen Klassenkontext, wie z.B. für leistungsschwächere Schüler*innen, wurden mit dem Projektteam diskutiert. Die Vorstellung der Unterrichtseinheit wurde durch wissenschaftliche Impulse des Projektteams unterstützt sowie durch gemeinsame Übungen zur Umsetzung der Gruppenarbeit.

5.1.2 Design und Datenerhebung

Das Projekt meRLe wurde als längsschnittliche quasi-experimentelle Interventionsstudie konzipiert, um die Wirksamkeit der Unterrichtseinheit in einer Treatmentgruppe (TG) im Vergleich zum regulären Deutschunterricht in einer Warte-Kontrollgruppe (W-KG) untersuchen zu können (Decristan et al., 2022). Das Interventionsdesign umfasst zwei Implementationsphasen, die mit Beginn des zweiten Schulhalbjahres im Januar 2019 begannen (vgl. Tabelle 1). Die Lehrkräfte der TG setzten in der ersten Phase nach der Teilnahme an den Fortbildungen die meRLe-Intervention in ihren Klassen um, während in den Klassen der W-KG regulärer Deutschunterricht stattfand. Nach Abschluss der Intervention durch die TG setzten die Lehrkräfte der W-KG die meRLe Intervention in der zweiten Phase des Projekts in ihren Klassen um. Die längsschnittlichen Wirkungen des Projekts können anhand dieses Designs durch die begleitenden Erhebungen vor und nach den Implementationsphasen untersucht werden. Für die vorliegende Untersuchung ist der Kontrollzeitraum des regulären Unterrichts jedoch nicht relevant. Stattdessen werden die Daten der TG und W-KG zusammengefasst, um die Implementation der Unterrichtseinheit an einer möglichst großen Stichprobe untersuchen zu

können. Im Folgenden wird daher auch nur auf die für die Untersuchung relevanten Erhebungszeitpunkte eingegangen.

Tabelle 1: Interventions- und Erhebungsdesign der Datengrundlage

	Phase 1 (01 bis 04/2019)	Phase 2 (04 bis 06 /2019)
TG	meRLe-Intervention	Regulärer Unterricht
W-KG	Regulärer Unterricht	meRLe-Intervention

In allen teilnehmenden Klassen wurden vor und nach der Unterrichtseinheit schriftliche Schüler*innenbefragung von geschulten Testleiter*innen durchgeführt. Die Befragungen wurden als Gruppenerhebungen mit der gesamten Klasse umgesetzt und umfassten Tests zur Lesekompetenz und kognitiven Grundfähigkeiten sowie Fragebogenskalen zu motivational-affektiven Merkmalen, der Bewertung und Umsetzung des Trainings und soziodemographischen Angaben. Die Leistungstests und Fragebogenitems wurden zur Verständnissicherung von den Testleiter*innen laut vorgelesen, während die Schüler*innen diese gleichzeitig ausfüllten. Ein Erhebungszeitpunkt umfasste 90 Minuten, in denen zwei kurze Pausen zur kognitiven und körperlichen Entspannung enthalten waren. Eine detaillierte Beschreibung der Erhebungsinstrumente, die für die Auswertung der vorliegenden Arbeit verwendet wurden, ist in Kapitel 5.3 zu finden.

Die erste Schüler*innenbefragung fand im Januar und Februar 2019 in den Klassen der TG statt. Diejenigen Schüler*innen, die eine Einwilligung zur Teilnahme an den schriftlichen Befragungen von ihren Eltern bzw. Erziehungsberechtigten erhalten hatten, nahmen an der 90-minütigen Paper-Pencil-Befragung teil. Nach Abschluss der ersten Implementationsphase und vor Beginn der Intervention in der W-KG fand die zweite Schüler*innenbefragung parallel in beiden Gruppen im März und April 2019 statt. Diese Befragung diente als Post-Interventionsbefragung der TG und als Prä-Interventionsbefragung in der W-KG vor Beginn der zweiten Implementationsphase (vgl. Tabelle 1). Nach Abschluss der Intervention in der W-KG wurde im Juni 2019 die Post-Interventionserhebung für diese Klassen durchgeführt. Zwischen den Prä- und Post-Erhebungen in beiden Untersuchungsgruppen lagen durchschnittlich 63.9 Tage ($SD = 10.4$; Min. = 49 Tage; Max = 92 Tage). Für die vorliegende Untersuchung werden die Prä- und Post-Erhebungsdaten für beide Interventionsgruppen jeweils zusammengefasst.

Die Lehrkräfte wurden zu drei Zeitpunkten im Verlauf des Projekts befragt. Vor und nach der Fortbildung beantworteten Lehrkräfte Fragen zu ihren Vorerfahrungen und Erwartungen sowie zur Evaluation der Fortbildung. Nach der Umsetzung der Intervention in den Klassen wurde Lehrkräften ein weiterer Fragebogen mit Fragen zur Implementation ausgeteilt. Die Eltern der teilnehmenden Schüler*innen erhielten außerdem einen Fragebogen zum soziodemographischen

Hintergrund und der Teilnahme ihres Kindes an außerunterrichtlichen Förderprogrammen. Zur Erfassung der Implementation auf Klassenebene wurde in jeder Klasse eine Unterrichtsstunde der Festigungsphase zwischen der achten und elften Sitzung, in der das Reziproke Lehren vollständig in den Gruppen angewendet wurde (vgl. Kapitel 5.1.1), videografiert oder beobachtet. Dies wird in Kapitel 5.1.3 genauer beschrieben.

5.1.3 Implementation der Unterrichtseinheit

Die Intervention wurde im Rahmen des merLe-Projekts in insgesamt 38 Klassen implementiert und empirisch begleitet. Die Unterrichtseinheit wurde im Klassenverbund im Deutschunterricht von den Lehrkräften selbst umgesetzt. Für die Umsetzung wurde ein Zeitraum von ca. 6 Wochen, mit jeweils 2 Sitzungen pro Woche, empfohlen. Aus den Angaben der Lehrkräfte geht hervor, dass der tatsächliche Umsetzungszeitraum im Durchschnitt bei 5.4 Wochen lag ($SD = 1.7$; Min. = 2 Wochen; Max. = 10 Wochen) mit durchschnittlich 2,7 Unterrichtsstunden pro Woche ($SD = 0,8$; Min. = 2 Unterrichtsstunden; Max. = 4 Unterrichtsstunden).

Um zu überprüfen, ob die Intervention in allen Klassen wie intendiert umgesetzt wurde, wurde die Wiedergabetreue auf Klassenebene erfasst. In jeder Klasse wurde dazu eine Unterrichtsstunde der Festigungsphase beobachtet bzw. videografiert und die Umsetzung von einzelnen Kernelementen des Stundenablaufs anhand einer Implementations-Checkliste dichotom bewertet (0 = nicht umgesetzt; 1 = umgesetzt). Die Kodierung wurde von geschulten Beobachter*innen durchgeführt. In 17 Klassen stimmten Lehrkräfte der Videografierung einer Unterrichtsstunde zu. In 21 Klassen lehnten Lehrkräfte die Videografierung ab, sodass in diesen Klassen eine Unterrichtsstunde beobachtet wurde. Die Implementations-Checkliste besteht aus 10 Items, die sich nach dem Stundenablauf im begleitenden Manual richten (vgl. Kapitel 5.1.1). Die Items sind in Tabelle 2 aufgeführt. Diese bilden die Strukturen ab, die im Stundenverlaufsplan für jeden Stunde vorgegeben wurden und stellen Kernelemente für die Wirksamkeit der Intervention dar, wie z. B. der Einstieg mit der Rahmenstory, das abschnittsweise Lesen und die Anwendung von Promptkarten. Da einige dieser Aktivitäten gruppenbasiert umgesetzt werden, erfolgte die Kodierung für diese Kernelemente an sogenannten Fokustischen (FT). Hierzu wurde in jeder Klasse ein Gruppentisch beobachtet.

Tabelle 2 zeigt den prozentualen Umsetzungsindex für jedes Kernelement der Intervention. Von 38 Klassen wurden 27 Klassen doppelt kodiert, um die Übereinstimmung der Ratings zu überprüfen. Über alle Items hinweg ist eine prozentuale Übereinstimmung von 97.0% zwischen den Rater*innen zu beobachten. Zeigten die Rater*innen keine Übereinstimmung, wurde für die folgenden Analysen der Mittelwert aus beiden Ratings verwendet. Da bei einigen Kodierungen vollständige Übereinstimmung vorlag, konnte Kappa nicht berechnet werden.

Tabelle 2: Prozentualer Implementationsindex der Kernelemente

Items nach Stundenablauf	<i>M</i> (%)	<i>(SD)</i> (%)	Übereinst. (%)
Einleitung: Rahmengeschichte mit Kontinent/Figur	93.4	(23.7)	96.3
Schüler*innen arbeiten im Logbuch (Ziele setzen)	94.7	(22.6)	100.0
Schüler*innen lesen Text	100.0	(0.0)	100.0
Schüler*innen wechseln sich beim Lesen ab (FT)	97.3	(11.5)	92.6
Schüler*innen bearbeiten Arbeitsblatt zu Lesestrategien	100.0	(0.0)	100.0
Schüler*innen arbeiten gemeinsam am Arbeitsblatt (FT)	100.0	(0.0)	100.0
Einsatz von Promptkarten	89.2	(31.5)	96.3
Schüler*innen erstellen ein Gruppenprodukt (FT)	93.4	(23.7)	96.3
Schüler*innen arbeiten im Logbuch (Ziele reflektieren)	88.2	(31.7)	92.6
Abschluss: Belohnungsticker im Logbuch	81.6	(39.3)	96.3
Gesamt	94.2	(10.0)	97.0

Der Implementationsindex aller 38 Klassen liegt bei durchschnittlich 94.2% ($SD = 10.0$) der Kernelemente, was eine hohe Wiedergabetreue des Unterrichtsablaufs auf Klassenebene widerspiegelt. Weiterhin wurde für jede Klasse ein prozentualer Implementationsindex berechnet, der angibt, wie viel Prozent der Kernelemente in dieser Klasse umgesetzt wurden. 24 Klassen zeigen einen Index von 100%. In diesen Klassen wurden alle Kernelemente wie intendiert umgesetzt. 13 Klassen zeigen einen Implementationsindex zwischen 95.0% und 70.0% und lediglich eine Klasse liegt mit 60.0% der umgesetzten Kernelemente unter dem empfohlenen Cut-Off-Wert von 70% für die Wiedergabetreue (vgl. Kapitel 3.3.1). Ein ausreichendes Maß an Wiedergabetreue gilt jedoch als Voraussetzung für eine valide Auswertung der Ergebnisse. Die Auswertung der Forschungsfragen soll daher nur auf Basis der Klassen erfolgen, in denen die Intervention wie intendiert umgesetzt wurde, um die Angaben der Schüler*innen zur Nutzung der Promptkarten und Akzeptanz der Intervention sinnvoll interpretieren zu können. Die Klasse mit einem Implementationsindex von 60.0% wird daher für alle folgenden Analysen ausgeschlossen.

Ein Item des Implementationschecks bezieht sich auf den Einsatz der Promptkarten. In vier Klassen wurde jedoch keine Nutzung der Promptkarten beobachtet. Da die Nutzung der Promptkarten in der vorliegenden Arbeit als zentrale Variable untersucht wird, muss zusätzlich geprüft werden, ob die Promptkarten in diesen Klassen lediglich in der beobachteten bzw. videografierten Unterrichtsstunde nicht zum Einsatz kamen oder von der Lehrkraft in der gesamten Unterrichtseinheit nicht eingeführt wurden. In Klassen, in denen die Lehrkraft die Promptkarten nicht eingeführt und während der Gruppenarbeit den Schüler*innen zur Verfügung gestellt hat, können Zusammenhänge zu individuel-

len Schüler*innenmerkmalen nicht sinnvoll untersucht und die Akzeptanzurteile nicht interpretiert werden. Daher werden in diesen vier Klassen zusätzlich die Angaben der Schüler*innen aus den begleitenden Erhebungen untersucht. Hierbei zeigt sich, dass in einer Klasse keine*r der Schüler*innen angegeben hat, die Promptkarten während der Intervention genutzt zu haben. Da sich diese Angabe auf den gesamten Verlauf der Intervention bezieht und nicht nur auf eine beobachtete Unterrichtsstunde, kann davon ausgegangen werden, dass die Promptkarten in dieser Klasse nicht eingeführt wurden. Diese Klasse wird daher ebenfalls für die weiteren Analysen ausgeschlossen, sodass eine Stichprobe von 36 Klassen erhalten bleibt.

5.2 Stichprobe

Ausgehend von den Standorten des Verbundprojekts wurden Klassen der TG in Nordrhein-Westfalen und Klassen der W-KG in Hessen rekrutiert. Die Rekrutierung der Stichprobe erfolgte via Telefon und E-Mail. Hierzu wurden Schulleitungen von Grundschulen in den Bundesländern zu Beginn des Schuljahres 2018/2019 kontaktiert, um Informationen zur Projektteilnahme an ihre Deutsch-Lehrkräfte der vierten Jahrgangsstufe weiterzuleiten. Insgesamt wurden Schulen angesprochen, in denen ein hoher Anteil mehrsprachiger Schüler*innen unterrichtet wird und die vierten Klassen an keinen anderen Interventionen oder Trainings zur Leseförderung teilnehmen.

In beiden Bundesländern nahmen insgesamt 38 Klassen am Projekt meRLe teil (23 in der TG; 15 in der W-KG). Für die Analysen der vorliegenden Untersuchung werden zwei Klassen ausgeschlossen, die unzureichende Implementation auf Klassenebene zeigten (vgl. Kapitel 5.1.3). Die Stichprobe der vorliegenden Arbeit umfasst demnach 36 Klassen mit insgesamt 634 Schüler*innen, die an mindestens einer der schriftlichen Erhebungen während des Projekts teilgenommen haben.³ 50.6 % der Schüler*innen sind Mädchen (48.9 % Jungen, 0.5 % fehlend) und die Schüler*innen sind im Mittel 10.1 Jahre alt ($SD = 0.6$). 70.8 % der Schüler*innen wurden als mehrsprachig kodiert, da sie angegeben haben, neben dem Deutschen noch mindestens eine weitere Sprache zu sprechen (28.5 % monolingual; 0.6 % fehlend). 27.8 % der mehrsprachigen Schüler*innen gaben an, eine andere Sprache als Deutsch zuerst gelernt zu haben (8.9 % gaben Deutsch als Erstsprache an; 43.0 % Deutsch und eine andere Sprache; 20.3 % fehlend). Insgesamt wurden 58 unterschiedliche Familiensprachen bzw. Varietäten von mehrsprachigen Schüler*innen angegeben. Am häufigsten wurde hierbei von 31.8 % ($n = 143$) der mehrsprachigen Schüler*innen Türkisch genannt und von 11.8 % ($n = 53$) Arabisch. Der An-

3 Im Rahmen vorbereitender Analysen wurden zusätzlich 12 Schüler*innen aus der Analyse ausgeschlossen, die nicht plausible Antworten gegeben haben. Die Angabe der Stichprobe erfolgt hier bereits unter Ausschluss dieser Fälle. Die betreffende Datenbereinigung wird in Kapitel 5.3.1 beschrieben.

teil mehrsprachiger Schüler*innen pro Klasse beträgt durchschnittlich 71.9% ($SD = 16.5$). Hierbei können jedoch nur die Schüler*innen berücksichtigt werden, die an den Erhebungen teilgenommen haben. Diejenigen Schüler*innen, die kein Einverständnis der Eltern zu Teilnahme erhalten hatten, gehen in diese Auswertungen nicht ein.

5.3 Erhebungsinstrumente und Operationalisierung

Im Folgenden werden die Erhebungsinstrumente für die zu untersuchenden Konstrukte vorgestellt. Die Instrumente werden bezüglich ihrer Itemkennwerte, der Dimensionalität und Reliabilität der Skalen sowie ihrer Messinvarianz zwischen den Gruppen monolingualer und mehrsprachiger Schüler*innen untersucht.

5.3.1 Nutzung der Promptkarten

Als Facette der Wiedergabebetreue wurde die Nutzung der Promptkarten durch die Schüler*innen während der Intervention erfasst. Zum einen wurde für alle Schüler*innen allgemein erhoben, ob sie die Promptkarten während der Intervention genutzt haben („*Hast du bei der Weltreise mit Merle die Hinweiskarten / Promptkarten zu den Lesestrategien genutzt?*“). Weiterhin wurde erhoben, ob mehrsprachige Schüler*innen die Promptkarten in einer Familiensprache genutzt haben, da diese bilingual gestaltet sind („*Hast du die Hinweiskarten / Promptkarten in einer anderen Sprache als Deutsch genutzt?*“). Beide Items wurden im dichotomen Antwortformat erhoben (0 = Nein; 1 = Ja).

Theoretisch wird angenommen, dass mit dem ersten Item eine allgemeine Nutzung der Promptkarten in der Gesamtstichprobe erfragt wird. Hierbei wird noch nicht zwischen unterschiedlichen Sprachen differenziert. Das zweite Item fragt schließlich, ob mehrsprachige Schüler*innen das spezifische Angebot zur Nutzung ihrer Familiensprachen angenommen und die Promptkarten (auch) auf einer anderen Sprache genutzt haben. Schüler*innen, die die Promptkarten auf einer anderen Sprache genutzt haben, sollten entsprechend im ersten Item bereits angegeben haben, die Promptkarten während der Intervention genutzt zu haben. Die Plausibilität der Angaben wurde auf diese Annahme hin untersucht. Tabelle 3 zeigt eine Kreuztabelle zwischen beiden Items in der Gesamtstichprobe. Der Tabelle ist zu entnehmen, dass zwölf Schüler*innen angaben, die Promptkarten während der Intervention nicht genutzt zu haben, jedoch auf einer anderen Sprache. Diese Angaben sind nicht plausibel und für diese Schüler*innen kann nicht abschließend festgestellt werden, ob sie die Promptkarten genutzt oder die Befragung gewissenhaft ausgefüllt haben. Aus diesem Grund werden die betreffenden Fälle für alle weitere Analysen ausgeschlossen. Die Gesamtstichprobe umfasst nach Ausschluss

dieser Fälle 634 Schüler*innen. Die Angaben zur Stichprobe in Kapitel 5.2 erfolgten bereits unter Ausschluss dieser Fälle.

Tabelle 3: Kreuztabelle zur Nutzung der Promptkarten und Nutzung der Promptkarten in einer anderen Sprache

		Nutzung der Promptkarten in einer anderen Sprache		
		Ja	Nein	Gesamt
		<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)
Nutzung der Promptkarten	Ja	347 (57.9 %)	156 (24.1 %)	503 (77.9 %)
	Nein	12 (1.9 %)	57 (8.7 %)	69 (10.7 %)
	Fehlend			74 (11.5 %)
	Gesamt			646 (100 %)

Die Promptkarten sind bilingual gestaltet und lagen in 33 Übersetzungen vor (vgl. Kapitel 5.1.1), die den Schüler*innen entsprechend ihrer Familiensprache(n) ausgeteilt wurden. Für 96.1% ($n = 441$) der mehrsprachigen Schüler*innen lagen Promptkarten in mindestens einer der angegebenen Familiensprachen vor. Lediglich für 17 mehrsprachige Schüler*innen (3.8 %) konnten aus organisatorischen Gründen keine Promptkarten in ihren Familiensprachen erstellt werden.⁴ Für einen Fall (0.2 %) konnte zusätzlich nicht eindeutig kodiert werden, ob die Promptkarten in der passenden Sprache vorlagen. Für diese 18 Schüler*innen kann die dritte und vierte Fragestellung nicht sinnvoll analysiert werden, da dort die Aktivierung der Familiensprachen über die Nutzung der Promptkarten in einer anderen Sprache im Fokus steht. Daher werden diese insgesamt 18 Schüler*innen aus den Analysen zur Subgruppe mehrsprachiger Schüler*innen in der dritten und vierten Fragestellung ausgeschlossen.

5.3.2 Akzeptanz

Zur Erfassung der Akzeptanzurteile wurden zwei Skalen eingesetzt, die die Akzeptanz als Gesamtbewertung am Ende der Intervention auf affektiver und kognitiver Ebene abbilden. Die Skalen wurden nach bestehenden Instrumenten für den spezifischen Kontext der Intervention im Projekt meRLe adaptiert (vgl. Eckert et al., 2017; Hebbecker & Souvignier, 2018) und als Likert-Skala mit vierstufigem Antwortformat (1 = stimmt gar nicht – 4 = stimmt genau) in der Post-Interventionsbefragung erhoben. Die affektive Akzeptanzdimension wird mit fünf Items ab-

4 Für folgende Sprachen konnten keine Übersetzungen erstellt werden: Amharisch, Aramäisch, Armenisch, Bengalisch, Burmesisch, Ghanaisch, Hausa, Mandinka, Mazedonisch, Philippinisch, Romanes, Senegalesisch, Tigrinisch.

gebildet, die sich auf das affektive Erleben während der Bearbeitung der Lese- und Lernstrategien beziehen. Die Skala der kognitiven Akzeptanzdimension umfasst drei Items, die die eigene Wirksamkeitserwartung und Nützlichkeit der Intervention zur Verbesserung des Leseverstehens abbilden.

Da die Skalen aus bestehenden Instrumenten adaptiert wurden, werden zunächst die Item- und Skalencharakteristiken geprüft. Diese sind Tabelle 4 zu entnehmen. Beide Subskalen weisen eine gute Reliabilität auf (affektive Akzeptanz: $\alpha = .87$; kognitive Akzeptanz: $\alpha = .87$).

Tabelle 4: Deskriptive Kennwerte der Items zu den Akzeptanzdimensionen

Item	<i>M</i>	(<i>SD</i>)	<i>r_{it}</i>	<i>N</i>
<i>Affektive Akzeptanz</i>				
Das Üben der Lese- und Lernstrategien hat mir gut gefallen.	3.29	(0.87)	.71	589
Durch die Weltreise mit Merle habe ich Lust, das Lesen weiter zu üben.	2.96	(1.10)	.64	585
Ich habe die Lese- und Lernstrategien gerne geübt.	3.10	(0.96)	.79	584
Es hat mir Spaß gemacht, die Lese- und Lernstrategien zu üben.	3.11	(0.99)	.80	583
Das Üben der Lese- und Lernstrategien war langweilig. (r)	3.10	(1.08)	.54	578
<i>Kognitive Akzeptanz</i>				
Durch die Weltreise mit Merle bin ich besser im Lesen geworden.	2.79	(1.14)	.74	579
Die Weltreise mit Merle hat mir geholfen, das Lesen zu üben.	2.85	(1.10)	.80	581
Durch die Weltreise mit Merle kann ich Texte besser verstehen.	2.99	(1.08)	.71	586

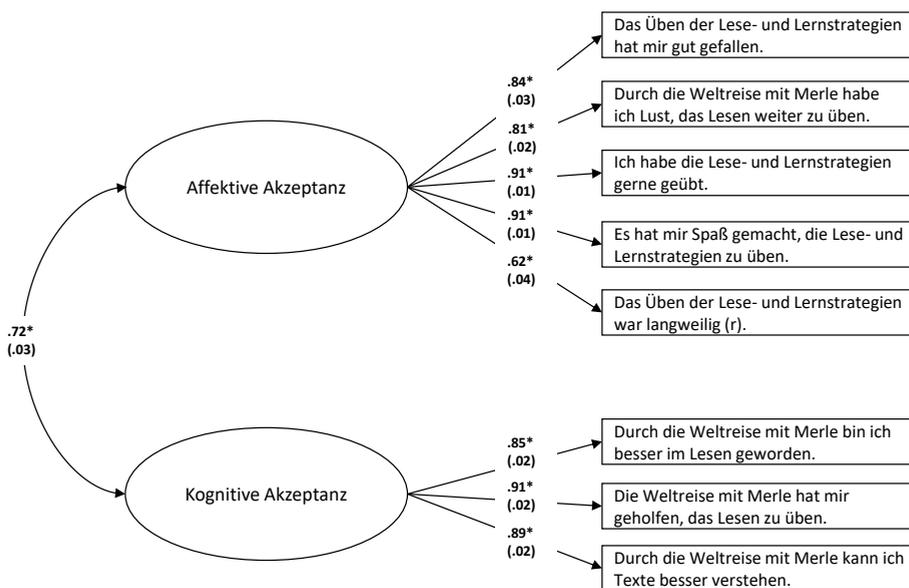
Anmerkung. (r): Item wurde vor der Auswertung rekodiert.

Insgesamt zeigen sich für die Items der affektiven Akzeptanz deskriptiv höhere Mittelwerte als für die Items der kognitiven Akzeptanz. Alle Items beider Subskalen weisen eine hohe Trennschärfe von $< .5$ auf (Bortz & Döring, 2006). Die Analysen der Trennschärfen geben insgesamt keine Hinweise darauf, Items für die weitere Untersuchung auszuschließen.

In Anlehnung an Eckert et al. (2021) wird angenommen, dass die Skalen eine affektive und kognitive Akzeptanz als getrennte Dimensionen abbilden. Zur empirischen Überprüfung dieser Annahme wird die faktorielle Struktur der Skalen mittels konfirmatorischer Faktorenanalyse getestet. Das methodische Vorgehen der Faktorenanalyse wird in Kapitel 5.4.2 beschrieben. Das theoretisch postu-

lierte Zweifaktorenmodell zeigt im Gesamtbild eine zufriedenstellende Modellgüte an und kann daher bestätigt werden: $\chi^2(19) = 125.822$ ($p < .001$), CFI = .976, RMSEA = .098, $N = 590$. Alle standardisierten Faktorladungen sind signifikant und $> .5$. Die Dimensionen korrelieren signifikant positiv miteinander ($r = .72$, $p < .001$) (vgl. Abbildung 3).

Abbildung 3: Messmodell der affektiven und kognitiven Akzeptanz



Anmerkung. * $p < .05$, Standardfehler in Klammern angegeben.

Da die Dimensionen in hohem Maße miteinander korrelieren, wird zusätzlich analysiert, ob ein einfaktorielles Modell die Daten besser repräsentieren kann, in welchem alle Items auf einen gemeinsamen Faktor laden. Die Modellstatistiken dieser genesteten Modelle sind Tabelle 5 zu entnehmen. Die Gütekriterien des einfaktoriellen Modells liegen nicht mehr im akzeptablen Bereich (vgl. Kapitel 5.4.2). Ein χ^2 -Differenztest zeigt zudem, dass sich die Modelle hinsichtlich ihres Modellfits signifikant voneinander unterscheiden. Vor diesem Hintergrund kann das theoretisch abgeleitete zweifaktorielle Modell beibehalten werden, da die Daten dadurch besser repräsentiert werden. Die Ergebnisse unterstützen die theoretische Annahme einer affektiven und kognitiven Akzeptanz auf Schüler*innenebene, die zwar miteinander zusammenhängen, inhaltlich und empirisch jedoch voneinander zu unterscheiden sind. Die Untersuchung der Daten und Analyse der Fragestellungen erfolgt demnach unter Berücksichtigung dieser getrennten Dimensionen.

Tabelle 5: Modellstatistiken zum Vergleich der Faktorenstruktur der Akzeptanz

	2-Faktorenmodell	1-Faktormodell
$\chi^2 (df)$	125.822 (19)	464.225 (20)
χ^2/df	7.622	23.211
$p(\chi^2)$	< .001	< .001
CFI	.976	.902
RMSEA	.098	.194
$\Delta\chi^2(df)$		125.072(1); $p < .001$

Anmerkung. $N = 590$.

Da ein Teil der Analysen in der Substichprobe mehrsprachiger Schüler*innen durchgeführt wird, wird zusätzlich überprüft, ob die Konstrukte messinvariant zwischen den Gruppen monolingualer und mehrsprachiger Schüler*innen sind. Die theoretischen und methodischen Grundlagen zur Überprüfung der Messinvarianz werden in Kapitel 5.4.2.1 im Rahmen der statistischen Analysen beschrieben. Zunächst wird der Modellfit in beiden Gruppen getrennt betrachtet. Es zeigt sich für beide Gruppen ein guter bis zufriedenstellender Modellfit (Monolingual: $\chi^2(19) = 62.845$ ($p < .001$), CFI = .982, RMSEA = .117, $N = 170$; Mehrsprachig: $\chi^2(19) = 85.928$, ($p < .001$), CFI = .979, RMSEA = .092, $N = 416$), wobei der RMSEA in der Gruppe monolingualer Schüler*innen den Cut-Off-Wert von .1 knapp übersteigt. Dies könnte auf die kleine Stichprobe für diese Gruppe zurückzuführen sein (vgl. Kapitel 5.4.2.1).

Tabelle 6 zeigt die Modellstatistiken zur Überprüfung der Messinvarianz zwischen den Gruppen. Der χ^2 -Differenztest gibt an, dass sich der Modellfit des Zweifaktorenmodells zwischen den Gruppen nicht signifikant unterscheidet und skalare Messinvarianz angenommen werden kann. Auch die geringen Veränderungen im CFI und RMSEA lassen auf skalare Messinvarianz zwischen den Gruppen schließen (vgl. Kapitel 5.4.2.1).

Tabelle 6: Modellstatistiken zur Überprüfung der Messinvarianz des Zweifaktorenmodells der Akzeptanz zwischen monolingualen und mehrsprachigen Schüler*innen

Modell	Konfigural	Metrisch	Skalar
χ^2 (df)	150.739 (38)	161.567 (44)	169.436 (58)
χ^2/df	3.967	3.672	2.921
$p(\chi^2)$	< .001	< .001	< .001
CFI	.981	.980	.981
RMSEA	.101	.095	.081
Faktorkorrelation (r)*	.810 / .678	.810 / .677	.810 / .678
N	170 / 416	170 / 416	170 / 416
$\Delta\chi^2(df)$		11.918(6); $p = .064$	13.595(14); $p = .480$

Anmerkung. *Faktorkorrelationen in den Gruppen monolingualer/mehrsprachiger Schüler*innen.

5.3.3 Individuelle Schüler*innenmerkmale

Im Folgenden wird die Operationalisierung der individuellen Schüler*innenmerkmale beschrieben, die in die Analyse der Forschungsfragen eingehen. Diese umfassen die individuelle Mehrsprachigkeit, das Leseverstehen, die Lesefreude, die Selbstwirksamkeit der Lesestrategieanwendung, die Einstellungen zu Mehrsprachigkeit sowie die Kontrollvariablen Geschlecht, kognitive Grundfähigkeiten und Anzahl an Büchern im Haushalt.

5.3.3.1 Mehrsprachigkeit

Die individuelle Mehrsprachigkeit der Schüler*innen wurde im Rahmen des meRLe-Projekts in einer dichotomen Variable kodiert (0 = Monolingual; 1 = Mehrsprachig). Als theoretischer Rahmen wurde hierbei die lebensweltliche Mehrsprachigkeit herangezogen (vgl. Kapitel 2.1.1). Sprachen, die im Schulkontext fremdsprachlich erworben wurden, werden daher nicht berücksichtigt. Die Kodierung basiert auf Angaben der Schüler*innen, Lehrkräfte und Eltern aus den begleitenden Befragungen. Als Grundlage für die Kodierung werden die Angaben der Schüler*innen aus der ersten Befragung zur offenen Frage „*Welche Sprachen sprichst du?*“ herangezogen. Mehrsprachigkeit wurde dann kodiert, wenn von den Schüler*innen neben dem Deutschen noch mindestens eine weitere Sprache angegeben wurde. Fehlende Angaben der Schüler*innen wurden zunächst durch Angaben der Eltern aufgefüllt, welche Sprachen zuhause gesprochen werden. Wenn hier neben dem Deutschen noch eine weitere Sprache angegeben wurde,

wurde der*die Schüler*in als mehrsprachig kodiert. Wenn zusätzlich zu fehlenden Schüler*innenangaben auch von der Elternbefragung keine Angabe vorlag, wurde in einem zweiten Schritt die Angabe der Lehrkraft zur Sprache des Kindes herangezogen. Diese wurde von den Lehrkräften für die Zusammensetzung der Gruppen nach Sprachpartner*innen erfragt.

Als zusätzliches Qualitätsmaß wurde eine Variable erstellt, die widersprüchliche Schlussfolgerungen zur Kodierung zwischen den Angaben der Schüler*innen, Eltern und Lehrkräfte erfasst. Die genaue Sprache wird dabei nicht berücksichtigt, sondern lediglich, ob die vorliegenden Perspektiven zur gleichen Kodierung (monolingual oder mehrsprachig) führen. Für die Stichprobe der vorliegenden Untersuchung zeigte sich in 12,8% der Fälle ($n = 81$) ein Widerspruch zwischen den Perspektiven. In 25,6% der Fälle ($n = 162$) lag keine weitere Perspektive für einen Abgleich vor, 61,7% der Fälle ($n = 391$) zeigten keinen Widerspruch zwischen den Perspektiven. Fälle, in denen widersprüchliche Angaben zwischen den Perspektiven identifiziert wurden oder keine Perspektive zum Abgleich vorhanden war, sowie Fälle, die aufgrund fehlender Angaben aller Perspektiven nicht kodiert werden konnten, wurden einer zusätzlichen Prüfung unterzogen. Für diese Fälle wurden Angaben der Sprachbiografie sowie Angaben aus einem adaptierten Wortschatztest für Familiensprachen (adaptiert nach: *Bilingual Verbal Ability Test*; Muñoz-Sandoval et al., 1998) herangezogen. Vorliegende Angaben zur Sprachbiografie umfassen Fragen zur Erstsprache (Erstsprache Deutsch, nicht-Deutsch oder geteilt) sowie die Frage nach der Sprachnutzung mit Familienmitgliedern. Zusätzlich wurde auch geprüft, ob die betreffenden Fälle den adaptierten Wortschatztest auf einer anderen Sprache ausgefüllt haben. Sobald eine der Informationsquellen einen Hinweis auf Mehrsprachigkeit zeigte, wurde der Fall als mehrsprachig kodiert. Andernfalls, wenn Angaben in den genannten Quellen gemacht wurden, sich hier jedoch kein Hinweis auf Mehrsprachigkeit zeigte, wurde der Fall als monolingual kodiert.

Fremdsprachlich erworbene Sprachen sollten in der Kodierung nicht berücksichtigt werden. Da die Schüler*innen auch im Rahmen des Grundschulunterrichts bereits Englisch lernen, musste zusätzlich geprüft werden, ob die Nennung der Sprache Englisch als Familiensprache oder Fremdsprache gilt. Wurde „Schulenglisch“ angegeben, wurde dies nicht als weitere Sprache gewertet. Wurde neben dem Deutschen nur noch „Englisch“ angegeben, wurden ebenfalls Angaben der Eltern, Lehrkräfte und sprachbiografische Variablen zur zusätzlichen Prüfung herangezogen. Dabei wurden die Schüler*innen als mehrsprachig kodiert, wenn Englisch auch in der Elternbefragung oder durch die Lehrkräfte angegeben wurde oder in den sprachbiografischen Variablen, die vorlagen, ein Hinweis auf familiäre Mehrsprachigkeit zu finden war. 7,9% der Fälle ($n = 50$) wurden auf die dargestellte Weise überprüft. In 82,0% dieser Fälle ($n = 41$) wurde Englisch als Schulenglisch identifiziert und die Schüler*innen als monolingual kodiert. In den restlichen neun Fällen wurde Englisch als Familiensprache identifiziert und die Schüler*innen als mehrsprachig kodiert. In der vorliegenden Stichprobe konnten

lediglich vier Fälle (0,6 %) hinsichtlich der Mehrsprachigkeit nicht kodiert werden, da keine der benötigten Angaben vorlag.

24,1% der Fälle ($n = 153$) in der vorliegenden Stichprobe wurden durch eine*n zweite*n geschulte Rater*in doppelt kodiert. Die Übereinstimmung beider Kodierungen liegt bei 93,5 %.

5.3.3.2 Leseverstehen

Das Leseverstehen vor und nach der Intervention wurde über einen standardisierten Lesetest erfasst. Hierbei wurde die Subskala „Textverständnis“ des Lesetests ELFE II eingesetzt (Lenhard et al., 2017). Der Test richtet sich an Schüler*innen der 1.–6. Klasse und ist als Speedtest konzipiert. Schüler*innen haben für die Subskala Textverständnis sieben Minuten Zeit, so viele Items wie möglich zu lösen. Die Items bestehen aus kurzen informativen Texten, zu denen eine bis drei inhaltliche Fragen mit je vier Antwortmöglichkeiten gestellt werden, die aus dem Text hervorgehen. Die Anzahl richtiger Antworten wird in einem Summenwert zusammengefasst.

5.3.3.3 Lesefreude

Die Lesefreude der Schüler*innen wurde über vier Items auf einer vierstufigen Likert-Skala (1 = stimmt gar nicht – 4 = stimmt genau) zu Beginn und am Ende der Intervention erhoben. Die Skala wurde adaptiert nach Bos et al. (2010) und Faber (2010). Die deskriptiven Kennwerte der Items sind in Tabelle 7 angegeben. Zu beiden Messzeitpunkten weisen die Items hohe Trennschärfen ($r_{it} < .5$) und die Skala eine gute Reliabilität auf (Prä: $\alpha = .86$; Post: $\alpha = .87$).

Tabelle 7: Deskriptive Kennwerte der Items zur Skala „Lesefreude“

Item	1. Messzeitpunkt (Prä)				2. Messzeitpunkt (Post)			
	<i>M</i>	(<i>SD</i>)	r_{it}	<i>N</i>	<i>M</i>	(<i>SD</i>)	r_{it}	<i>N</i>
Lesen macht mir Spaß.	3.15	(0.91)	.78	583	3.13	(0.89)	.82	590
Ich lese gerne.	3.09	(0.96)	.79	573	3.10	(0.94)	.79	589
Lesen finde ich langweilig. (r)	3.37	(0.95)	.70	580	3.28	(0.97)	.71	585
Ich hätte gerne noch mehr Zeit zum Lesen.	2.59	(1.07)	.60	579	2.56	(1.07)	.60	587

Konfirmatorische Faktorenanalysen bestätigen für beide Messzeitpunkte in der Gesamtstichprobe eine eindimensionale Struktur und eine gute (Prä: $\chi^2(2) = 4.558$ ($p = .102$), CFI = .999, RMSEA = .047, $N = 583$) bis zufriedenstellende Modellgüte (Post: $\chi^2(2) = 15.514$ ($p < .001$), CFI = .997, RMSEA = .107, $N = 593$).

Dies bestätigt sich auch bei getrennter Untersuchung der Konstrukte für mehrsprachige und monolinguale Schüler*innen (vgl. Tabelle A1 im Anhang). Zudem kann für beide Messzeitpunkte skalare Messinvarianz zwischen den Gruppen angenommen werden (vgl. Tabelle A2 im Anhang).

5.3.3.4 Selbstwirksamkeit der Lesestrategieanwendung

Die Selbstwirksamkeit der Lesestrategieanwendung wurde über vier Items auf einer vierstufigen Likert-Skala (1 = stimmt gar nicht – 4 = stimmt genau) zu Beginn und am Ende der Intervention erhoben. Die Skala wurde nach Koch und Spörer (2016) adaptiert und auf die im Projekt meRLe eingesetzten Lesestrategien „Fragen stellen“ und „Zusammenfassen“ angepasst. Für jede Lesestrategie wurden zwei Items eingesetzt. Die deskriptiven Kennwerte der Items sind in Tabelle 8 angegeben. Zu beiden Messzeitpunkten weisen die Items hohe Trennschärfen ($r_{ii} < .5$) auf und die Skala zeigt gute Reliabilität (Prä: $\alpha = .82$; Post: $\alpha = .85$).

Tabelle 8: Deskriptive Kennwerte der Items zur Skala „Selbstwirksamkeit der Lesestrategieanwendung“

Item	1. Messzeitpunkt (Prä)				2. Messzeitpunkt (Post)			
	<i>M</i>	(<i>SD</i>)	r_{ii}	<i>N</i>	<i>M</i>	(<i>SD</i>)	r_{ii}	<i>N</i>
Wenn ich einen Text zusammenfassen soll, gelingt mir das gut.	2.92	(0.83)	.61	573	3.08	(0.83)	.67	586
Ich schaffe es, mir Fragen zu einem Text zu überlegen.	3.14	(0.85)	.63	577	3.40	(0.80)	.67	586
Ich schaffe es, einen Text zusammenzufassen.	3.06	(0.87)	.66	576	3.21	(0.82)	.72	589
Wenn ich mir Fragen zu einem Text überlegen soll, gelingt mir das gut.	3.17	(0.84)	.66	577	3.37	(0.80)	.73	583

Konfirmatorische Faktorenanalysen zeigen, dass die Modellanpassung hinsichtlich χ^2 -Wert und RMSEA für beide Messzeitpunkte in der Gesamtstichprobe von einer akzeptablen Modellanpassung abweicht. Der CFI zeigt für beide Messzeitpunkte jedoch einen guten Modellfit des eindimensionalen Modells an (Prä: $\chi^2(2) = 63.732$ ($p < .001$), CFI = .964, RMSEA = .230, $N = 584$; Post: $\chi^2(2) = 46.804$ ($p < .001$), CFI = .989, RMSEA = .194, $N = 593$).

Eine für monolinguale und mehrsprachige Schüler*innen getrennte Analyse der Modellanpassung zeigt auch deutliche Abweichungen des χ^2 -Werts und des RMSEA. Die Interpretation des CFIs spricht jedoch in beiden Gruppen für eine gute Modellanpassung (vgl. Tabelle A3 im Anhang). Auf Basis des χ^2 -Differenztests kann für beide Messzeitpunkte keine Invarianz zwischen den Gruppen angenommen werden. Die Veränderung des CFIs und RMSEAs bewegt sich allerdings bis zur skalaren Messinvarianz im Rahmen der beschriebenen Cut-Off-Werte (vgl. Tabelle A4 im Anhang). Für die Interpretation der Analysen muss berücksichtigt werden, dass sich für die Selbstwirksamkeit der Lesestrategien ein gewisser „Misfit“ in den Daten zeigt und die Konstrukte zwischen den Gruppen monolingualer und mehrsprachiger Schüler*innen womöglich Unterschiede aufweisen.

5.3.3.5 Einstellungen zu Mehrsprachigkeit

Die Einstellungen zu Mehrsprachigkeit wurde über vier Items auf einer vierstufigen Likert-Skala (1 = stimmt gar nicht – 4 = stimmt genau) zu Beginn und am Ende der Intervention erhoben. Die Skala wurde im Projekt meRLe selbst entwickelt und deckt affektive und kognitive Überzeugungen zu Mehrsprachigkeit ab. Die Items beziehen sich hierbei auf Mehrsprachigkeit im Allgemeinen und nicht auf die eigene Mehrsprachigkeit, damit die Items sowohl von mehrsprachigen als auch monolingualen Schüler*innen beantwortet werden können. Die deskriptiven Kennwerte der Items sind in Tabelle 9 angegeben. Zu beiden Messzeitpunkten weisen die Items mittlere bis hohe Trennschärfe auf ($r_{it} < .4$) und die Skala zeigt zufriedenstellende Reliabilität (Prä: $\alpha = .72$; Post: $\alpha = .72$).

Tabelle 9: Deskriptive Kennwerte der Items zur Skala „Einstellungen zu Mehrsprachigkeit“

Item	1. Messzeitpunkt (Prä)				2. Messzeitpunkt (Post)			
	<i>M</i>	(<i>SD</i>)	r_{it}	<i>N</i>	<i>M</i>	(<i>SD</i>)	r_{it}	<i>N</i>
Ich bewundere es, wenn jemand noch eine andere Sprache als Deutsch sprechen kann.	3.00	(1.05)	.51	576	3.21	(0.98)	.54	584
Ich würde gerne noch mehr über die Sprachen der anderen Kinder in meiner Klasse wissen.	3.21	(0.99)	.50	578	3.33	(0.94)	.48	589
Ich finde es toll, wenn die Kinder in meiner Klasse im Deutschunterricht in einer anderen Sprache sprechen dürfen.	2.55	(1.21)	.44	576	2.72	(1.21)	.42	582
Ich finde es cool, wenn jemand mehr als eine andere Sprache sprechen kann.	3.29	(0.98)	.58	580	3.39	(0.93)	.62	584

Konfirmatorische Faktorenanalysen zeigen für beide Messzeitpunkte eine gute Modellpassung des eindimensionalen Modells in der Gesamtstichprobe (Prä: $\chi^2(2) = 0.526$ ($p = .768$), CFI = 1.000, RMSEA = .000, $N = 581$; Post: $\chi^2(2) = 0.062$ ($p = .966$), CFI = 1.000, RMSEA = .000, $N = 591$).

Der Modellfit ist auch bei einer für monolinguale und mehrsprachige Schüler*innen getrennten Analyse in beiden Gruppen als gut zu beurteilen (vgl. Tabelle A5 im Anhang). Lediglich der erste Messzeitpunkt in der Gruppe monolingualer Schüler*innen weicht hinsichtlich des RMSEA von einer zufriedenstellenden Modellpassung ab. Für den ersten Messzeitpunkt kann auf Basis aller einbezogenen Gütekriterien skalare Messinvarianz zwischen monolingualen und mehrsprachigen Schüler*innen bestätigt werden. Für den zweiten Messzeitpunkt zeigt der χ^2 -Differenztest eine signifikante Verschlechterung zwischen den Modellen konfiguraler und metrischer, sowie metrischer und skalarer Messinvarianz an. Auf Basis der Veränderungen im CFI und RMSEA könnte jedoch auch für den zweiten Messzeitpunkt skalare Messinvarianz zwischen den Gruppen angenommen werden (vgl. Tabelle A6 im Anhang). Dieser mögliche „Misfit“ in der Modellspezifikation ist bei der Interpretation der Ergebnisse zu berücksichtigen.

5.3.3.6 Kontrollvariablen

Zur Erklärung individueller Unterschiede in der Nutzung der Promptkarten und Ausprägung der Akzeptanzdimensionen werden in der zweiten und vierten Fragestellung individuelle Voraussetzungen der Schüler*innen als Prädiktoren untersucht. Um die interessierenden Zusammenhänge der Mehrsprachigkeit und lerngegenstandsbezogener Variablen sinnvoll interpretieren zu können, werden Kontrollvariablen in die Analyse eingeschlossen, um alternative Erklärungen für die Effekte der unabhängigen Variablen zu berücksichtigen. Theoriegeleitet werden das Geschlecht, kognitive Grundfähigkeiten und die Anzahl an Büchern im Haushalt als Indikator für den sozioökonomischen Status berücksichtigt, die sich in bisheriger Forschung als einflussreich für lesebezogene schulische Leistungen im Kontext von (migrationsbedingter) Mehrsprachigkeit erwiesen haben (vgl. Hesse et al., 2008; Rauch et al., 2010; Segerer et al., 2013). Im Folgenden werden die verwendeten Items und Instrumente zur Erfassung der Kontrollvariablen dargestellt.

Geschlecht. Das Geschlecht der Schüler*innen wurde dichotom erhoben, indem Schüler*innen in der ersten Befragung die Frage „Bist du ein Mädchen oder ein Junge?“ beantworteten. Die dichotome Variable geht mit der Kodierung „0 = Mädchen“ und „1 = Junge“ in die Analysen ein.

Kognitive Grundfähigkeiten. Als Indikator für kognitive Grundfähigkeiten wurde der Grundintelligenztest CFT 20-R (Weiß, 2006) herangezogen und die Subskala „Matrizen“ des ersten Testteils eingesetzt. Der CFT 20-R ist ein sprachfreier Intelligenztest und kann ab der 3. Jahrgangsstufe eingesetzt werden (Weiß, 2006). Der Test ist als Speedtest konzipiert. Zur Bearbeitung der Subskala Matrizen

stehen vier Minuten zur Verfügung, in denen von 15 Items so viele wie möglich beantwortet werden sollen. Die Items bestehen aus Darstellungen von Formen und Mustern, die fortgesetzt werden müssen. Für jedes Item sind fünf Antwortmöglichkeiten gegeben, aus der eine richtige Antwort ausgewählt wird. Die Anzahl der in der vorgegebenen Zeit richtig gelösten Items geht als Summenwert in die Analysen ein.

Anzahl an Büchern im Haushalt. Die Anzahl an Büchern im Haushalt wird häufig als Indikator für den sozioökonomischen Hintergrund in Schüler*innenbefragung erhoben, da dies als Maß für kulturelle Ressourcen im Haushalt einen hohen Zusammenhang zum sozioökonomischen Status aufweist (z. B. Evans et al., 2010; Fauth et al., 2021; McElvany et al., 2009). Die Frage wurde aus dem internationalen Schüler*innenfragebogen der IGLU-Erhebung (Bos et al., 2005) übernommen und als kategoriales Item in der ersten Schüler*innenbefragung erhoben. Die Kategorien reichen von „1 = Keine oder nur sehr wenige Bücher (0–10 Bücher)“ bis „5 = Genug, um drei oder mehr Regale zu füllen (über 200 Bücher)“. Neben jeder Antwortmöglichkeit war eine Abbildung als visuelle Hilfe abgedruckt, die die Anzahl an Büchern bzw. Bücherregalen schematisch darstellt. Die Variable geht dichotomisiert in die Analysen ein. Die ersten zwei Kategorien (bis 25 Bücher) wurden als „0 = Weniger als 26 Bücher im Haushalt“ kodiert. Dies stellt eine niedrige Anzahl an Büchern im Haushalt dar. Die Kategorien drei bis fünf (26 bis über 200 Bücher) wurden als „1 = Mehr als 25 Bücher im Haushalt“ zusammengefasst.

5.4 Statistische Auswertung

Die statistische Auswertung erfolgt mit den Programmen SPSS (Version 29.0) (IBM Corp., 2021) und Mplus (Version 8.4) (Muthén & Muthén, 2012–2021). Im Folgenden wird das Vorgehen beschrieben und die Auswahl der einzelnen Analysen begründet.

5.4.1 Berücksichtigung der hierarchischen Datenstruktur

Die Daten der vorliegenden Untersuchung weisen eine hierarchische Struktur auf, da die Schüler*innen eindeutig einer übergeordneten Struktur (ihrer Klasse) zugeordnet sind. Die Daten sind demnach statistisch nicht unabhängig (Hochweber & Hartig, 2012), was jedoch für viele statistische Verfahren eine wichtige Voraussetzung ist (Snijders & Bosker, 2012). Für Schüler*innen einer Klasse ist keine Unabhängigkeit gegeben, da sie im Vergleich zu Schüler*innen anderer Klassen höhere Ähnlichkeiten in ihren Beurteilungen und Leistungen aufweisen können (Hochweber & Hartig, 2012). Wird dies nicht berücksichtigt, können Standardfehler der untersuchten Merkmale unterschätzt und die Signifikanz der Ergebnisse infolge einer Alpha-Fehler-Inflation überschätzt werden (Hochweber & Hartig, 2012). In-

haltlich sind sogenannte ökologische Fehlschlüsse möglich, wenn Varianzquellen unterschiedlicher Ebenen nicht differenziert und Zusammenhänge von Variablen einer höheren Ebene auf die Individualebene übertragen werden (Eid et al., 2010; Snijders & Bosker, 2012).

In welchem Ausmaß die Daten eine hierarchische Abhängigkeit aufweisen, kann über die Höhe der Intraklassenkorrelation (ICC) bestimmt werden (Snijders & Bosker, 2012). Mit der ICC(1) wird der Anteil der Gesamtvarianz quantifiziert, „der auf Unterschiede zwischen den Klassen zurückzuführen ist“ (Lüdtke et al., 2006, S. 87). Um das Ausmaß der geteilten Wahrnehmungsanteile in der vorliegenden Arbeit zu beurteilen, sind die ICC(1)-Werte der eingesetzten Konstrukte in Tabelle 10 dargestellt. Theoretisch sind alle Konstrukte auf Individualebene zu verorten. Als Teil der Implementation kann für die Akzeptanz und Nutzung der Promptkarten jedoch angenommen werden, dass die Konstrukte ein gewisses Maß an Abhängigkeit vom Klassenkontext aufweisen, da die Intervention auf Klassenebene umgesetzt wurde und je nach Implementation der Lehrkraft zwischen den Klassen variieren kann (z. B. bezüglich der Wiedergabetreue oder Qualität der Instruktion). Für die lerngegenstandsbezogenen Voraussetzungen und Schüler*innenoutcomes ist ebenfalls zu erwarten, dass sich Abhängigkeiten von der Klassenebene zeigen. Nach LeBreton und Senter (2008) kann der ICC(1)-Wert als Effektstärke interpretiert werden. Die Autor*innen beschreiben einen Wert von .01 dabei als kleinen Effekt, einen Wert von .10 als mittleren Effekt und ein Wert von .25 als großen Effekt (LeBreton & Senter, 2008).

Tabelle 10: Varianzkomponentenzerlegung der untersuchten Konstrukte

Variable	ICC(1)	Clustergröße (\emptyset)	N
<i>Facetten des Implementationserfolgs</i>			
Affektive Akzeptanz	.18	16.40	590
Kognitive Akzeptanz	.13	16.40	590
Nutzung der Promptkarten	.08	16.08	579
Nutzung der Promptkarten in einer anderen Sprache	.15	11.25	405
<i>Lerngegenstandsbezogene Voraussetzungen</i>			
Leseverstehen (Prä)	.13	16.08	579
Lese Freude (Prä)	.04	16.19	583
Selbstwirksamkeit der Lesestrategie-anwendung (Prä)	.09	16.22	584
Einstellungen zu Mehrsprachigkeit (Prä)	.06	16.14	581
<i>Interventionsspezifische Schüler*innenoutcomes</i>			
Leseverstehen (Post)	.10	16.33	588
Lese Freude (Post)	.07	16.47	593
Selbstwirksamkeit der Lesestrategie-anwendung (Post)	.05	16.47	593
Einstellungen zu Mehrsprachigkeit (Post)	.05	16.42	591

Tabelle 10 zeigt, dass für die affektive Akzeptanz 18 % der Varianz auf die Klassenebene zurückzuführen ist. Für die kognitive Akzeptanz ist 13 % der Varianz auf die Klassenebene zurückzuführen. Somit zeigen beide Konstrukte substantielle Abhängigkeit von der Klassenzugehörigkeit im Rahmen eines mittleren Effekts. Die Nutzung der Promptkarten in der Gesamtstichprobe zeigt mit 8 % geteilter Varianz eine geringere Abhängigkeit von der Klassenebene auf, die Nutzung der Promptkarten in einer anderen Sprache zeigt dagegen mit 15 % geteilter Varianz einen mittleren Effekt an. Die ICC (1)-Werte der lerngegenstandsbezogenen Voraussetzungen und Schüler*innenoutcomes bewegen sich ebenfalls im Rahmen kleiner Effekte, die größte Abhängigkeit von der Klassenebene zeigt sich hierbei mit 13 % im Leseverstehen zu Beginn der Intervention.

Insgesamt ist zu beobachten, dass ein substantieller Anteil der Varianz in den Facetten des Implementationserfolgs auf Klassenebene zu verorten ist. Dies muss für die Analyse der Fragestellungen methodisch angemessen berücksichtigt werden. Da die Stichprobe der vorliegenden Untersuchung mit 36 Klassen für eine mehrebenenanalytische Auswertung mit vielen Parametern und mediierenden Zusammenhängen jedoch nicht ausreicht (Pötschke, 2019), wird die hierarchische Datenstruktur über die Korrektur der Standardfehler berücksichtigt. Die Modelle

und Analysen werden daher in Mplus über *type = complex* spezifiziert, um eine Alpha-Fehler-Inflation und mögliche inhaltliche Fehlschlüsse zu vermeiden.

5.4.2 Konfirmatorische Faktorenanalysen

Zur Prüfung der Faktorenstruktur der untersuchten Konstrukte wurden konfirmatorische Faktorenanalysen in Mplus durchgeführt. Hierbei wurde das *Weighted-Least-Square-Mean-Variance*-Verfahren (WLSMV) angewendet, welches als besonders präzise Schätzmethode zur Modellierung von kategorialen und ordinalskalierten Daten gilt (Brown, 2006). Für Likert-skalierte Indikatoren mit weniger als fünf Antwortkategorien haben sich robuste kategoriale Schätzmethode, wie das WLSMV-Verfahren, im Vergleich zu robusten kontinuierlichen Methoden, wie dem robusten *Maximum-Likelihood*-Schätzer (MLR), als beste Wahl erwiesen (Rhemtulla et al., 2012). Zudem muss keine multivariate Normalverteilung der Daten vorliegen, die bei vierstufigen Items häufig nicht gegeben ist. Zur Identifizierung des Modells wurde per Voreinstellung in Mplus jeweils die Faktorladung des ersten Indikators eines latenten Faktors auf 1 fixiert, sodass die Varianzen der Faktoren frei geschätzt werden (Geiser, 2011). Die hierarchische Datenstruktur (Schüler*innen in Klassen) wurde bei den Berechnungen durch die Analyse mit *type = complex* berücksichtigt (vgl. Kapitel 5.4.1). Fehlende Werte werden beim WLSMV-Verfahren ausgeschlossen und die Modellanpassung wird nur auf Basis der vorliegenden Antworten geschätzt. Die Faktor-Scores der Modelle stellen jedoch kontinuierliche Variablen dar, die für die Auswertung der Forschungsfragen extrahiert werden (vgl. Kapitel 5.4.2.2).

Bei der Gesamtbeurteilung der Modellpassung wird untersucht, inwiefern die empirischen und geschätzten Varianzen und Kovarianzen voneinander abweichen (Weiber & Mühlhaus, 2014). Mit dem χ^2 -Wert kann die Abweichung „zwischen empirischer Varianz-Kovarianz-Matrix [...] und modelltheoretisch errechneter Varianz-Kovarianz-Matrix“ abgebildet werden (Weiber & Mühlhaus, 2014, S. 204). Eine bessere Anpassung zeigt sich dabei je geringer der χ^2 -Wert ausfällt und sofern das Signifikanzniveau $> .05$ liegt. Hierbei wird jedoch die Anpassung des gesamten Modells überprüft, daher kann der χ^2 -Test auch dann zur Ablehnung führen, wenn sich die Abweichung nur auf Teile des Modells beziehen (Weiber & Mühlhaus, 2014). Zudem reagiert der Test bei großen Stichproben sehr sensibel und auch kleinere Abweichungen der geschätzten Varianzen und Kovarianzen können zur Ablehnung des Modells führen (Weiber & Mühlhaus, 2014). Daher sollte der χ^2 -Wert nur im Verhältnis zu den Freiheitsgraden des Modells interpretiert werden (Weiber & Mühlhaus, 2014). Ein Quotient von < 3 gilt hierbei als erstrebenswert für einen guten Modellfit (Homburg & Giering, 1996), wobei auch hier mögliche Verzerrungen bzgl. der Verletzung der Normalverteilungsannahme beachtet werden müssen. Weiterhin kann die Modellanpassung hinsichtlich unterschiedlicher Fit-Werte interpretiert werden, für die unterschiedliche Cut-Off-Werte als Emp-

fehlungen vorliegen. Eine gute Modellgüte wird nach Hu und Bentler (1999) erreicht, wenn die Werte des *Comparative Fit Index* (CFI) $> .95$ und des *Root-Mean-Square-Error of Approximation* (RMSEA) $< .08$ betragen. Eine akzeptable Modellgüte wird auch noch bei Werten des CFI von $> .90$ (Homburg & Baumgartner, 1995) und des RMSEA von $< .1$ (Browne & Cudeck, 1992) angenommen.

Für die Analyse der vorliegenden Daten muss insgesamt berücksichtigt werden, dass die Empfehlungen der Cut-Off-Werte häufig in Simulationsstudien unter Verwendung kontinuierlicher Daten mit dem *Maximum-Likelihood*-Schätzverfahren (ML) abgeleitet wurden. Klare Richtlinien zur Modellanpassung unter Verwendung kategorialer Schätzmethoden liegen dagegen nicht vor (vgl. Sauerwein, 2017). Die Werte sind daher mit Vorsicht und im Gesamtbild des Modells zu interpretieren.

5.4.2.1 Messinvarianz

Die dritte und vierte Fragestellung werden in der vorliegenden Untersuchung in der Substichprobe mehrsprachiger Schüler*innen untersucht. Damit Konstrukte in unterschiedlichen Substichproben valide untersucht werden können, ist es notwendig, die eingesetzten Instrumente auf Messinvarianz zu untersuchen. Damit kann sichergestellt werden, dass die Instrumente in verschiedenen Gruppen die gleichen Konstrukte messen und Unterschiede in den Ergebnissen nicht auf Verzerrungen der erhobenen Konstrukte zurückzuführen sind (Weiber & Mühlhaus, 2014). Für die vorliegende Untersuchung bedeutet dies, dass die untersuchten Konstrukte sowohl in der Gesamtstichprobe als auch in der Teilstichprobe mehrsprachiger Schüler*innen die gleichen Konstrukte darstellen sollten und somit messinvariant zwischen den Gruppen monolingualer und mehrsprachiger Schüler*innen sind.

Nach Widaman und Reise (1997) werden verschiedene Grade an Messinvarianz unterschieden, die mit zunehmenden Gleichheitsrestriktionen aufeinander aufbauen (vgl. auch Geiser, 2011; Weiber & Mühlhaus, 2014). Die (1) *konfigurale Invarianz* beschreibt die schwächste Form der Invarianz und ist gegeben, wenn die Faktorenstruktur in beiden Gruppen identisch ist und die gleiche Anzahl an Faktoren bestätigt werden kann. Bei mehrdimensionalen Modellen sollten die Korrelationen zwischen den latenten Faktoren zur Beschreibung diskriminanter Validität zudem signifikant von 1 verschieden sein. (2) *Metrische Invarianz* beschreibt den nächsthöheren Grad an Invarianz, bei welchem neben der Faktorenstruktur die Faktorladungen in beiden Gruppen gleichgesetzt werden. Bei Vorliegen der (3) *skalaren Invarianz* sind neben den bereits genannten Restriktionen auch die Intercepts (bei kontinuierlichen Indikatoren) bzw. die Schwellenwerte (bei kategorialen Indikatoren) über beide Gruppen identisch. Den höchsten Grad an Messinvarianz beschreibt die (4) *strikte faktorielle Invarianz*, die vorliegt, wenn zusätzlich auch die Varianzen der Messfehler über die Gruppen hinweg gleichgesetzt werden. In der vorliegenden Untersuchung wird skalare Messinvarianz als ausreichend angesehen,

um mögliche Verzerrungen zwischen den Stichproben zu identifizieren und für die Interpretation der Ergebnisse berücksichtigen zu können. Gruppenvergleiche zwischen mehrsprachigen und monolingualen Schüler*innen werden nicht vorgenommen.

Die Prüfung der Messinvarianz zwischen den Gruppen wurde auf Basis aufeinander aufbauender Mehrgruppen-Konfirmatorischer Faktorenanalysen (MGCFAs) mithilfe von Mplus vorgenommen. Die Modelle wurden unter Verwendung der WLSMV-Schätzmethode mit der Theta-Parametrisierung geschätzt. In einem ersten Schritt wird dabei überprüft, ob das theoretisch postulierte Modell in beiden Gruppen eine gute Passung auf die Daten aufweist (Milfont & Fischer, 2010). Die anschließende Untersuchung der Invarianzstufen erfolgt als MGCFAs anhand des *Grouping*-Befehls in Mplus und über die Spezifikation der unterschiedlichen Messinvarianzstufen (*Model = Configural Metric Scalar*) sowie unter Berücksichtigung der hierarchischen Datenstruktur (*type = complex*).

Vergleiche der Invarianzstufen der Modelle können mit dem χ^2 -Differenztest vorgenommen werden. Der Grad an vorliegender Messinvarianz wird folglich über das Modell mit den meisten Restriktionen erfasst, dessen Modellfit sich im Vergleich zur vorigen Invarianzstufe nicht signifikant verschlechtert hat (Geiser, 2011). Kommt es mit zunehmender Restriktion zu einer signifikanten Verschlechterung des Datenfits, kann der getestete Grad an Invarianz in den verwendeten Daten nicht angenommen werden (Weiber & Mühlhaus, 2014). Unter Verwendung der WLSMV-Schätzmethode wird der χ^2 -Differenztest in Mplus über die Option *DIFFTTEST* durchgeführt.

Weiterhin können zur Beurteilung der Messinvarianz auch die genannten Fit-Werte herangezogen werden (vgl. Kapitel 5.4.2). Nach Cheung und Rensvold (1999) kann Invarianz angenommen werden, wenn sich der CFI zwischen zwei geschachtelten Modellen mit zunehmenden Gleichheitsrestriktionen um weniger als .01 verschlechtert. Zusätzlich kann die Veränderung im RMSEA beurteilt werden, die nach Chen (2007) eine maximale Zunahme um .015 betragen darf, um die nächsthöhere Invarianzstufe anzunehmen. Hierbei ist jedoch zu beachten, dass der RMSEA in kleinen Stichproben schnell hohe Werte annehmen und daher verzerrt sein kann (Kenny et al., 2015). Die genannten Cut-Off Werte sind schließlich vor dem Hintergrund zu interpretieren, dass diese auf Simulationsstudien mit metrischen Daten basieren, häufig unter Verwendung des ML-Schätzverfahrens. Es ist bislang noch wenig untersucht, ob sich die Werte auch für Studien mit kategorialen bzw. ordinalen Daten eignen, weswegen auch hier die Cut-Off-Werte mit Vorsicht zu interpretieren sind (vgl. Sauerwein, 2017; vgl. Kapitel 5.4.2).

5.4.2.2 Faktor-Scores

Die Analyse der Fragestellungen erfolgte auf Basis manifester Faktor-Scores, die aus den konfirmatorischen Faktorenanalysen extrahiert wurden. Dieses Vorgehen wurde gewählt, da die Stichprobe für eine latente Modellierung aller Konstrukte in einem Strukturgleichungsmodell mit mehreren Prädiktoren und unter Berücksichtigung der hierarchischen Datenstruktur nicht ausreicht. Die Anzahl der freien Parameter übersteigt in diesem Fall schnell die Anzahl der Cluster (36 Klassen), was Verzerrungen der Modellergebnisse zur Folge haben kann. In diesem Fall wird in Mplus eine Warnung ausgegeben, da noch wenig untersucht ist, welche Folgen dies auf die Parameterschätzung haben kann. Im Kontext von Mehrebenenanalysen konnte bereits gezeigt werden, dass mehr Parameter auf der Between-Ebene als Cluster in der Stichprobe die Schätzungen verzerren kann (Meuleman & Billiet, 2009).

Im Gegensatz zur Verwendung manifester Skalenmittelwerte führt die Verwendung der Faktor-Scores zu präziseren Schätzungen der individuellen Ausprägung der untersuchten Merkmale, da die Faktorladungen gewichtet in die Bildung der Werte eingehen und Messfehler der Indikatoren minimiert werden können (Moosbrugger & Kelava, 2020; Vieluf et al., 2013).

5.4.3 Regressionsmodelle

Für die Analyse der Fragestellungen wurden in Mplus Regressionsmodelle spezifiziert. Für die erste und dritte Fragestellung wurden die Zusammenhänge zwischen den Facetten des Implementationserfolgs und den interventionsspezifischen Schüler*innenoutcomes schrittweise über lineare autoregressive Regressionsmodelle untersucht. Für die zweite und vierte Fragestellung wurden multivariate Regressionsanalysen modelliert, um die Facetten des Implementationserfolgs durch individuelle Prädiktoren zu erklären. Da für die Nutzung der Promptkarten dichotome Items vorliegen, wurden gemäß der Standardeinstellungen in Mplus für diese Modelle Probit-Regressionen unter der WLSMV-Schätzmethode durchgeführt. Im Folgenden wird zunächst die Spezifikation der Regressionsmodelle für die erste und dritte Fragestellung und anschließend für die zweite und vierte Fragestellung beschrieben, da hier jeweils das gleiche Vorgehen unter Verwendung unterschiedlicher Stichproben gewählt wurde.

Alle Regressionsmodelle wurden in Mplus unter Berücksichtigung der hierarchischen Datenstruktur mit der Funktion *type = complex* spezifiziert (vgl. Kapitel 5.4.1). Für alle Analysen werden standardisierte Regressionskoeffizienten angegeben (stdyx-Standardisierung in Mplus).

Zur Untersuchung der Vorhersagekraft der Facetten des Implementationserfolgs für interventionsspezifische Schüler*innenoutcomes (erste und dritte Fragestellung) wurden einzelne autoregressive lineare Regressionsmodelle spezifiziert,

in denen die Ausprägung der Outcomes am Ende der Intervention jeweils als abhängige Variable modelliert wurden. Die Modelle wurden für jedes untersuchte Outcome einzeln spezifiziert, um die Anzahl freier Parameter im Vergleich zur Anzahl der Cluster nicht zu überschreiten (vgl. Kapitel 5.4.2.2). Die Prädiktoren wurden den Modellen schrittweise hinzugefügt. Im ersten Modell wurde lediglich das Eingangsniveau der abhängigen Variable als Kontrollvariable berücksichtigt. Im zweiten und dritten Modell wurden schrittweise die Nutzung der Promptkarten (Modell 2) und Akzeptanzdimensionen (Modell 3) als Prädiktoren eingeführt. Weitere Kontrollvariablen wurden nicht berücksichtigt, da angenommen wird, dass alle Einflüsse möglicher Kontrollvariablen (z. B. Geschlecht) bereits durch das Eingangsniveau erklärt werden.

Da die Akzeptanzdimensionen in der Gesamtstichprobe und in der Substichprobe mehrsprachiger Schüler*innen hohe Korrelationen aufweisen (vgl. Kapitel 5.3.2) und gemeinsam als Prädiktoren in die Regressionsmodelle eingehen, muss das Vorliegen möglicher Multikollinearität beachtet werden. Die Korrelationen der Akzeptanzdimensionen liegen nur knapp unter der häufig genannten kritischen Grenze von $r \geq .8$ (z. B. Field, 2009). Daher wurden zusätzliche Analysen zur Kollinearitätsdiagnostik durchgeführt. Der *Variance Inflation Factor* (VIF) liegt in den Regressionsmodellen für beide Akzeptanzdimensionen bei Werten von < 10 (Gesamtstichprobe: 2.67 – 2.88; Substichprobe mehrsprachiger Schüler*innen: 2.42 – 2.59). Die Werte liegen nach Field (2009) im akzeptablen Bereich und zeigen keine Multikollinearität an. Der Toleranzwert der Kollinearitätsdiagnostik liegt in den Modellen ebenfalls für beide Akzeptanzdimensionen im Rahmen der von Field (2009) beschriebenen Cut-Off-Werte von $> .2$ und deuten damit nicht auf Multikollinearität hin (Gesamtstichprobe: .35 – .38; Substichprobe mehrsprachiger Schüler*innen: .39 – .41). Die Akzeptanzdimensionen korrelieren mit der Nutzung der Promptkarten nur in geringem Maße (vgl. Kapitel 6.1.4), daher wird das Risiko möglicher Verzerrungen aufgrund von Multikollinearität als gering eingestuft.

Da die Faktor-Scores der Akzeptanzdimensionen und Schüler*innenoutcomes als kontinuierliche manifeste Variablen in die Modelle eingehen, wurde das ML-Schätzverfahren verwendet. Fehlende Werte werden dabei mittels des *Full-Information-Maximum-Likelihood*-Verfahrens (FIML) geschätzt (vgl. Kapitel 5.4.5). Zur Ermittlung robuster Standardfehler wurde in den Regressionsmodellen das Bootstrapping-Verfahren mit der Ziehung von 1000 Bootstrapping-Stichproben angewendet (Lai, 2018). Dies hat sich bei leichten Missspezifikationen und Modellabweichungen auch im Vergleich zum MLR-Verfahren als besonders robust erwiesen (Lai, 2018). In der vorliegenden Untersuchung zeigte sich insbesondere für die Selbstwirksamkeit der Lesestrategieanwendung und die Einstellungen zu Mehrsprachigkeit Abweichungen von den Cut-Off-Werten (vgl. Kapitel 5.3.3), die über die Verwendung des Bootstrap-Verfahren berücksichtigt werden sollen.

Die Analyse der zweiten und vierten Fragestellung erfolgte ebenfalls über schrittweise Regressionsmodelle, bei denen in zwei getrennten Modellen die Nutzung der Promptkarten und die Akzeptanzdimensionen durch die individuellen

Schüler*innenmerkmale erklärt wurden. Für die Vorhersage der Akzeptanzdimensionen wurde ein multivariates Regressionsmodell mit beiden Akzeptanzdimensionen als abhängige Variablen spezifiziert. Analog zu den Modellen der ersten und dritten Fragestellung wurde auch hier das ML-Schätzverfahren mit robusten Standardfehlern durch die Ziehung von 1000 Bootstrapping-Stichproben angewendet (Lai, 2018) und fehlende Werte modellbasiert durch FIML berücksichtigt. Da in der Gesamtstichprobe und in der Substichprobe mehrsprachiger Schüler*innen die höchste Korrelation von $r = .44$ ($p < .01$) weiter unter der kritischen Grenze von $r \geq .8$ (z.B. Field, 2009) liegt, kann davon ausgegangen werden, dass keine Multikollinearität zwischen den Konstrukten vorliegt. Die Korrelationen zwischen den Prädiktoren kann den Tabellen A7 und A8 im Anhang entnommen werden.

Da die Nutzung der Promptkarten als dichotome Items erhoben wurden, wurden für ihre Vorhersage in der zweiten und vierten Fragestellung gemäß der Standardeinstellungen in Mplus Probit-Regressionen unter der WLSMV-Schätzmethode durchgeführt. Probit- und Logit-Regressionen, die standardmäßig zur Erklärung abhängiger dichotomer Variablen verwendet werden, liegt die Annahme zugrunde, „dass eine latente [...] Variable y^* existiert, die dazu führt, dass Personen Zustände annehmen oder Entscheidungen treffen, deren Auftreten empirisch als dichotome Variable y beobachtet werden kann“, wenn ein bestimmter Schwellenwert überschritten wird (Best & Wolf, 2012, S. 379). Die latente Variable y^* wird dabei linear modelliert, sodass die standardisierten Regressionskoeffizienten den Effekt des Prädiktors auf die latente Variable y^* angeben (Best & Wolf, 2012). Zwischen der latenten und beobachteten Variable besteht jedoch ein nicht-linearer Zusammenhang, mit dem die Wahrscheinlichkeit des Auftretens der dichotomen Variable beschrieben wird. Die Regressionskoeffizienten können daher inhaltlich nur über das Signifikanzniveau und das Vorzeichen interpretiert werden, die Stärke des Werts kann inhaltlich nicht eingeordnet werden (Best & Wolf, 2012).

Eine inhaltliche Interpretation der Probit-Regressionkoeffizienten, die z. B. nicht in häufig verwendete Odds Ratios umgewandelt werden können, ist durch die Berechnung des *average marginal effects* (AME) möglich. Der AME gibt die durchschnittliche Veränderung der Wahrscheinlichkeit in Prozentpunkten an, die abhängige Variable zu beobachten, wenn sich die Kovariate um eine Einheit ändert und alle anderen Variablen konstant bleiben (Leeper, 2021). Best und Wolf (2012, S. 383) beschreiben den Vorteil von AMEs darin, „eine intuitive Interpretation zu ermöglichen“. Sie können zudem zwischen geschachtelten Probit-Modellen verglichen werden und gelten als robust gegenüber Abweichungen von der Normalverteilung. Es muss jedoch beachtet werden, dass der AME einen Durchschnittswert angibt, der keinem linearen Verlauf der Wahrscheinlichkeit entspricht. Für die Vorhersage der Nutzung der Promptkarten wurden neben den standardisierten Regressionskoeffizienten auch AMEs berechnet. Folgende Formel wurde hierbei über einen *model constraint* in Mplus angegeben (Keyserlingk et al., 2020):

$$AME = \phi(-a + \beta x)x$$

Hierbei stellt a den Schwellenwert der dichotomen Variable dar, β den Regressionskoeffizienten und x den Mittelwert des Prädiktors. Diese werden mit der Standardnormalverteilung (ϕ) multipliziert.

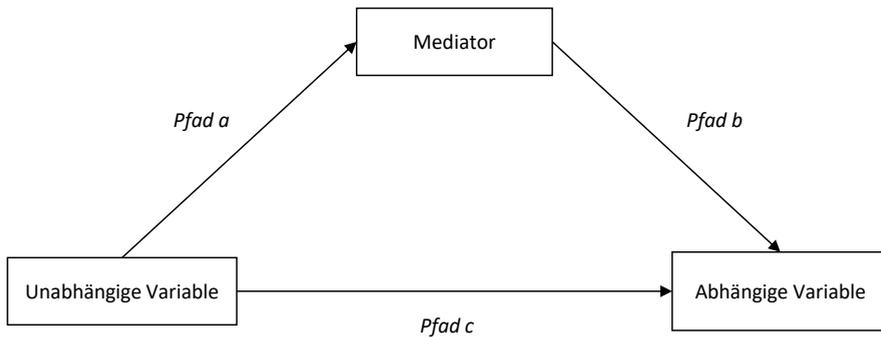
5.4.4 Mediationsmodelle

Zur Beantwortung der Forschungsfragen werden neben den direkten Zusammenhängen der Variablen auch indirekte Zusammenhänge in Mediationsmodellen untersucht. Im Folgenden werden zunächst theoretische Grundlagen zur Untersuchung von Mediationsmodellen vorgestellt und darauf aufbauend ein Vorgehen zum Testen indirekter Zusammenhänge für die vorliegende Arbeit abgeleitet.

Ein Zusammenhang zwischen zwei Variablen kann durch weitere Mechanismen vermittelt werden, die in der psychologischen Forschung anhand von Mediationsanalysen untersucht werden. Die Kausalität der Wirkungszusammenhänge ist methodisch hierbei eine wichtige Voraussetzung. Die medierende Variable wird dabei kausal durch die unabhängige Variable erklärt und die abhängige Variable wiederum ebenfalls kausal durch die medierende Variable (Hayes, 2018). Da jedoch nicht immer längsschnittliche Daten vorliegen und die Datenlage dahingehend eingeschränkt ist, kann diese Voraussetzung oft nicht erfüllt werden. Nach Hayes (2018) können Mediationshypothesen dennoch überprüft werden, wenn ein kausaler Zusammenhang zumindest theoretisch plausibilisiert werden kann, obwohl die Datenerhebung keine zeitliche Abfolge wiedergibt.

Im klassischen Mediationsmodell nach Baron und Kenny (1986) wird der Zusammenhang zwischen einer unabhängigen und einer abhängigen Variable durch eine dritte, medierende Variable vermittelt und kausal erklärt. Hierbei wird zwischen einem *totalen Effekt* (c) der unabhängigen Variable auf die abhängige Variable unterschieden, der den alleinigen Zusammenhang zwischen der abhängigen und der unabhängigen Variable beschreibt, und dem *direkten Effekt* (c') der unabhängigen auf die abhängige Variable, welcher unter Kontrolle der medierenden Variable bestehen bleibt (Baron & Kenny, 1986). Im klassischen Vorgehen nach Baron und Kenny (1986) werden vier Schritte beschrieben, mit denen Mediationsanalysen anhand von Regressionsgleichungen überprüft werden (vgl. Abbildung 4).

Abbildung 4: Mediationsmodell nach Baron und Kenny (1986)



Im ersten Schritt muss die unabhängige Variable die abhängige Variable signifikant vorhersagen. Im zweiten Schritt muss die unabhängige Variable die mediiierende Variable signifikant vorhersagen (Pfad a). Im dritten Schritt muss die mediiierende Variable einen signifikanten Zusammenhang zur abhängigen Variable aufweisen (Pfad b), während auch die unabhängige Variable als Prädiktor in die Gleichung miteingeht. Sind alle drei Bedingungen gegeben, muss im vierten Schritt überprüft werden, ob der Zusammenhang der unabhängigen Variable auf die abhängige Variable unter Kontrolle des Mediators nicht mehr gegeben ist (Pfad c). Ist dies der Fall, liegt eine vollständige Mediation vor. Reduziert sich der Zusammenhang der unabhängigen und abhängigen Variable unter Kontrolle des Mediators, bleibt jedoch zusätzlich bestehen, wird dies als partielle Mediation bezeichnet (Preacher & Hayes, 2004). In diesem Fall erklärt die mediiierende Variable den Zusammenhang zwischen der unabhängigen und abhängigen Variable nicht vollständig und weitere mediiierende Mechanismen können vorliegen (Baron & Kenny, 1986). Die Ausprägung des indirekten Zusammenhangs wird schließlich als Produkt der Pfade a und b berechnet und auf Signifikanz geprüft.

Neben der Untersuchung eines Mediators zwischen zwei Variablen können auch mehrere potenziell mediiierende Variablen in einem Mediationsmodell mit *multiplen Mediatoren* analysiert werden, das als Erweiterung des klassischen Mediationsmodells gilt (MacKinnon et al., 2007). Hierbei ist zwischen *parallelen* und *seriellen* Mediatoren zu unterscheiden (Hayes, 2018). In einem Mediationsmodell mit multiplen parallelen Mediatoren wird die abhängige Variable durch mehrere Mediatoren vorhergesagt, die in keinem kausalen Zusammenhang miteinander stehen und kein Mediator einen anderen Mediator vorhersagt. Ein Mediationsmodell mit *seriellen* multiplen Mediatoren modelliert dagegen den kausalen Zusammenhang zwischen zwei Variablen über mehrere mediiierende Variablen, die ebenfalls in kausalem Zusammenhang stehen. Durch die Integration mehrerer potenzieller Mediatoren in einem Modell kann die Ausprägung spezifischer indirekter Zusammenhänge unter der Kontrolle der anderen Mediatoren geschätzt und

somit der eigenständige Beitrag eines jeden Mediators isoliert betrachtet werden (Hayes, 2018). Zu beachten gilt jedoch, dass der Einbezug vieler potentieller Mediatoren auch zu einer größeren Stichprobenvarianz führt, die mit einer geringeren Power zur Detektion indirekter Zusammenhänge einhergehen kann (Hayes, 2018). Weisen die indirekten Zusammenhänge zudem unterschiedliche Vorzeichen auf, können sie sich gegenseitig aufheben und in keinem totalen indirekten Zusammenhang resultieren, obwohl spezifische indirekte Effekte zu verzeichnen sind (Hayes, 2018).

Das klassische Vorgehen zur Überprüfung von Mediationshypothesen nach Baron und Kenny (1986), das in der psychologischen Forschung etabliert ist, setzt einen direkten Zusammenhang der unabhängigen Variable auf die abhängige Variable voraus. Diese Voraussetzung wird jedoch zunehmend kritisiert, da indirekte Zusammenhänge auch dann bestehen können, wenn sich kein direkter Effekt zwischen der unabhängigen und der abhängigen Variable zeigt (MacKinnon et al., 2007; Zhao et al., 2010). Zhao et al. (2010) argumentieren, dass das Bestehen des indirekten Zusammenhangs und dessen Stärke der Hauptindikator zur Beurteilung einer Mediation sein sollte und nicht das Ausbleiben eines direkten Zusammenhangs unter Kontrolle des Mediators. Ein direkter Zusammenhang kann z. B. trotz indirektem Zusammenhang ausbleiben, wenn sich direkte und indirekte Zusammenhänge mit unterschiedlichen Vorzeichen und ähnlicher Stärke gegenseitig aufheben (Zhao et al., 2010). Hierzu ist es von Bedeutung, die Art der Mediation zu berücksichtigen. Zhao et al. (2010) unterscheiden drei Arten von Mediationen. Als *complementary mediation* beschreiben die Autor*innen eine Mediation mit signifikantem indirekten Zusammenhang und direktem Zusammenhang, die die gleichen Vorzeichen aufweisen. Bei der *competitive mediation* dagegen haben ein signifikanter indirekter und direkter Zusammenhang unterschiedliche Vorzeichen. Eine dritte Form der Mediation beschreiben Zhao und Kolleg*innen (2010) als *indirect-only-mediation*, die dann zu klassifizieren ist, wenn zwar der indirekte Zusammenhang signifikant ist, jedoch kein signifikanter direkter Zusammenhang besteht.

Da das klassische Vorgehen nach Baron und Kenny (1986) mit der Voraussetzung eines signifikanten totalen Zusammenhangs die Problematik mit sich bringt, dass signifikante indirekte Zusammenhänge unentdeckt bleiben könnten, soll ein signifikanter Zusammenhang zwischen der unabhängigen und der abhängigen Variable in der vorliegenden Arbeit keine Voraussetzung zur Überprüfung von Mediationshypothesen darstellen. Stattdessen sollen die Schritte zwei bis drei nach Baron und Kenny (1986) überprüft und sofern diese erfüllt sind, die Signifikanz der indirekten Zusammenhänge berechnet werden.

In linearen Regressionsmodellen können Veränderungen der Parameter in schrittweise erweiterten Modellen Hinweise auf mediiierende Zusammenhänge geben. In Probit-Regressionen ist dies nicht möglich, da die Ausprägung von Probit-Koeffizienten nicht zwischen Modellen verglichen werden kann (Best & Wolf, 2012). In Probit-Modellen wird die Fehlervarianz fixiert, um die latente abhängige

Variable y^* schätzen zu können. Dies bedeutet, dass anders als in linearen Regressionsgleichungen, die Gesamtvarianz zwischen Probit-Modellen der gleichen abhängigen Variablen variieren kann und die Fehlervarianz nicht mit steigender erklärter Varianz sinkt, sondern konstant bleibt (Best & Wolf, 2012). Die Regressionskoeffizienten geschachtelter Probit-Modelle können demnach nicht sinnvoll miteinander verglichen werden, da unterschiedliche Skalierungen zugrunde liegen (Best & Wolf, 2012). Zum Vergleich der Zusammenhänge in hierarchischen Probit-Modellen wird in der vorliegenden Arbeit daher der *AME* herangezogen.

Zur Überprüfung der Signifikanz indirekter Zusammenhänge wird das 95 %-Konfidenzintervall unter dem Bootstrapping-Verfahren mit der Ziehung von 1000 Bootstrapping-Stichproben interpretiert (Lai, 2018; MacKinnon et al., 2007). Das Bootstrapping-Verfahren wurde gewählt, da die indirekten Effekte aus einer nicht-linearen Funktion zweier Variablen hervorgehen (Muthén et al., 2016). Die indirekten Effekte sind entsprechend häufig nicht normalverteilt, wie bei der Berechnung des regulär symmetrischen Konfidenzintervalls angenommen wird (Muthén et al., 2016). Symmetrische Konfidenzintervalle zeichnen sich durch eine einheitliche Entfernung der oberen und unteren Grenze vom geschätzten Wert aus und sind angemessen, wenn die Daten normalverteilt sind (Muthén et al., 2016). Bootstrapping-Verfahren gelten als Alternative zur Schätzung symmetrischer Konfidenzintervalle und können für die Schätzung indirekter Zusammenhänge geeigneter sein, da sie asymmetrische Konfidenzintervalle darstellen. Asymmetrische Konfidenzintervalle mit Bootstrap führen dabei zu präziseren Ergebnissen als symmetrische Konfidenzintervalle (MacKinnon et al., 2004; Muthén et al., 2016; Shrout & Bolger, 2002). Schließt die obere und untere Grenze dieses Konfidenzintervalls Null nicht ein, ist der indirekte Effekt auf dem Niveau $p < .05$ signifikant.

5.4.5 Umgang mit fehlenden Werten

Die untersuchten Konstrukte gingen über die Faktor-Scores als manifeste Variablen in die Analysen ein (mit Ausnahme des Leseverstehens und der kognitiven Grundfähigkeiten – hier wurden Summenwerte verwendet; vgl. Kapitel 5.3.3). Da die Faktor-Scores unter Verwendung des WLSMV-Verfahrens in Mplus berechnet wurden und dieses Verfahren den Einbezug fehlender Werte nicht erlaubt, weisen Schüler*innen, die an einem der Erhebungstage nicht anwesend waren, fehlende Werte auf diesen Konstrukten auf. Da die Faktor-Scores kontinuierliche Variablen darstellen, können fehlende Werte in den Regressionsmodellen, die mit dem ML-Verfahren geschätzt werden, mittels FIML einbezogen werden. Bei diesem Verfahren werden alle beobachteten Daten zur Schätzung der Populationsparameter herangezogen (Lüdtke & Robitzsch, 2011). Fehlende Werte müssen daher nicht ausgeschlossen werden und präzisere Modellschätzungen sind möglich. Lediglich Fälle, die fehlende Werte auf allen Variablen der jeweiligen Modelle aufweisen, können nicht in die Analysen eingeschlossen werden.

Der Anteil fehlender Faktor-Scores liegt für die Facetten des Implementationserfolgs in der Gesamtstichprobe zwischen 6.9% und 8.7% und in der Substichprobe mehrsprachiger Schüler*innen zwischen 7.7% und 8.8%. Die Faktor-Scores der individuellen Lernvoraussetzungen und interventionsspezifischen Schüler*innenoutcomes, die in die Analysen eingingen, weisen in der Gesamtstichprobe zwischen 6.5% und 8.7% und in der Substichprobe mehrsprachiger Schüler*innen zwischen 7.0% und 8.6% fehlende Werte auf. Bei den Kontrollvariablen fehlen in der Gesamtstichprobe zwischen 0.5% und 6.8% der Werte, in der Substichprobe mehrsprachiger Schüler*innen zwischen 0.2% und 7.2%.

Der Anteil fehlender Werte ist insgesamt als gering einzuschätzen. Fehlende Werte wurden hauptsächlich dadurch verursacht, dass Schüler*innen an einem der Erhebungstage abwesend waren. Daher kann angenommen werden, dass die fehlenden Werte nicht systematisch entstanden sind und *Missing at Random* als Voraussetzung zur Verwendung der FIML-Prozedur gegeben ist (Lüdtke & Robitzsch, 2011).

5.4.6 Signifikanztestung

Zur Analyse der Fragestellungen wurden Hypothesen und Forschungsfragen aufgestellt. Obwohl auch gerichtete Hypothesen formuliert wurden, werden für alle Analysen zweiseitige Signifikanztestungen ($\alpha = 5\%$) vorgenommen. Dies ermöglicht, die Signifikanztestung innerhalb der Modelle und über die Modelle hinweg einheitlich zu halten. Da das Forschungsfeld zudem insgesamt noch dünn besiedelt ist, soll auch die Interpretation unerwarteter Ergebnisse möglich bleiben.

Bei multiplen Signifikanztestungen an einer Stichprobe, wie sie in der vorliegenden Arbeit vorgenommen wurden, muss auch das Risiko einer Alpha-Fehler-Kumulierung berücksichtigt werden. Aufgrund der kleinen Stichprobe (insbesondere in der Subgruppe mehrsprachiger Schüler*innen) und der verminderten statistischen Power, signifikante Effekte zu entdecken, wurde jedoch entschieden, keine Alpha-Fehler-Korrektur vorzunehmen. Die zweiseitige Signifikanztestung stellt jedoch einen konservativen Maßstab zur Bewertung der Signifikanzen dar (Pillemer, 1991), mit dem in der vorliegenden Arbeit die Gefahr einer Alpha-Fehler-Kumulierung vermindert werden soll.

6 Ergebnisse

Im folgenden Kapitel werden die Ergebnisse der statistischen Analysen dargestellt. Zur Beschreibung der Datengrundlage werden in Kapitel 6.1 zunächst die deskriptiven Kennwerte und bivariaten Zusammenhänge der untersuchten Konstrukte angegeben. In den Kapiteln 6.2 bis 6.5 sind die Ergebnisse der abgeleiteten Fragestellungen beschrieben.

6.1 Deskriptive Statistik

Im ersten Schritt werden die deskriptiven Kennwerte der untersuchten Variablen dargestellt. Zur besseren Interpretation der Kennwerte werden hierbei die Mittelwerte der eingesetzten Skalen berichtet, bevor für die bivariaten Zusammenhänge und Analysen der Fragestellungen die Faktor-Scores der latenten Konstrukte verwendet werden (vgl. Kapitel 5.4.2.2). Für die dichotomen Variablen werden die Häufigkeitsverteilungen angegeben. Alle Angaben erfolgen für die Gesamtstichprobe sowie für die Substichprobe mehrsprachiger Schüler*innen, die im Rahmen der dritten und vierten Fragestellung zur Nutzung der Promptkarten in einer anderen Sprache separat untersucht werden. Die Substichprobe mehrsprachiger Schüler*innen bezieht sich nur auf Schüler*innen, die die Promptkarten in (mindestens) einer ihrer Familiensprachen vorliegen hatten (vgl. Kapitel 5.3.1).

6.1.1 Nutzung der Promptkarten

In Tabelle 11 sind die Häufigkeiten der selbstberichteten Nutzung der Promptkarten dargestellt. In der Gesamtstichprobe gab die Mehrheit der Schüler*innen an, die Promptkarten während der Intervention genutzt zu haben (80.4%). Ebenso gab ein Großteil der mehrsprachigen Schüler*innen an, die Promptkarten während der Intervention in einer anderen Sprache als Deutsch genutzt zu haben (64.5%).

Tabelle 11: Häufigkeitsverteilungen zur Nutzung der Promptkarten

	Nutzung der Promptkarten ¹		Nutzung der Promptkarten in einer anderen Sprache ²	
	<i>n</i>	(%)	<i>n</i>	(%)
Ja	510	(8.4)	278	(64.5)
Nein	69	(10.9)	110	(25.5)
Fehlend	55	(8.7)	43	(10.0)
Gesamt	634	(100.0)	431	(100.0)

Anmerkung. ¹Gesamtstichprobe, ²Substichprobe mehrsprachiger Schüler*innen.

6.1.2 Akzeptanz

Die Mittelwerte der Akzeptanzdimensionen sind in Tabelle 12 dargestellt. Insgesamt zeigt sich, dass die Schüler*innen der Gesamtstichprobe die Intervention in hohem Maße akzeptieren, da die durchschnittliche Ausprägung für beide Dimensionen über dem theoretischen Mittelwert der vierstufigen Skala liegt. Die Ausprägung der affektiven Akzeptanz ist hierbei höher als der Mittelwert der kognitiven Akzeptanz. Dieses Muster zeigt sich auch in der Substichprobe mehrsprachiger Schüler*innen, die ebenfalls ein hohes Maß an affektiver und kognitiver Akzeptanz aufweisen. In der Gruppe mehrsprachiger Schüler*innen sind die Werte der Akzeptanzdimensionen deskriptiv höher als in der Gesamtstichprobe. Inwiefern die Mehrsprachigkeit der Schüler*innen das Akzeptanzniveau vorhersagt, wird in Fragestellung 2 untersucht (vgl. Kapitel 6.3).

Tabelle 12: Deskriptive Kennwerte der Akzeptanzdimensionen

Skala	Gesamtstichprobe			Substichprobe mehrsprachiger Schüler*innen		
	<i>M</i>	<i>(SD)</i>	<i>N</i>	<i>M</i>	<i>(SD)</i>	<i>N</i>
Affektive Akzeptanz	3.11	(0.81)	588	3.18	(0.77)	398
Kognitive Akzeptanz	2.88	(0.98)	586	2.98	(0.95)	397

6.1.3 Individuelle Schüler*innenmerkmale

Tabelle 13 stellt die Häufigkeiten der individuellen Schüler*innenmerkmale dar, die als dichotome Variable erhoben wurden.

Tabelle 13: Häufigkeiten der dichotomen individuellen Merkmale

Merkmal	Gesamtstichprobe		Substichprobe mehrsprachiger Schüler*innen	
	<i>n</i>	(%)	<i>n</i>	(%)
<i>Mehrsprachigkeit</i>				
Monolingual (0)	181	(28.6)		
Mehrsprachig (1)	449	(70.8)		
Fehlend	4	(0.6)		
<i>Geschlecht</i>				
Mädchen (0)	321	(50.6)	221	(51.3)
Jungen (1)	310	(48.9)	209	(48.5)
Fehlend	3	(0.5)	1	(0.2)
<i>Anzahl an Büchern im Haushalt</i>				
Weniger als 26 Bücher im Haushalt	221	(34.9)	173	(40.1)
Mehr als 25 Bücher im Haushalt	370	(58.4)	230	(53.4)
Fehlend	43	(6.8)	28	(6.5)
Gesamt	634	(100.0)	431	(100.0)

In Tabelle 14 sind die deskriptiven Kennwerte der eingesetzten Leistungstests und Skalen zur Erfassung individueller Voraussetzungen und Outcomes angegeben. Die kognitiven Grundfähigkeiten wurden als Kontrollvariable eingesetzt. Das Leseverstehen, die Lesefreude, die Selbstwirksamkeit der Lesestrategieanwendung sowie die Einstellungen zu Mehrsprachigkeit wurden im Prä-Post-Design zu Beginn und unmittelbar nach Abschluss der Intervention erhoben.

Über den Verlauf der Intervention verbessern sich in der Gesamtstichprobe das Leseverstehen ($t(541) = 15.48$, $p < .001$, $d = 0.67$), die Selbstwirksamkeit der Lesestrategieanwendung ($t(546) = 7.80$, $p < .001$, $d = 0.33$) und die Einstellungen zu Mehrsprachigkeit ($t(543) = 4.01$, $p < .001$, $d = 0.17$) signifikant. Die Lesefreude bleibt im Verlauf der Intervention stabil ($t(545) = 0.74$, $p = .458$, $d = 0.03$).

In der Substichprobe mehrsprachiger Schüler*innen zeigt sich das gleiche Ergebnismuster. Das Leseverstehen ($t(368) = 12.15$, $p < .001$, $d = 0.63$), die Selbstwirksamkeit der Lesestrategieanwendung ($t(372) = 5.47$, $p < .001$, $d = 0.28$) und Einstellungen zu Mehrsprachigkeit ($t(369) = 3.573$, $p < .001$, $d = 0.19$) verbessern sich signifikant, während die Lesefreude stabil bleibt ($t(371) = 1.55$, $p = .121$, $d = 0.08$).

Tabelle 14: Deskriptive Kennwerte der Leistungstests und eingesetzten Skalen zur Erfassung individueller Merkmale

Konstrukt	Gesamtstichprobe			Substichprobe mehrsprachiger Schüler*innen		
	<i>M</i>	(<i>SD</i>)	<i>N</i>	<i>M</i>	(<i>SD</i>)	<i>N</i>
Kognitive Grundfähigkeiten	8.66	(2.73)	594	8.40	(2.71)	400
Leseverstehen (Prä)	12.99	(4.53)	579	12.41	(4.28)	394
Leseverstehen (Post)	14.53	(4.48)	588	13.90	(4.23)	398
Lesefreude (Prä)	3.05	(0.82)	582	3.05	(0.80)	397
Lesefreude (Post)	3.01	(0.82)	591	2.98	(0.84)	400
Selbstwirksamkeit Lesestrategieanwendung (Prä)	3.07	(0.68)	582	3.06	(0.69)	397
Selbstwirksamkeit Lesestrategieanwendung (Post)	3.26	(0.68)	592	3.23	(0.68)	401
Einstellungen zu Mehrsprachigkeit (Prä)	3.01	(0.78)	581	3.10	(0.75)	395
Einstellungen zu Mehrsprachigkeit (Post)	3.16	(0.75)	590	3.25	(0.70)	399

6.1.4 Bivariate Zusammenhänge der untersuchten Konstrukte

Nachdem die deskriptiven Statistiken der interessierenden Konstrukte dargestellt wurden, werden im Folgenden die bivariaten Zusammenhänge analysiert. Hierzu werden bivariate Korrelationen nach Pearson berichtet. Für die eingesetzten Likert-Skalen werden die aus den Faktorenanalysen extrahierten Faktor-Scores verwendet, um Messfehler in manifesten Analysen zu verringern (vgl. Kapitel 5.4.2.2).

Tabelle 15: Bivariate Korrelationen der Akzeptanzdimensionen und Nutzung der Promptkarten

Variable	Nutzung der Promptkarten ¹			Nutzung der Promptkarten in einer anderen Sprache ²		
	<i>r</i>	<i>p</i>	<i>N</i>	<i>r</i>	<i>p</i>	<i>N</i>
Affektive Akzeptanz	.27	< .001	578	.14	.006	387
Kognitive Akzeptanz	.24	< .001	578	.13	.003	387

Anmerkungen. ¹Gesamtstichprobe; ²Substichprobe mehrsprachiger Schüler*innen.

In Tabelle 15 sind die bivariaten Zusammenhänge zwischen den Akzeptanzdimensionen und der Nutzung der Promptkarten dargestellt. Beide Akzeptanzdimensionen weisen signifikant positive und ähnlich ausgeprägte Zusammenhänge zu der Nutzung der Promptkarten in der Gesamtstichprobe auf. Auch die Nutzung der Promptkarten in einer anderen Sprache hängt in der Substichprobe der mehrsprachigen Lernenden signifikant positiv mit den Akzeptanzdimensionen zusammen. Die Zusammenhänge sind etwas geringer ausgeprägt als in der Gesamtstichprobe.

Tabelle 16: Bivariate Korrelationen der untersuchten individuellen Merkmale und der Facetten des Implementationserfolgs

	Affektive Akzeptanz ¹		Kognitive Akzeptanz ¹		Nutzung der Promptkarten ¹		Nutzung der Promptkarten in einer anderen Sprache ²	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Mehrsprachigkeit	.12	.005	.18	< .001	.06	.160		
<i>Kontrollvariablen</i>								
Geschlecht (Jungen)	-.10	.018	-.07	.089	-.15	< .001	-.02	.632
Kognitive Grundfähigkeiten	-.04	.301	-.14	< .001	-.004	.931	.02	.668
Mehr als 25 Bücher im Haushalt	-.05	.228	-.09	.029	-.10	.020	.06	.295
<i>Lerngegenstandsbezogene Voraussetzungen</i>								
Leseverstehen (Prä)	-.10	.020	-.18	< .001	-.07	.087	.06	.232
Lese Freude (Prä)	.25	< .001	.11	.008	.08	.058	.10	.057
Selbstwirksamkeit Lese-strategieranwendung (Prä)	.11	.010	.05	.276	.003	.951	.11	.033
Einstellungen zu Mehrsprachigkeit (Prä)	.21	< .001	.24	< .001	.10	.023	.13	.012
<i>Interventionsspezifische Outcomes</i>								
Leseverstehen (Post)	-.11	.010	-.18	< .001	-.04	.365	.07	.169
Lese Freude (Post)	.26	< .001	.15	< .001	.10	.020	.19	< .001
Selbstwirksamkeit Lese-strategieranwendung (Post)	.25	< .001	.14	< .001	.02	.580	.07	.158
Einstellungen zu Mehrsprachigkeit (Post)	.27	< .001	.28	< .001	.12	.005	.12	.023

Anmerkungen. ¹ $N_{\text{Min}} = 533$, $N_{\text{Max}} = 590$; ² $N_{\text{Min}} = 362$, $N_{\text{Max}} = 388$.

Tabelle 16 gibt die Zusammenhänge zwischen den individuellen Merkmalen der Schüler*innen und den Akzeptanzdimensionen sowie der Nutzung der Promptkarten an. Die affektive und kognitive Akzeptanz korrelieren signifikant positiv mit dem Merkmal Mehrsprachigkeit. Zwischen der Mehrsprachigkeit und Nutzung der Promptkarten zeigt sich dagegen kein signifikanter Zusammenhang. Hinsichtlich der ausgewählten Kontrollvariablen zeigen sich für die Akzeptanzdimensionen unterschiedliche Muster. Die affektive Akzeptanz korreliert signifikant negativ mit dem Geschlecht. Mädchen weisen demnach eine höhere Ausprägung der affektiven Akzeptanz auf als Jungen. Die kognitive Akzeptanz korreliert signifikant negativ mit den kognitiven Grundfähigkeiten sowie mit der Anzahl an Büchern im Haushalt. Schüler*innen, die eine geringere Ausprägung der kognitiven Grundfähigkeit erzielt haben, bewerten die Intervention demnach als wirksamer. Gleiches gilt für Schüler*innen, die weniger als 26 Bücher im Haushalt angegeben haben. Die Nutzung der Promptkarten zeigt einen negativen Zusammenhang zum Geschlecht und der Anzahl an Büchern im Haushalt.

Die affektive Akzeptanz hängt mit allen untersuchten lerngegenstandsbezogenen Merkmalen signifikant zusammen. Positive Zusammenhänge zeigen sich hier für die Lesefreude, die Selbstwirksamkeit der Lesestrategieanwendung und für die Einstellungen zu Mehrsprachigkeit. Negative Zusammenhänge sind für das Leseverstehen zu verzeichnen. Bei allen lerngegenstandsbezogenen Merkmalen hängt sowohl die Ausprägung zu Beginn der Intervention als auch die Ausprägung am Ende der Intervention signifikant mit der affektiven Akzeptanz zusammen. Auch die kognitive Akzeptanz hängt mit fast allen lerngegenstandsbezogenen Merkmalen signifikant zusammen. Positive Zusammenhänge zeigen sich hierbei für die Lesefreude und die Einstellungen zu Mehrsprachigkeit vor und nach der Intervention und zur Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention. Die Ausprägung der Selbstwirksamkeit vor der Intervention hängt allerdings nicht signifikant mit der kognitiven Akzeptanz zusammen. Negative Korrelationen zeigen sich, wie bereits für die affektive Akzeptanz, auch hier für das Leseverstehen vor und nach der Intervention. Insbesondere in Hinblick auf die Kontrollvariablen liegen für die Akzeptanzdimensionen unterschiedliche Ergebnisse vor, was eine getrennte Auswertung der Dimensionen unterstützt.

Die Nutzung der Promptkarten und die Lesefreude am Ende der Intervention hängen positiv zusammen. Darüber hinaus zeigen sich positive Zusammenhänge zwischen der Nutzung der Promptkarten und den Einstellungen zu Mehrsprachigkeit zu Beginn und am Ende der Intervention. Die anderen lerngegenstandsbezogenen Merkmale hängen nicht signifikant mit der Nutzung der Promptkarten während der Intervention zusammen.

In der Substichprobe mehrsprachiger Schüler*innen korreliert die Nutzung der Promptkarten in einer anderen Sprache mit keiner der Kontrollvariablen signifikant. Bezüglich der lerngegenstandsbezogenen Voraussetzungen zeigen sich positive Zusammenhänge zur Selbstwirksamkeit der Lesestrategieanwendung und den Einstellungen zu Mehrsprachigkeit. Die Einstellungen zu Mehrsprachigkeit hängen

auch am Ende der Intervention positiv mit der Nutzung der Promptkarten in einer anderen Sprache zusammen. Zusätzlich zeigt sich ein positiver Zusammenhang der Lesefreude am Ende der Intervention zur Nutzung der Promptkarten in einer anderen Sprache. Obwohl die Zusammenhänge in der Substichprobe mehrsprachiger Schüler*innen nicht signifikant sind, fällt vergleichend auf, dass die Nutzung der Promptkarten in einer anderen Sprache (entgegen den Zusammenhängen in der Gesamtstichprobe) positiv mit der Ausprägung des Leseverstehens vor und nach der Intervention zusammenhängt.

6.2 Fragestellung 1: Sagen die Facetten des Implementationserfolgs interventionsspezifische Schüler*innenoutcomes vorher?

Die erste Fragestellung der vorliegenden Arbeit soll klären, ob die Nutzung der Promptkarten und Akzeptanzdimensionen als Facetten des Implementationserfolgs zur Erklärung des Lernerfolgs auf kognitiver und affektiv-motivationaler Ebene beitragen. Zur Überprüfung der direkten Zusammenhänge werden schrittweise autoregressive Regressionsmodelle spezifiziert (vgl. Kapitel 5.4.3). Neben dem Leseverstehen als kognitive Zielvariable werden die Lesefreude, die Selbstwirksamkeit der Lesestrategieanwendung und die Einstellungen zu Mehrsprachigkeit als affektiv-motivationale Merkmale in den Blick genommen. Die genannten Schüler*innenoutcomes werden in einzelnen Regressionsmodellen als abhängige Variablen modelliert, die schrittweise durch die Nutzung der Promptkarten und die Akzeptanzdimensionen erklärt werden sollen (vgl. Kapitel 6.2.1). Im ersten Modell wird zunächst nur für das Eingangsniveau der zu untersuchenden Variable kontrolliert. In Modell II wird die Nutzung der Promptkarten hinzugefügt und in Modell III die Akzeptanzdimensionen. Neben den direkten Zusammenhängen wird zusätzlich die vermittelnde Rolle der Akzeptanz zwischen der Nutzung der Promptkarten und der Schüler*innenoutcomes untersucht. Hierfür werden Mediationsmodelle spezifiziert (vgl. Kapitel 6.2.2).

6.2.1 Direkte Vorhersagekraft der Nutzung der Promptkarten und Akzeptanzdimensionen für interventionsspezifische Schüler*innenoutcomes

Leseverstehen. Tabelle 17 gibt die Ergebnisse der schrittweisen Regressionsmodelle an, um das Leseverstehen durch die Facetten des Implementationserfolgs zu erklären. 68% der Varianz der abhängigen Variable werden im ersten Modell bereits durch die Ausprägung des Leseverstehens zu Beginn der Intervention erklärt. Durch Hinzunahme der Nutzung der Promptkarten als Prädiktor in Modell II sowie im nächsten Schritt der Akzeptanzdimensionen in Modell III kann keine weitere Varianz erklärt werden. Die Facetten des Implementationserfolgs stellen

damit keine signifikanten Prädiktoren für den Lernerfolg im Leseverstehen dar. Entgegen der Erwartung ist der Koeffizient der kognitiven Akzeptanz negativ, dieser Zusammenhang ist jedoch nicht signifikant.

Tabelle 17: Lineare Regressionsmodelle zur Vorhersage des Leseverstehens am Ende der Intervention

	Modell I ¹			Modell II ²			Modell III ²		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Prä-Wert	.82	(.01)	<.001	.82	(.01)	<.001	.82	(.02)	<.001
Nutzung der Promptkarten				.01	(.02)	.675	.02	(.03)	.464
Affektive Akzeptanz							.004	(.04)	.925
Kognitive Akzeptanz							-.05	(.04)	.250
R ²	.68			.68			.68		

Anmerkungen. ¹N = 625; ²N = 626.

Lesefreude. In Tabelle 18 sind die Ergebnisse der schrittweisen Regressionsmodelle zur Erklärung der Lesefreude am Ende der Intervention abgebildet. Das Eingangsniveau der Lesefreude erklärt im ersten Modell 52 % der Varianz in der abhängigen Variable. Durch Hinzunahme der Nutzung der Promptkarten können in Modell II keine weiteren Varianzanteile aufgeklärt werden. Die Hinzunahme der Akzeptanzdimensionen in Modell III führt nur zu einer zusätzlichen Varianzaufklärung von 1 % und die untersuchten Variablen stellen keine signifikanten Prädiktoren dar. Die Lesefreude am Ende der Intervention wird demnach unter Kontrolle des Eingangsniveaus nicht signifikant durch Facetten des Implementationserfolgs erklärt.

Tabelle 18: Lineare Regressionsmodelle zur Vorhersage der Lesefreude am Ende der Intervention

	Modell I			Modell II			Modell III		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Prä-Wert	.72	(.03)	<.001	.72	(.03)	<.001	.70	(.03)	<.001
Nutzung der Promptkarten				.04	(.03)	.265	.02	(.03)	.655
Affektive Akzeptanz							.08	(.05)	.076
Kognitive Akzeptanz							-.01	(.05)	.907
R ²	.52			.52			.53		

Anmerkung. N = 628.

Selbstwirksamkeit der Lesestrategieanwendung. In Tabelle 19 sind die Ergebnisse der schrittweisen Regressionsmodelle zur Erklärung der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention angegeben. Der Prä-Wert erklärt im ersten Modell 32% der Varianz der abhängigen Variable. Durch Hinzunahme der Nutzung der Promptkarten wird in Modell II keine weitere Varianz aufgeklärt. Die Integration der Akzeptanzdimensionen führt in Modell III schließlich zu einer geringfügigen Zunahme an erklärter Varianz auf insgesamt 36%. Hierbei kann die affektive Akzeptanz als signifikant positiver Prädiktor identifiziert werden ($\beta = .26$, $p < .001$). Die Selbstwirksamkeit in der Lesestrategieanwendung ist demnach am Ende der Intervention höher ausgeprägt, je höher die affektive Akzeptanz angegeben wurde. Die kognitive Akzeptanz stellt dagegen keinen signifikanten Prädiktor dar, allerdings ist auch hier die Richtung des Zusammenhangs entgegen der Erwartung negativ.

Tabelle 19: Lineare Regressionsmodelle zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention

	Modell I			Modell II			Modell III		
	β	(SE)	p	β	(SE)	p	β	(SE)	p
Prä-Wert	.57	(.03)	<.001	.57	(.03)	<.001	.54	(.03)	<.001
Nutzung der Promptkarten				.02	(.04)	.653	-.03	(.04)	.431
Affektive Akzeptanz							.26	(.05)	<.001
Kognitive Akzeptanz							-.08	(.05)	.140
R ²	.32			.32			.36		

Anmerkung. $N = 628$.

Einstellungen zu Mehrsprachigkeit. Tabelle 20 stellt die Ergebnisse der schrittweisen Regressionsmodelle zur Untersuchung der Einstellungen zu Mehrsprachigkeit am Ende der Intervention dar. Im ersten Modell erklärt der Prä-Wert 32% der Varianz der abhängigen Variable. Durch Hinzunahme der Nutzung der Promptkarten in Modell II wird keine weitere Varianz aufgeklärt. Die Nutzung der Promptkarten wird nur marginal signifikant, der Zusammenhang ist wie erwartet positiv. Die Hinzunahme der Akzeptanzdimensionen trägt insgesamt nur geringfügig zur Varianzaufklärung auf insgesamt 35% bei. Weder die affektive noch die kognitive Akzeptanzdimension stellen dabei signifikante Prädiktoren dar, erwartungsgemäß zeigen sich positive Koeffizienten.

Tabelle 20: Lineare Regressionsmodelle zur Vorhersage der Einstellungen zu Mehrsprachigkeit am Ende der Intervention

	Modell I			Modell II			Modell III		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Prä-Wert	.57	(.03)	<.001	.56	(.03)	<.001	.52	(.03)	<.001
Nutzung der Promptkarten				.07	(.04)	.051	.03	(.03)	.424
Affektive Akzeptanz							.10	(.06)	.108
Kognitive Akzeptanz							.08	(.06)	.211
R ²	.32			.32			.35		

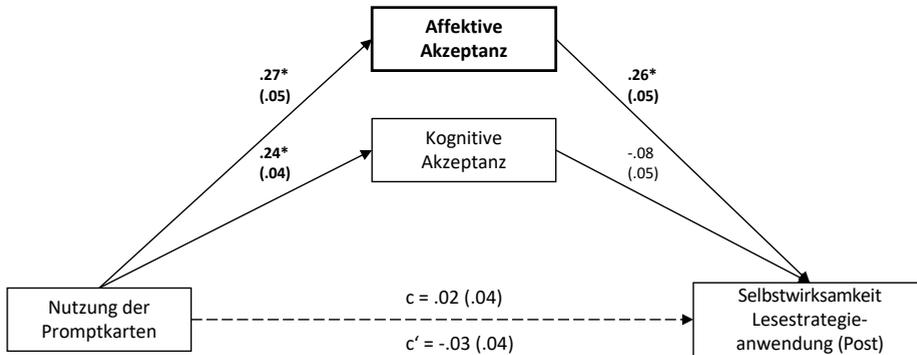
Anmerkung. *N* = 627.

6.2.2 Vermittelnde Rolle der Akzeptanz

Theoretisch wurde abgeleitet, dass die Akzeptanzdimensionen als Mediatoren zwischen der Nutzung der Promptkarten und den Schüler*innenoutcomes fungieren könnten (vgl. Kapitel 4.1.2). Nach dem beschriebenen Vorgehen zum Testen von indirekten Zusammenhängen (vgl. Kapitel 5.4.4) muss kein direkter Zusammenhang zwischen der unabhängigen und abhängigen Variable bestehen, um einen indirekten Zusammenhang sinnvoll zu überprüfen. Es muss jedoch ein signifikanter Zusammenhang zwischen der unabhängigen Variable und der mediiierenden Variable sowie zwischen der mediiierenden Variable und der abhängigen Variable bestehen. Für die vorliegende Arbeit bedeutet dies, dass die Nutzung der Promptkarten die Ausprägung der Akzeptanz signifikant vorhersagen muss und weiterhin die Akzeptanzdimensionen signifikante Prädiktoren für die Schüler*innenoutcomes darstellen müssen. In den autoregressiven Regressionsmodellen (vgl. Tabelle 17 bis Tabelle 20) zeigte sich nur für die Selbstwirksamkeit der Lesestrategieanwendung ein signifikanter Zusammenhang der affektiven Akzeptanz (vgl. Tabelle 19). Aus diesem Grund kann nur für diese Variable ein vermittelnder Zusammenhang der Akzeptanz sinnvoll geprüft werden.

Dazu wird ein Mediationsmodell spezifiziert, in dem die Nutzung der Promptkarten als unabhängige Variable, die Akzeptanzdimensionen als multiple Mediatoren und die Selbstwirksamkeit der Lesestrategieanwendung als abhängige Variable modelliert werden (vgl. Abbildung 5). Die kognitive Akzeptanz wird im Modell zusätzlich spezifiziert, um für diese Dimension der Akzeptanz zu kontrollieren und das Vorgehen analog zu den Regressionsmodellen zu halten. Zudem wurde für die Ausprägung der Selbstwirksamkeit zu Beginn der Intervention im Modell kontrolliert. Die Koeffizienten des gesamten Mediationsmodells sind dem Anhang zu entnehmen (vgl. Tabelle A9).

Abbildung 5: Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung durch die Nutzung der Promptkarten, mediiert über die Akzeptanzdimensionen und unter Kontrolle des Eingangswerts der Selbstwirksamkeit



Anmerkungen. $N = 628$, $*p < .05$; Korrelationen zwischen den Akzeptanzdimensionen: $r = .77$ ($p < .001$).

Im Mediationsmodell in Abbildung 5 zeigt sich, dass zwischen der Nutzung der Promptkarten und der affektiven Akzeptanz ein signifikanter Zusammenhang besteht. Damit ist auch die Voraussetzung eines signifikanten Zusammenhangs zwischen der unabhängigen und mediiierenden Variable erfüllt (vgl. Kapitel 5.4.4). Tabelle 21 zeigt die totalen, direkten und spezifischen indirekten Zusammenhänge im Mediationsmodell. Der vermittelnde Zusammenhang der affektiven Akzeptanz ist hierbei signifikant, da das 95 %-Bootstrap-Konfidenzintervall Null nicht einschließt. Da der totale Zusammenhang nicht signifikant ist, handelt es sich hierbei um eine *indirect-only-mediation* (Zhao et al., 2010; vgl. Kapitel 5.4.4).

Tabelle 21: Indirekte Zusammenhänge im Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung am Ender der Intervention

UV	Mediator	Selbstwirksamkeit Lesestrategieanwendung (Post)			
		β	(SE)	95 %-KI	
				Untere 2.5 %	Obere 2.5 %
Nutzung der Promptkarten	Totaler Zusammenhang (c)	.020	(.039)	-.053	.100
	Totaler indirekter Zusammenhang	.053	(.016)	.022	.085
	Indirekt: Affektive Akzeptanz	.071	(.018)	.035	.101
	Indirekt: Kognitive Akzeptanz	-.018	(.012)	-.042	.006
	Direkter Zusammenhang (c')	-.033	(.042)	-.111	.055

Anmerkungen. $N = 628$, KI = Konfidenzintervall.

6.2.3 Zusammenfassung

Im Rahmen der ersten Fragestellung wurde die Vorhersagekraft der Nutzung der Promptkarten und der Akzeptanz für interventionsspezifische Schüler*innen-outcomes am Ende der Intervention untersucht. Neben den direkten Zusammenhängen wurde zusätzlich die vermittelnde Rolle der Akzeptanz zwischen der Nutzung der Promptkarten und den Schüler*innenoutcomes überprüft.

Unter Kontrolle des jeweiligen Eingangsniveaus zeigte sich entgegen der Annahmen (vgl. Kapitel 4.1.1) keine signifikant positive Vorhersagekraft der untersuchten Facetten des Implementationserfolgs für das Leseverstehen (H1.1a & b), die Lesefreude (H1.2a & b) und die Einstellungen zu Mehrsprachigkeit (H1.4a & b). Lediglich für die Selbstwirksamkeit zur Lesestrategieanwendung konnte die affektive Akzeptanzdimension als positiver Prädiktor bestätigt werden (H1.3b), die Nutzung der Promptkarten hatte entgegen der Erwartungen hier jedoch auch keine Vorhersagekraft (H1.3a). Schüler*innen, die die Intervention auf affektiver Ebene positiv bewerten, zeigen demnach am Ende der Intervention ein höheres Ausmaß an Selbstwirksamkeit in der Lesestrategieanwendung. Die vermittelnde Rolle der Akzeptanz konnte zusätzlich für die Selbstwirksamkeit der Strategieanwendung bestätigt werden (H1.5c). Da der direkte Zusammenhang zwischen der Nutzung der Promptkarten und der Selbstwirksamkeit der Lesestrategieanwendung nicht signifikant ist, ist hierbei von einer *indirect-only-mediation* (Zhao et al., 2010) zu sprechen. Da die Voraussetzung zur Überprüfung von indirekten Zusammenhängen für die anderen interventionsspezifischen Schüler*innenoutcomes nicht gegeben waren, konnten Hypothesen H1.5a, b & d dagegen nicht bestätigt werden.

6.3 Fragestellung 2: Welche individuellen Schüler*innenmerkmale sagen die Facetten des Implementationserfolgs vorher?

Im Rahmen der zweiten Forschungsfrage wird analysiert, welche individuellen Lernvoraussetzungen der Schüler*innen die Nutzung der Promptkarten und die Ausprägung der Akzeptanzdimensionen vorhersagen. Hierbei wird das Merkmal Mehrsprachigkeit als besonders innovatives Element der Intervention betrachtet sowie lerngegenstandsbezogene Voraussetzungen auf kognitiver und affektiv-motivationaler Ebene.

Die direkte Vorhersagekraft der Mehrsprachigkeit wird in einem linearen Regressionsmodell unter Berücksichtigung von Kontrollvariablen (Geschlecht, kognitive Grundfähigkeiten und Anzahl an Büchern im Haushalt) untersucht. Darauf aufbauend werden in einem zweiten Modell zusätzlich lerngegenstandsbezogene Voraussetzungen als Prädiktoren aufgenommen (Leseverstehen, Lesefreude, Selbstwirksamkeit der Lesestrategieanwendung und Einstellungen zu Mehrsprachigkeit), um die postulierten Zusammenhänge zu prüfen. Diese gehen gleichzeitig in die Modelle ein, da aufgrund geringer theoretischer Vorarbeiten zu Facet-

ten des Implementationserfolgs keine Hypothesen über die Stärke der einzelnen Zusammenhänge der Variablen gemacht werden können. Vermutete vermittelnde Zusammenhänge zwischen der Mehrsprachigkeit und den lerngegenstandsbezogenen Variablen als proximale Merkmale werden anschließend in Mediationsmodellen mit multiplen Mediatoren überprüft.

6.3.1 Individuelle Determinanten der Nutzung der Promptkarten

Zur Vorhersage der Nutzung der Promptkarten durch individuelle Schüler*innenmerkmale werden zwei Regressionsmodelle spezifiziert, in denen jeweils die Nutzung der Promptkarten als abhängige Variable eingeht und schrittweise Prädiktoren hinzugefügt werden. Da die Nutzung der Promptkarten als dichotome Variable erhoben wurde, werden die Modelle in Mplus über Probit-Regressionen geschätzt. Zur besseren Interpretation der Regressionskoeffizienten werden die Zusammenhänge in einem *AME* ausgedrückt (vgl. Kapitel 5.4.3). Dieser gibt die durchschnittliche prozentuale Veränderung der Übertrittswahrscheinlichkeit in der abhängigen Variable an, wenn der untersuchte Prädiktor um eine Einheit steigt. Die Ergebnisse der Regressionsmodelle sind Tabelle 22 zu entnehmen.

Tabelle 22: Regressionsmodelle zur schrittweisen Vorhersage der Nutzung der Promptkarten durch lerngegenstandsbezogene Voraussetzungen

Modell I	Nutzung der Promptkarten			
	β	(SE)	<i>p</i>	<i>AME</i>
Mehrsprachigkeit	.06	(.07)	.349	.126
Geschlecht (Jungen)	-.24	(.07)	.001	-.424
Kognitive Grundfähigkeiten	.02	(.06)	.812	.005
Mehr als 25 Bücher im Haushalt	-.17	(.08)	.033	-.306
R ²	.09			
<hr/>				
Modell II				
Mehrsprachigkeit	.01	(.07)	.846	.030
Geschlecht (Jungen)	-.21	(.07)	.005	-.389
Kognitive Grundfähigkeiten	.05	(.07)	.444	.019
Mehr als 25 Bücher im Haushalt	-.18	(.08)	.018	-.354
Leseverstehen	-.13	(.08)	.090	-.026
Lese Freude	.15	(.08)	.051	.189
Selbstwirksamkeit der Lesestrategieanwendung	-.01	(.10)	.908	-.018
Einstellungen zu Mehrsprachigkeit	.13	(.08)	.078	.226
R ²	.14			

Anmerkung. *N* = 634.

Unter Berücksichtigung der Kontrollvariablen sagt das Merkmal Mehrsprachigkeit in Modell I die Nutzung der Promptkarten nicht signifikant vorher. Dagegen stellen die Kontrollvariablen Geschlecht (Jungen) und eine hohe Anzahl an Büchern im Haushalt negative Prädiktoren dar. Die Wahrscheinlichkeit, die Promptkarten während der Intervention genutzt zu haben, ist bei Jungen im Vergleich zu Mädchen durchschnittlich um 42,4 % geringer. Bei Schüler*innen mit weniger als 26 Büchern im Haushalt ist die durchschnittliche Wahrscheinlichkeit zudem um 30,6 % geringer als bei Schüler*innen mit einer höheren Anzahl an Büchern im Haushalt.

In Modell II werden zusätzlich kognitive sowie affektiv-motivationale lerngegenstandsbezogene Voraussetzungen der Schüler*innen als Prädiktoren aufgenommen. Das Merkmal Mehrsprachigkeit ist hierbei weiterhin nicht signifikant, die Stärke des AMEs hat sich jedoch verringert, was ein Hinweis auf mediiierende Zusammenhänge sein könnte. Die kognitive lerngegenstandsbezogene Voraussetzung des Leseverstehens sagt die Nutzung der Promptkarten nur marginal signifikant vorher, der negative Zusammenhang deutet jedoch darauf hin, dass Schüler*innen mit einem höheren Leseverstehen zu Beginn der Intervention seltener angegeben haben, die Promptkarten zu nutzen. Der AME gibt mit 2,6 % jedoch nur eine geringe durchschnittliche Veränderung in der Wahrscheinlichkeit an. Positiv, allerdings auch nur marginal signifikant, wird die Nutzung der Promptkarten durch die affektiv-motivationalen Merkmale Lesefreude und Einstellungen zu Mehrsprachigkeit vorhergesagt. Je Standardabweichung steigt die Wahrscheinlichkeit, die Promptkarten genutzt zu haben, für die Lesefreude um 18,9 % und für die Einstellungen zu Mehrsprachigkeit um 22,6 %. Die Selbstwirksamkeit zur Lesestrategieanwendung steht dagegen in negativem Zusammenhang zur Nutzung der Promptkarten und stellt keinen signifikanten Prädiktor dar.

Darauf aufbauend soll geprüft werden, ob lerngegenstandsbezogene Voraussetzungen den Zusammenhang zwischen dem Merkmal Mehrsprachigkeit und der Nutzung der Promptkarten mediiieren. Zwar zeigte sich in den Probit-Regressionen kein signifikanter direkter Zusammenhang der Mehrsprachigkeit mit der abhängigen Variable, dies stellt in der vorliegenden Arbeit jedoch keine Voraussetzung zur Untersuchung indirekter Zusammenhänge dar (vgl. Zhao et al., 2010; vgl. 5.4.4). Lediglich die Voraussetzungen signifikanter Zusammenhänge zwischen der unabhängigen und mediiierenden Variable sowie zwischen der mediiierenden und der abhängigen Variable müssen gegeben sein, um indirekte Zusammenhänge sinnvoll zu untersuchen. Da keine der lerngegenstandsbezogenen Voraussetzungen signifikant mit der Nutzung der Promptkarten zusammenhängt, können die postulierten vermittelnden Zusammenhänge nicht überprüft werden. Die beschriebenen Voraussetzungen zur Überprüfung von Mediationsmodellen sind damit für die mediiierenden Variablen nicht erfüllt.

6.3.2 Individuelle Determinanten der Akzeptanz

Zur Vorhersage der Akzeptanzdimensionen durch individuelle Schüler*innenmerkmale werden zwei multivariate Regressionsmodelle spezifiziert, in denen beide Akzeptanzdimensionen als abhängige Variablen eingehen. Die Ergebnisse der Regressionsmodelle sind Tabelle 23 zu entnehmen.

Tabelle 23: Multivariate Regressionsmodelle zur schrittweisen Vorhersage der Akzeptanzdimensionen durch lerngegenstandsbezogene Voraussetzungen

	Affektive Akzeptanz			Kognitive Akzeptanz		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Modell I						
Mehrsprachigkeit	.11	(.04)	.016	.15	(.04)	<.001
Geschlecht (Jungen)	-.10	(.04)	.011	-.08	(.04)	.039
Kognitive Grundfähigkeiten	-.03	(.04)	.495	-.12	(.04)	.001
Mehr als 25 Bücher im Haushalt	-.02	(.04)	.567	-.04	(.04)	.299
R ²	.02			.05		
Modell II						
Mehrsprachigkeit	.04	(.05)	.366	.09	(.04)	.038
Geschlecht (Jungen)	-.04	(.04)	.280	-.05	(.04)	.214
Kognitive Grundfähigkeiten	.01	(.04)	.735	-.07	(.04)	.078
Mehr als 25 Bücher im Haushalt	-.08	(.04)	.023	-.06	(.04)	.122
Leseverstehen	-.15	(.06)	.008	-.16	(.05)	.003
Lesefreude	.27	(.05)	<.001	.14	(.05)	.007
Selbstwirksamkeit der Lesestrategieanwendung	.03	(.05)	.589	.02	(.05)	.661
Einstellungen zu Mehrsprachigkeit	.17	(.05)	.001	.19	(.05)	<.001
R ²	.13			.13		

Anmerkungen. *N* = 634, Korrelation zwischen den Akzeptanzdimensionen: $r = .77/.78$ (Modell I/Modell II) ($p < .001$).

Unter Berücksichtigung der Kontrollvariablen sagt das Merkmal Mehrsprachigkeit in Modell I beide Akzeptanzdimensionen signifikant positiv vorher. Mehrsprachige Schüler*innen weisen demnach höhere Ausprägungen in der Freude an der Intervention auf (affektive Akzeptanz) und geben eine höhere Wirksamkeitserwartung der Intervention zur Verbesserung ihrer Lesekompetenz an (kognitive Akzeptanz) als monolinguale Schüler*innen in der Stichprobe. Die untersuchten Prädiktoren erklären in Modell I insgesamt 2% der Varianz in der affektiven Akzeptanz und 5% der Varianz in der kognitiven Akzeptanz. In Modell II wurden zusätzlich kognitive sowie affektiv-motivationale lerngegenstandsbezogene Voraussetzungen

der Schüler*innen als Prädiktoren in das Regressionsmodell aufgenommen. In diesem Modell reduziert sich die Vorhersagekraft des Merkmals Mehrsprachigkeit für beide Akzeptanzdimensionen. Die Mehrsprachigkeit stellt für die affektive Akzeptanz dabei keinen signifikanten Prädiktor mehr dar. Der Zusammenhang der Mehrsprachigkeit zur kognitiven Akzeptanz ist jedoch weiterhin signifikant. Dies deutet auf einen vollständig vermittelnden Zusammenhang durch die kognitiven und affektiv-motivationalen Merkmale für die affektive Akzeptanz sowie auf eine partielle Mediation für die kognitive Akzeptanz hin (vgl. Kapitel 5.4.4).

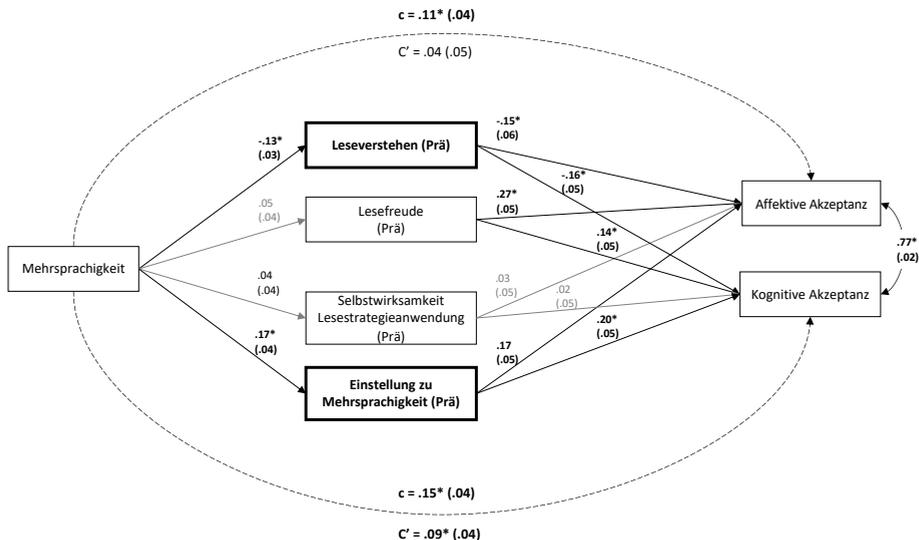
Die kognitive lerngegenstandsbezogene Voraussetzung des Leseverstehens sagt in Modell II sowohl die affektive als auch die kognitive Akzeptanzdimension signifikant negativ vorher. Je höher der Eingangswert im Leseverstehen desto niedriger wird die Akzeptanz der Schüler*innen angegeben. Signifikant positiv werden die Akzeptanzdimensionen dagegen durch die affektiv-motivationalen Merkmale Lesefreude und Einstellungen zu Mehrsprachigkeit vorhergesagt. Je höher die Lesefreude zu Beginn der Intervention angegeben wird, desto höher fällt auch die Ausprägung der Akzeptanzdimensionen aus, gleiches gilt für die Einstellungen zu Mehrsprachigkeit. Lediglich die Selbstwirksamkeit zur Lesestrategieanwendung zu Beginn der Intervention stellt keinen signifikanten Prädiktor für die affektive und kognitive Akzeptanz dar. Die Richtung des Zusammenhangs ist hingegen auch für dieses Merkmal für beide Akzeptanzdimensionen positiv. Insgesamt wird in Modell II durch Einbezug der lerngegenstandsbezogenen Voraussetzungen eine Steigerung der Varianzaufklärung erzielt und jeweils 13 % der Varianz in der affektiven und kognitiven Akzeptanzdimension können aufgeklärt werden.

Darauf aufbauend wird differenzierter untersucht, welche indirekten Zusammenhänge sich zwischen der Mehrsprachigkeit und den Akzeptanzdimensionen nachweisen lassen. Das Leseverstehen, die Lesefreude und die Einstellungen zu Mehrsprachigkeit erfüllen hierbei die Voraussetzung des signifikanten Zusammenhangs zwischen einer mediierenden und einer abhängigen Variable. In einem Mediationsmodell wird daher ein vermittelnder Zusammenhang zwischen der Mehrsprachigkeit und den Akzeptanzdimensionen durch die lerngegenstandsbezogenen Voraussetzungen als multiple Mediatoren spezifiziert. Zunächst zeigt sich, dass die Mehrsprachigkeit signifikante Zusammenhänge zum Leseverstehen und den Einstellungen zu Mehrsprachigkeit aufweist. Für diese Variablen sind demnach alle Voraussetzungen zur Prüfung mediierender Zusammenhänge gegeben. Alle Mediatoren gehen simultan in die Analyse ein, um spezifische indirekte Effekte unter Kontrolle aller Variablen untersuchen zu können (vgl. Kapitel 5.4.4).

Abbildung 6 zeigt die Zusammenhänge im Mediationsmodell für die unabhängige Variable Mehrsprachigkeit. Analog zum Regressionsmodell (vgl. Tabelle 23) werden auch hier die Kontrollvariablen berücksichtigt, zur besseren Lesbarkeit sind die Zusammenhänge der Kontrollvariablen allerdings nicht dargestellt. Für die Mediatoren und abhängigen Variablen wurden Korrelationen zugelassen. Die vollständigen Ergebnisse des Mediationsmodells sind dem Anhang zu entnehmen (vgl. Tabelle A10). Für beide Akzeptanzdimensionen reduziert sich der signifikante

totale Zusammenhang (c) der Mehrsprachigkeit durch den Einbezug der Mediatoren und resultiert für die affektive Akzeptanz in einem nicht signifikanten direkten Zusammenhang (c').

Abbildung 6: Mediationsmodell zur schematischen Darstellung der signifikanten Zusammenhänge zwischen Mehrsprachigkeit und Akzeptanz, unter Berücksichtigung der Kontrollvariablen



Anmerkungen. $N = 634$, $*p < .05$, Korrelation zwischen den Akzeptanzdimensionen: $r = .77$ ($p < .001$).

Die indirekten Zusammenhänge der Mediatoren sind in Tabelle 24 dargestellt. Für beide Akzeptanzdimensionen zeigen sich signifikante spezifische indirekte Zusammenhänge für das Leseverstehen und die Einstellungen zu Mehrsprachigkeit, da hier das 95%-Bootstrap-Konfidenzintervall Null nicht einschließt (vgl. Kapitel 5.4.4). Die indirekten Zusammenhänge sind für beide Mediatoren positiv. Für das Leseverstehen ergibt sich dies aus der Multiplikation zweier negativer Zusammenhänge. Mehrsprachige Schüler*innen weisen zu Beginn der Intervention niedrigere Ausprägungen im Leseverstehen als monolinguale Schüler*innen auf. Eine hohe Ausprägung im Leseverstehen sagt wiederum die Akzeptanzdimensionen negativ vorher, weshalb sich aus beiden negativen direkten Zusammenhängen ein positiver vermittelnder Zusammenhang des Leseverstehens ergibt. Die im Durchschnitt niedrigere Lesekompetenz von mehrsprachigen Schüler*innen hängt auf diese Weise mit einer höheren Akzeptanz der Intervention zusammen. Weiterhin geben mehrsprachige Schüler*innen positivere Einstellungen zu Mehrsprachigkeit zu Beginn der Intervention an als monolinguale Schüler*innen. Diese im Durchschnitt positiveren Einstellungen hängen schließlich mit einer höheren Akzeptanz der Intervention zusammen. Hierbei ergibt sich aus den positiven direkten Zusammenhängen ein positiver indirekter Zusammenhang. Nach Barron und Kenny

(1986) wird der Zusammenhang zwischen der Mehrsprachigkeit und der affektiven Akzeptanz vollständig durch das Leseverstehen und die Einstellungen zu Mehrsprachigkeit mediiert, da in beiden Fällen der direkte Zusammenhang nicht mehr signifikant ist. Die kognitive Akzeptanz wird durch die Mediatoren im Modell dagegen nur partiell mediiert (vgl. Tabelle 24).

Tabelle 24: Spezifische indirekte Zusammenhänge lerngegenstandsbezogener Voraussetzungen im Mediationsmodell zur Vorhersage der Akzeptanzdimensionen durch das Merkmal Mehrsprachigkeit

		Affektive Akzeptanz			
		95 % KI			
UV	Mediator	β	(SE)	Untere 2.5 %	Obere 2.5 %
Mehr- sprachigkeit	Totaler Zusammenhang (c)	.107	(.044)	.019	.195
	Totaler indirekter Zusammenhang	.063	(.018)	.032	.100
	<i>Indirekt: Leseverstehen</i>	.019	(.010)	.004	.043
	<i>Indirekt: Lesefreude</i>	.014	(.010)	-.005	.033
	<i>Indirekt: Selbstwirksamkeit der Lesestrategieanwendung</i>	.001	(.003)	-.004	.008
	<i>Indirekt: Einstellungen zu Mehrsprachigkeit</i>	.029	(.011)	.011	.055
	Direkter Zusammenhang (c')	.044	(.049)	-.051	.138
		Kognitive Akzeptanz			
		95 % KI			
UV	Mediator	β	SE	Untere 2.5 %	Obere 2.5 %
Mehr- sprachigkeit	Totaler Zusammenhang (c)	.147	(.039)	.072	.223
	Totaler Indirekter Zusammenhang	.062	(.014)	.037	.094
	<i>Indirekt: Leseverstehen</i>	.021	(.010)	.006	.044
	<i>Indirekt: Lesefreude</i>	.007	(.006)	-.002	.020
	<i>Indirekt: Selbstwirksamkeit der Lesestrategieanwendung</i>	.001	(.003)	-.006	.006
	<i>Indirekt: Einstellungen zu Mehrsprachigkeit</i>	.033	(.011)	.015	.057
	Direkter Zusammenhang (c')	.085	(.041)	.007	.167

Anmerkungen. N = 634, KI = Konfidenzintervall.

Zusammenfassend ist festzuhalten, dass mehrsprachige Schüler*innen zwar auch unter Berücksichtigung der Kontrollvariablen eine höhere Ausprägung in den Akzeptanzdimensionen aufweisen als monolinguale Schüler*innen in der Stichprobe, dies jedoch nicht ausschließlich auf das Merkmal der Mehrsprachigkeit zurückzuführen ist. Vielmehr zeigt sich, dass eine niedrigere Lesekompetenz im Deutschen sowie positivere Einstellungen zu Mehrsprachigkeit diesen Zusammenhang vermitteln. Diese Ergebnisse tragen zur Erklärung individueller Unterschiede in der Ausprägung der Akzeptanz auf Schüler*innenebene bei, indem insbesondere lerngegenstandsbezogene Voraussetzungen auf kognitiver sowie affektiv-motivationaler Ebene die Konstrukte vorhersagen. Für die kognitive Akzeptanz sind schließlich weitere mediiierende Mechanismen denkbar. Auch unter Kontrolle der Mediatoren geben mehrsprachige Schüler*innen ein höheres Ausmaß an kognitiver Akzeptanz an.

6.3.3 Zusammenfassung

Die zweite Fragestellung untersuchte individuelle Determinanten der Facetten des Implementationserfolgs auf Schüler*innenebene, um zu analysieren, welche individuellen Merkmale zu einem hohen Implementationserfolg der Intervention beitragen. Für den spezifischen Kontext einer mehrsprachigkeitssensiblen Intervention wurde explorativ untersucht, ob die Mehrsprachigkeit der Schüler*innen die Facetten des Implementationserfolgs vorhersagt. Zudem wurden lerngegenstandsbezogene Voraussetzungen auf kognitiver und affektiv-motivationaler Ebene, die zu Beginn der Intervention erhoben wurden, als Prädiktoren untersucht.

Für die Nutzung der Promptkarten zeigte die explorative Untersuchung des Merkmals Mehrsprachigkeit keinen direkten Zusammenhang (F2.1). Weiterhin konnten die Hypothesen H2.1 bis H2.3 nicht bestätigt werden, da das Leseverstehen (H2.1), die Lesefreude (H2.2) und die Selbstwirksamkeit der Lesestrategieanwendung (H2.3) die Nutzung der Promptkarten nicht signifikant vorhersagten. Auch die Einstellungen zu Mehrsprachigkeit sagten die Nutzung der Promptkarten nicht signifikant vorher (F2.2). Vermittelnde Zusammenhänge, die in Forschungsfrage F2.3 untersucht werden sollten, konnten aufgrund fehlender signifikanter Zusammenhänge der angenommenen mediiierenden Variablen nicht überprüft werden.

Die Mehrsprachigkeit sagt unter alleiniger Berücksichtigung der Kontrollvariablen die affektive und kognitive Akzeptanzdimension positiv vorher (F2.4). Bei Hinzunahme der lerngegenstandsbezogenen Voraussetzungen war dieser Zusammenhang im zweiten Modell jedoch nur noch für die kognitive Akzeptanz signifikant. Das Leseverstehen zu Beginn der Intervention sagte beide Akzeptanzdimensionen entgegen den Annahmen negativ vorher, weswegen die Hypothese H2.4 nicht bestätigt werden konnte. Die Lesefreude sagte beide Akzeptanzdimensionen positiv vorher, daher konnte H2.5 bestätigt werden. Hypothese H2.6 konnte nicht

bestätigt werden, da die Selbstwirksamkeit der Lesestrategieanwendung keine signifikante Vorhersagekraft zeigte. Die Einstellungen zu Mehrsprachigkeit, die explorativ untersucht wurden, sagten beide Akzeptanzdimensionen positiv vorher (F2.5). Zusätzlich konnten das Leseverstehen zu Beginn der Intervention und die Einstellungen zu Mehrsprachigkeit als Mediatoren zwischen der Mehrsprachigkeit und der affektiven und kognitiven Akzeptanz bestätigt werden (F2.6).

6.4 Fragestellung 3: Sagt die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen interventionsspezifische Schüler*innenoutcomes vorher?

Ein besonders innovatives Element der Intervention stellt die Integration der Familiensprachen der Schüler*innen dar, die u. a. über bilinguale Promptkarten aktiviert werden sollten. Im Folgenden wird spezifisch für die Gruppe mehrsprachiger Schüler*innen untersucht, welche Bedeutung die Nutzung der Promptkarten in einer anderen Sprache für den Lernerfolg mehrsprachiger Schüler*innen einnimmt. Analog zu den Untersuchungen der ersten und zweiten Fragestellung werden hierbei das Leseverstehen, die Lesefreude, die Selbstwirksamkeit zur Lesestrategieanwendung und die Einstellungen zu Mehrsprachigkeit fokussiert und die Vorhersagekraft der Nutzung der Promptkarten in einer anderen Sprache und Akzeptanzdimensionen als Facetten des Implementationserfolgs untersucht.

6.4.1 Direkte Vorhersagekraft der Nutzung der Promptkarten in einer anderen Sprache für interventionsspezifische Schüler*innenoutcomes

Analog zum Vorgehen in Fragestellung 1 (vgl. Kapitel 6.2) werden schrittweise autoregressive Regressionsmodelle spezifiziert. In diesen werden die Ausprägungen der Outcomes am Ende der Intervention unter Kontrolle des Prä-Werts durch die Nutzung der Promptkarten in einer anderen Sprache und die Akzeptanzdimensionen vorhergesagt. In die Analysen gehen nur mehrsprachige Schüler*innen ein, die die Promptkarten in (mindestens) einer ihrer Familiensprachen vorliegen hatten (vgl. Kapitel 5.1.3).

Leseverstehen. Tabelle 25 gibt die Ergebnisse der schrittweisen Regressionsmodelle zur Vorhersage des Leseverstehens an. Der Wert im Leseverstehen am Ende der Intervention wird in Modell I in hohem Maße durch das Eingangsniveau vorhergesagt. Hierbei werden 63% der Varianz in der abhängigen Variable erklärt. Durch Hinzunahme der Nutzung der Promptkarten in einer anderen Sprache kann in Modell II kein weiterer Anteil an Varianz aufgeklärt werden und die Variable stellt keinen signifikanten Prädiktor dar. Gleiches gilt für die Integration der Akzeptanzdimensionen in Modell III. Auch hier stellen die Variablen keine signi-

fikanten Prädiktoren dar und kein weiterer Anteil an Varianz wird aufgeklärt. Das Regressionsgewicht der kognitiven Akzeptanz zeigt entgegen der Erwartung einen negativen Zusammenhang an, der jedoch nicht signifikant ist.

Tabelle 25: Lineare Regressionsmodelle zur Vorhersage des Leseverstehens am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen

	Modell I ¹			Modell II ²			Modell III ²		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Prä-Wert	.79	(.02)	<.001	.79	(.02)	<.001	.79	(.03)	<.001
Nutzung der Promptkarten in einer anderen Sprache				.03	(.03)	.360	.03	(.03)	.253
Affektive Akzeptanz							.04	(.05)	.487
Kognitive Akzeptanz							-.08	(.05)	.124
R ²	.63			.63			.63		

Anmerkungen. ¹N = 423, ²N = 424.

Lesefreude. In Tabelle 26 sind die Ergebnisse der schrittweisen Regressionsmodelle zur Vorhersage der Lesefreude am Ende der Intervention angegeben. Der Prä-Wert der Lesefreude sagt die abhängige Variable in Modell I in hohem Maße vorher. Dabei werden 48 % der Varianz aufgeklärt. In Modell II zeigt sich die Nutzung der Promptkarten in einer anderen Sprache als signifikant positiver Prädiktor ($\beta = .12$, $p = .004$), es wird jedoch nur 1% an Varianz zusätzlich aufgeklärt. In Modell III gehen zusätzlich die Akzeptanzdimensionen ein, die jedoch keine signifikanten Prädiktoren darstellen. Die Nutzung der Promptkarten in einer anderen Sprache bleibt hierbei als signifikanter Prädiktor bestehen.

Tabelle 26: Lineare Regressionsmodelle zur Vorhersage der Lesefreude am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen

	Modell I			Modell II			Modell III		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Prä-Wert	.69	(.04)	<.001	.68	(.04)	<.001	.66	(.04)	<.001
Nutzung der Promptkarten in einer anderen Sprache				.12	(.04)	.004	.11	(.04)	.009
Affektive Akzeptanz							.06	(.06)	.257
Kognitive Akzeptanz							.03	(.06)	.619
R ²	.48			.49			.50		

Anmerkung. N = 426.

Selbstwirksamkeit der Lesestrategieanwendung. Tabelle 27 stellt die Ergebnisse der schrittweisen Regressionsmodelle zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention dar. Der Prä-Wert erklärt im ersten Modell 30 % der Varianz in der Ausprägung am Ende der Intervention. Durch Hinzunahme der Nutzung der Promptkarten in einer anderen Sprache wird in Modell II keine weitere Varianz aufgeklärt und der Prädiktor ist nicht signifikant. Die Integration der Akzeptanzdimensionen in Modell III führt schließlich zu einer geringfügigen Zunahme an erklärter Varianz auf insgesamt 35 %. Die affektive Akzeptanz stellt hierbei einen signifikanten Prädiktor mit positivem Zusammenhang dar ($\beta = .23, p < .001$). Die Selbstwirksamkeit der Lesestrategieanwendung ist demnach am Ende der Intervention höher, je positiver die affektive Akzeptanz bewertet wurde. Die kognitive Akzeptanz stellt dagegen keinen signifikanten Prädiktor dar.

Tabelle 27: Lineare Regressionsmodelle zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen

	Modell I			Modell II			Modell III		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Prä-Wert	.55	(.04)	<.001	.54	(.04)	<.001	.52	(.04)	<.001
Nutzung der Promptkarten in einer anderen Sprache				.01	(.04)	.732	-.02	(.04)	.616
Affektive Akzeptanz							.23	(.07)	<.001
Kognitive Akzeptanz							.01	(.06)	.903
R ²	.30			.30			.35		

Anmerkung. *N* = 425.

Einstellungen zu Mehrsprachigkeit. Tabelle 28 stellt die Ergebnisse der schrittweisen Regressionsmodelle zur Vorhersage der Einstellungen zu Mehrsprachigkeit dar. Im ersten Modell erklärt die Ausprägung zu Beginn der Intervention 28 % der Varianz in der abhängigen Variable. Durch Hinzunahme der Nutzung der Promptkarten in einer anderen Sprache in Modell II wird keine weitere Varianz aufgeklärt und der zusätzliche Prädiktor ist nicht signifikant. Die Hinzunahme der Akzeptanzdimensionen in Modell III führt nur zu einer Steigerung erklärter Varianz auf insgesamt 30 %. Beide Dimensionen sind nicht signifikant.

Tabelle 28: Lineare Regressionsmodelle zur Vorhersage der Einstellungen zu Mehrsprachigkeit am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen

	Modell I			Modell II			Modell III		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Prä-Wert	.53	(.05)	<.001	.51	(.05)	<.001	.50	(.05)	<.001
Nutzung der Promptkarten in einer anderen Sprache				.05	(.05)	.329	.03	(.05)	.519
Affektive Akzeptanz							.08	(.07)	.268
Kognitive Akzeptanz							.07	(.07)	.355
R ²	.28			.28			.30		

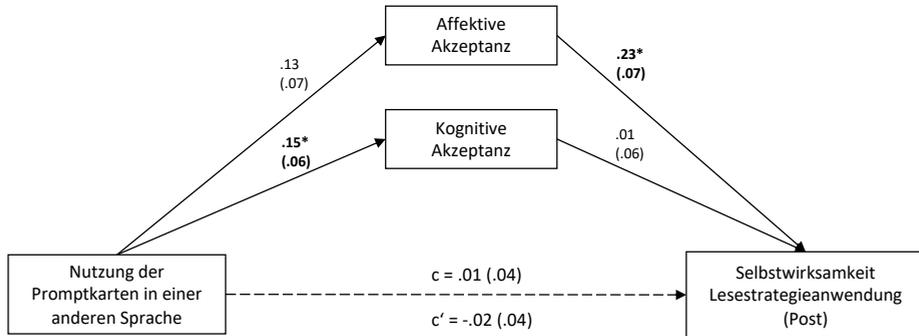
Anmerkung. *N* = 424.

6.4.2 Vermittelnde Rolle der Akzeptanz

In Kapitel 6.2.2 wurde die angenommene medierende Rolle der Akzeptanzdimensionen für den Zusammenhang zwischen der Nutzung der Promptkarten und den Schüler*innenoutcomes in der Gesamtstichprobe untersucht. Nur die Selbstwirksamkeit der Lesestrategieanwendung erfüllte die im Methodenkapitel beschriebenen Voraussetzungen zur Untersuchung indirekter Zusammenhänge. Für die Nutzung der Promptkarten in einer anderen Sprache soll ebenfalls untersucht werden, ob die Akzeptanzdimensionen den Zusammenhang zwischen der Nutzung und den Schüler*innenoutcomes medieren. Hierfür werden ebenfalls zunächst die Voraussetzungen überprüft. In den autoregressiven Regressionsmodellen (vgl. Tabelle 25 bis Tabelle 28) zeigte sich nur für die Selbstwirksamkeit der Lesestrategieanwendung ein signifikanter Zusammenhang der affektiven Akzeptanz (vgl. Tabelle 27). Aus diesem Grund kann nur für diese Variable ein vermittelnder Zusammenhang sinnvoll geprüft werden.

Hierzu wird erneut ein Mediationsmodell spezifiziert, in dem die Nutzung der Promptkarten in einer anderen Sprache als unabhängige Variable, die Akzeptanzdimensionen als multiple Mediatoren und die Selbstwirksamkeit der Lesestrategieanwendung als abhängige Variable spezifiziert werden (vgl. Abbildung 7). Die kognitive Akzeptanz wird im Modell zusätzlich spezifiziert, um für diese Dimension der Akzeptanz zu kontrollieren und das Vorgehen analog zu den Regressionsmodellen zu halten. Zudem wird für die Ausprägung der Selbstwirksamkeit zu Beginn der Intervention im Modell kontrolliert. Die Koeffizienten des gesamten Mediationsmodells sind Tabelle A11 im Anhang zu entnehmen.

Abbildung 7: Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung durch die Nutzung der Promptkarten in einer anderen Sprache, mediiert über die Akzeptanzdimensionen und unter Kontrolle des Eingangswerts der Selbstwirksamkeit



Anmerkungen. $N = 425$, $*p < .05$, Korrelationen zwischen den Akzeptanzdimensionen: $r = .77$ ($p < .001$).

Im Mediationsmodell in Abbildung 7 zeigt sich, dass zwischen der Nutzung der Promptkarten in einer anderen Sprache und der kognitiven Akzeptanz ein signifikanter positiver Zusammenhang besteht. Mehrsprachige Schüler*innen, die die Promptkarten in einer anderen Sprache genutzt haben, bewerten die Intervention demnach als nützlicher für das Lesen als mehrsprachige Schüler*innen, welche die Promptkarten während der Intervention nicht in einer anderen Sprache genutzt haben. Die Voraussetzung zur Überprüfung indirekter Zusammenhänge ist damit nur für die kognitive Akzeptanz gegeben, da nur hier ein Zusammenhang zwischen der unabhängigen und mediiierenden Variablen besteht. Die kognitive Akzeptanz sagt jedoch die abhängige Variable – die Selbstwirksamkeit der Lesestrategieanwendung – nicht signifikant vorher. Damit sind die Voraussetzungen zur Überprüfung indirekter Zusammenhänge nicht gegeben. Die Akzeptanzdimensionen stellen demnach keine Mediatoren zwischen der Nutzung der Promptkarten in einer anderen Sprache und der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention dar.

6.4.3 Zusammenfassung

Im Rahmen der dritten Fragestellung wurde die Vorhersagekraft der Nutzung der Promptkarten in einer anderen Sprache für den Lernerfolg mehrsprachiger Schüler*innen untersucht. Neben den direkten Zusammenhängen wurde zusätzlich die vermittelnde Rolle der Akzeptanz überprüft.

Unter Kontrolle des jeweiligen Eingangsniveaus zeigte sich entgegen den Annahmen (vgl. Kapitel 4.3) keine signifikante positive Vorhersagekraft der untersuchten Facetten des Implementationserfolgs für das Leseverstehen (H3.1a & b)

und die Einstellungen zu Mehrsprachigkeit (H3.4a & b). Dagegen sagte die Nutzung der Promptkarten in einer anderen Sprache die Lesefreude am Ende der Intervention signifikant positiv vorher, weswegen Hypothese H3.1a bestätigt werden konnte. Die Akzeptanz hatte dabei keine Vorhersagekraft und Hypothese H3.2b wurde abgelehnt. Von den Akzeptanzdimensionen zeigte lediglich die affektive Akzeptanz positive Vorhersagekraft für die Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention und Hypothese H3.3b wurde bestätigt. Die vermittelnde Rolle der Akzeptanz konnte für keine der untersuchten Outcomes bestätigt werden, weswegen Hypothese H3.5 abgelehnt wurde.

6.5 Fragestellung 4: Welche individuellen Schüler*innenmerkmale sagen die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen vorher?

Schließlich werden mit der vierten Fragestellung individuelle Determinanten mehrsprachiger Schüler*innen für die Nutzung der Promptkarten in einer anderen Sprache untersucht, um einen näheren Einblick zu erhalten, welche Schüler*innen das innovative Angebot zur Integration ihrer Familiensprachen im Lernprozess annehmen. Analog zum Vorgehen in Fragestellung 2 (vgl. Kapitel 6.3) wird die direkte Vorhersagekraft von lerngegenstandsbezogenen Voraussetzungen (Leseverstehen, Lesefreude, Selbstwirksamkeit der Lesestrategieanwendung und Einstellungen zu Mehrsprachigkeit) unter Berücksichtigung der Kontrollvariablen (Geschlecht, kognitive Grundfähigkeiten und Anzahl an Büchern im Haushalt) untersucht. Hierbei werden ebenfalls nur Schüler*innen einbezogen, die die Promptkarten in (mindestens) einer ihrer Familiensprachen vorliegen hatten (vgl. Kapitel 5.1.3).

6.5.1 Individuelle Determinanten für die Nutzung der Promptkarten in einer anderen Sprache bei mehrsprachigen Schüler*innen

Zur Vorhersage der Nutzung der Promptkarten in einer anderen Sprache durch individuelle Lernvoraussetzungen werden zwei Probit-Regressionsmodelle berechnet, in denen jeweils die Nutzung der Promptkarten in einer anderen Sprache als abhängige Variable spezifiziert ist und schrittweise Prädiktoren hinzugefügt werden. Die Ergebnisse der Regressionsmodelle sind Tabelle 29 zu entnehmen.

Im ersten Modell zeigt keine der untersuchten Kontrollvariablen einen signifikanten Zusammenhang zur Nutzung der Promptkarten in einer anderen Sprache. Im zweiten Modell sagen die Einstellungen zu Mehrsprachigkeit die Nutzung der Promptkarten in einer anderen Sprache signifikant positiv vorher. Der *AME* zeigt an, dass sich die Wahrscheinlichkeit, die Promptkarten während der Intervention

in einer anderen Sprache genutzt zu haben, je Standardabweichung des Prädiktors um 18 % erhöht. Alle anderen untersuchten lerngegenstandsbezogenen Voraussetzungen stellen keine signifikanten Prädiktoren dar.

Tabelle 29: Regressionsmodelle zur schrittweisen Vorhersage der Nutzung der Promptkarten in einer anderen Sprache durch lerngegenstandsbezogene Voraussetzungen in der Substichprobe mehrsprachiger Schüler*innen

Modell I	Nutzung Promptkarten in einer anderen Sprache			
	β	(SE)	<i>p</i>	AME
Geschlecht (Jungen)	-.03	(.07)	.643	-.043
Kognitive Grundfähigkeiten	.02	(.06)	.803	.004
Mehr als 25 Bücher im Haushalt	.07	(.08)	.353	.102
R ²	.01			
Modell II				
Geschlecht (Jungen)	-.01	(.07)	.910	-.011
Kognitive Grundfähigkeiten	.02	(.06)	.709	.006
Mehr als 25 Bücher im Haushalt	.01	(.08)	.877	.016
Leseverstehen	.04	(.08)	.634	.006
Lese Freude	.08	(.09)	.360	.068
Selbstwirksamkeit der Lesestrategieanwendung	.06	(.07)	.371	.064
Einstellungen zu Mehrsprachigkeit	.15	(.07)	.036	.180
R ²	.05			

Anmerkung. *N* = 431.

6.5.2 Zusammenfassung

Wie erwartet, konnten die Einstellungen zu Mehrsprachigkeit zu Beginn der Intervention bei mehrsprachigen Schüler*innen als positiver Prädiktor für die Nutzung der Promptkarten in einer anderen Sprache festgestellt werden. Hypothese H4.4 konnte demnach bestätigt werden. Mehrsprachige Schüler*innen mit positiveren Einstellungen zu Mehrsprachigkeit gaben demnach häufiger an, die Promptkarten in einer anderen Sprache genutzt zu haben als mehrsprachige Schüler*innen mit negativeren Einstellungen zu Mehrsprachigkeit. Die anderen lerngegenstandsbezogenen Voraussetzungen zeigten dagegen für die Nutzung der Promptkarten in einer anderen Sprache keine Vorhersagekraft, weswegen die Hypothesen H4.1 (Leseverstehen), H4.2 (Lese Freude) und H4.3 (Selbstwirksamkeit der Lesestrategieanwendung) nicht bestätigt werden konnten.

7 Diskussion

Ziel der vorliegenden Arbeit war es, den Implementationserfolg beim mehrsprachigkeitssensiblen Reziproken Lehren aus Schüler*innenperspektive zu untersuchen. Als Facetten von Implementationserfolg wurden die Wiedergabetreue und die Akzeptanz der Schüler*innen analysiert. Beim mehrsprachigkeitssensiblen Reziproken Lehren üben Schüler*innen in kleinen Gruppen Lese- und Lernstrategien ein und dürfen dabei alle Sprachen verwenden, die sie sprechen. Zur Bearbeitung der Lesestrategien werden in der Unterrichtseinheit sogenannte Promptkarten eingesetzt, auf denen die wichtigsten Schritte zum Ausführen der Strategien beschrieben sind und lernförderliche kognitive Aktivitäten in der kooperativen Arbeit begünstigen sollen (vgl. Kapitel 5.1.1). Die Promptkarten sind bilingual gestaltet, sodass neben der deutschen Beschreibung der Strategien jeweils auch die Übersetzung in einer Familiensprache der Schüler*innen angegeben ist. Die Nutzung der Promptkarten verknüpft als innovatives Element der Intervention das Einüben der Lesestrategien mit der Integration von Mehrsprachigkeit. Es wird angenommen, dass die Nutzung der Promptkarten ein interventionsspezifisches Kernelement für die Wirksamkeit der Intervention darstellt. Sie ist als eine Facette der Wiedergabetreue auf Seiten der Empfänger*innen der Intervention angesiedelt, die als *participant adherence* (Jones et al., 2008; zit. nach Sanetti & Kratochwill, 2009; vgl. Kapitel 3.3) beschrieben werden kann. Die Akzeptanz der Schüler*innen wurde als Gesamtevaluation der Intervention konzeptualisiert und als einstellungsähnliches Konstrukt in Anlehnung an die *Treatment-Acceptability*-Forschung auf affektiver und kognitiver Ebene untersucht (Eckert et al., 2017; Eckert et al., 2021). Die aufgestellten Forschungsfragen untersuchen zum einen die Bedeutung der Facetten des Implementationserfolgs für den Lernerfolg der Schüler*innen und zum anderen, wie individuelle Unterschiede im Implementationserfolg durch individuelle Lernvoraussetzungen erklärt werden können. Ein besonderer Fokus wird hierbei auf die Mehrsprachigkeit als innovatives Element der Untersuchung gelegt.

Im Folgenden werden die Ergebnisse der Forschungsfragen vor dem Hintergrund des in den Kapiteln 2 und 3 vorgestellten Forschungsstandes zum Umgang mit Mehrsprachigkeit und zur Implementation von Interventionen im Unterricht interpretiert und diskutiert. Der erste Teil der Diskussion nähert sich aus einem breiteren Blick der Implementationsforschung an die untersuchten Facetten des Implementationserfolgs an, indem die Bedeutung dieser für interventionsspezifische Schüler*innenoutcomes sowie individuelle Determinanten auf Schüler*innenebene untersucht werden (vgl. Kapitel 7.1). Im zweiten Teil der Diskussion wird die Rolle der Mehrsprachigkeit fokussiert. Hierbei wird zum einen die Bedeutung der Mehrsprachigkeit als Prädiktor für den Implementationserfolg analysiert sowie aus einer differenziellen Perspektive für mehrsprachige Schüler*innen die Aktivierung von Mehrsprachigkeit über die Nutzung der bilingualen Promptkarten in einer anderen Sprache untersucht (vgl. Kapitel 7.2). Anschließend werden praktische Implikationen der Ergebnisse abgeleitet (vgl. Kapitel

7.3) sowie die Limitationen der Untersuchung diskutiert, an die sich ausblickende Forschungsfragen anschließen (vgl. Kapitel 7.4). In einem abschließenden Fazit werden die zentralen Ergebnisse der Arbeit übergeordnet zusammengefasst (vgl. Kapitel 7.5).

7.1 Evaluation des Implementationserfolgs auf Schüler*innenebene

Die Nutzung der Promptkarten und Akzeptanz der Schüler*innen stellen Facetten einer evaluativen Begleitung des innovativen didaktischen Konzepts zum mehrsprachigkeitssensiblen Reziproken Lehren dar. Als Kernelement der Intervention wurde die Nutzung der Promptkarten ausgewählt, mit denen die Lesestrategien systematisch eingeübt und lernförderliche kognitive Aktivitäten in der Gruppenarbeit begünstigt werden sollen (vgl. Kapitel 4). Gemessen an diesen Facetten zeigen die deskriptiven Ergebnisse der Untersuchung insgesamt ein hohes Maß an Implementationserfolg auf Ebene der Schüler*innen. Die große Mehrheit der Schüler*innen gab an, die Promptkarten während der Intervention genutzt zu haben und die affektive und kognitive Akzeptanz als Gesamtevaluation der Intervention ist im Durchschnitt hoch ausgeprägt (vgl. Kapitel 6.1). Dies deutet auf eine hohe Praxiskompatibilität und soziale Validität der Intervention auf Ebene der Schüler*innen hin (Kazdin, 1980; Wolf, 1978).

Ausgehend von Modellen der Implementationsforschung (Witt & Elliott, 1985) und Unterrichtsforschung (Vieluf et al., 2020) wurde erwartet, dass die Nutzung der Promptkarten und Akzeptanzdimensionen interventionsspezifische Schüler*innenoutcomes vorhersagen. Hierbei wurden das Leseverstehen, die Lesefreude, Selbstwirksamkeitserwartungen der Lesestrategieanwendung und Einstellungen zu Mehrsprachigkeit untersucht. Obwohl erwartet wurde, dass alle untersuchten interventionsspezifischen Schüler*innenoutcomes durch die Nutzung der Promptkarten und Akzeptanzdimensionen erklärt werden, konnte dies lediglich für die Selbstwirksamkeit der Lesestrategieanwendung bestätigt werden (vgl. Kapitel 6.2). Die Ergebnisse für die Selbstwirksamkeit schließen damit an bestehende Forschungsbefunde an, in denen das Ausmaß an Implementationserfolg von Unterrichtsinterventionen mit einem höheren Lernerfolg assoziiert ist (Eckert et al., 2021; Low et al., 2014; Mautone et al., 2009).

Die Nutzung der Promptkarten sagte die Ausprägung der Selbstwirksamkeit der Lesestrategieanwendung jedoch nur vermittelt über die affektive Akzeptanz vorher, ein direkter Zusammenhang war in den Analysen nicht signifikant. Im Rahmen des Mediationsmodells (vgl. Abbildung 5) bestätigte sich hierbei zunächst die Annahme, dass die Nutzung der Promptkarten mit einer positiven Gesamtevaluation der Intervention auf affektiver und kognitiver Ebene zusammenhängt. Schüler*innen, die angaben, die Promptkarten während der Intervention genutzt zu haben, zeigen demnach höhere Ausprägungen der affektiven und kognitiven Akzeptanz. Dies ist anschlussfähig an konzeptuelle Modelle und Befunde der

Implementationsforschung, nach denen Facetten der Wiedergabetreue und Akzeptanzurteile von Lehrkräften positiv zusammenhängen (Hebbecke et al., 2020; Henninger, 2010; Mautone et al., 2009). Der Forschungsstand wird mit den vorliegenden Ergebnissen dahingehend erweitert, dass sich diese Zusammenhänge für die untersuchten Konstrukte auch auf Ebene der Schüler*innen zeigen.

Die Nutzung der Promptkarten kann zudem als interventionsspezifisches Kernelement und Wirkfaktor der Intervention bestätigt werden, da sie wünschenswerte Veränderungen in der Praxis begünstigt (Blase & Fixsen, 2013; Murrain et al., 2016). Die wünschenswerten Veränderungen sind hier jedoch nur auf die Selbstwirksamkeitserwartungen, vermittelt über die affektive Akzeptanz, beschränkt. Selbstwirksamkeitserwartungen werden durch positive Erfahrungen bei der Bearbeitung von Aufgaben sowie positive Leistungsrückmeldungen geformt (Bandura, 1977). Die Nutzung der Promptkarten scheint demnach mit positiven Erfahrungen bei der Bearbeitung der Lesestrategien zusammenhängen, die in der Untersuchung als affektive Akzeptanz erhoben wurden. Aus der Logik der Angebot-Nutzungsmodelle der Unterrichtsforschung kann die Nutzung der Promptkarten als Gestaltung des Lernangebots demnach innere Nutzungsprozesse anregen, die in stetigem reziproken Austausch mit individuellen Wahrnehmungsurteilen den Lernerfolg beeinflussen (Vieluf et al., 2020; vgl. Kapitel 3.5.1). Selbstwirksamkeitserwartungen können weiterhin auch zukünftiges Verhalten beeinflussen (Bandura, 1977), weswegen angenommen werden kann, dass sich die Nutzung der Promptkarten, die mit höheren Selbstwirksamkeitserwartungen am Ende der Intervention assoziiert ist, positiv auf das zukünftige Leseverhalten auswirkt. Dadurch können auch das Leseverstehen und die Lesemotivation gefördert werden (Goy et al., 2017; McElvany et al., 2008; Unrau et al., 2018).

Dass sich der positive Zusammenhang der Facetten des Implementationserfolgs nur für die Selbstwirksamkeit der Lesestrategieanwendung zeigte, könnte methodisch darauf zurückzuführen sein, dass die Skala zur Selbstwirksamkeit im Vergleich zu den anderen untersuchten Outcomes besonders instruktionssensitiv ist und Zusammenhänge zum Implementationserfolg daher besser erfasst werden können (Naumann et al., 2019). Die Skala bildet die spezifischen Aktivitäten der Intervention im Vergleich zu den anderen untersuchten Schüler*innenoutcomes besonders nah ab und ist auf praktischer Ebene mit der Nutzung der Promptkarten eng verwoben. Die Lesefreude und das Leseverstehen stellen dagegen eher übergeordnete Merkmale dar, die im kurzen Interventionszeitraum von durchschnittlich 5,4 Wochen (vgl. Kapitel 5.1.3) durch die Intervention weniger verändert werden können und weiter vom konkreten Unterrichtsgegenstand – dem Einüben der Lesestrategien – entfernt sind. Diese Annahme wird auch dadurch unterstützt, dass die Lesefreude zwischen Prä- und Post-Erhebung stabil bleibt und sich nicht signifikant verbessert (vgl. Kapitel 6.1.3). In einer Metaanalyse zum Reziproken Lehren von Rosenshine und Meister (1994) zeigte sich bereits, dass die Effekte höher ausfallen, wenn interventionsnahe Instrumente verwendet werden

anstelle standardisierter Tests. Dies könnte den Ergebnissen zufolge auch für Zusammenhänge zum Implementationserfolg gegeben sein.

Es muss jedoch berücksichtigt werden, dass sich der Zusammenhang zwischen der Nutzung der Promptkarten und der Selbstwirksamkeitserwartung der Lesestrategieanwendung nur im Rahmen einer *indirect-only-mediation* (Zhao et al., 2010; vgl. Kapitel 5.4.4), vermittelt über die affektive Akzeptanz, zeigte. Dies kann ein Indikator für weitere und komplexe Zusammenhänge zwischen den Konstrukten sein. Da die Nutzung der Promptkarten als *participant adherence* nur eine Facette der Wiedergabetreue widerspiegelt, könnten weitere Facetten moderierend wirken, die in dieser Studie nicht erhoben wurden. Wiedergabetreue wird als mehrdimensionales Konstrukt konzeptualisiert, welches qualitative sowie quantitative Dimensionen umfasst (vgl. Kapitel 3.3.1). In einer Übersichtsarbeit zur Wiedergabetreue bei schulbasierten Programmen zur Verbesserung der psychischen Gesundheit zeigt die Adhärenz im Vergleich zu qualitativen Dimensionen nur geringe Zusammenhänge zu den gewünschten Outcomes bei Schüler*innen (Rojas-Andrade & Bahamondes, 2019). Auch Humphrey et al. (2018) führen an, dass qualitative Dimensionen der Wiedergabetreue, wie die Responsivität der Teilnehmenden, die wünschenswerten Veränderungen in einer Intervention zur Förderung sozialer und emotionaler Fähigkeiten bei Schüler*innen signifikant vorhersagen. Aus einer Studie von Dowling und Barry (2020) zur Förderung sozial-emotionaler Entwicklung geht zudem hervor, dass die Qualität der Umsetzung, gefolgt von der Responsivität der Teilnehmenden, die einflussreichsten Dimensionen der Wiedergabetreue für die untersuchten Outcomes darstellen. Auch im Kontext kooperativer Lernarrangements erscheint die Qualität der Nutzung der Promptkarten und der resultierenden Interaktionsqualität zwischen den Schüler*innen von besonderer Bedeutung (vgl. Adl-Amini, 2018; Lipowsky et al., 2022). Die Qualität der Umsetzung könnte dabei an lernförderlichen kognitiven Aktivitäten, wie z. B. gegenseitigem Erklären, Diskutieren und Argumentieren gemessen werden. Lipowsky et al. (2022) fassen zusammen, dass die Interaktionsqualität beim kooperativen Lernen jedoch häufig unzureichend ist, wenn Schüler*innen nicht präzise angeleitet und beim kooperativen Arbeiten unterstützt werden. In der vorliegenden Untersuchung sind keine Aussagen zur Qualität oder Häufigkeit der Nutzung der Promptkarten möglich, die von Relevanz sein könnten. Um die Zusammenhänge genauer untersuchen zu können, wäre es wichtig, mehr Facetten der Wiedergabetreue auf Ebene der Schüler*innen zu erheben und zudem die Qualität der Nutzung der Promptkarten zu untersuchen. Die Dosierung könnte im Sinne einer Häufigkeit der Nutzung ebenso einen relevanten Faktor darstellen, wie sich aus bisherigen Untersuchungen der Implementationsforschung ableiten lässt (Durlak & DuPre, 2008) und auch im Rahmen der Lehr-Lernforschung für eine effektive und hohe Lernzeit zu erwarten ist (z. B. Ditton, 2000).

Das Ausbleiben des direkten Zusammenhangs als *indirect-only-mediation* zwischen der Nutzung der Promptkarten und der Selbstwirksamkeit zur Lesestrategieanwendung könnte statistisch auch darauf zurückzuführen sein, dass sich indirekte

Zusammenhänge unterschiedlicher Vorzeichen aufheben. Im Mediationsmodell ist der Zusammenhang zwischen kognitiver Akzeptanz und der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention zwar nicht signifikant, entgegen den Erwartungen ist dieser jedoch negativ. Entsprechend ist der spezifische indirekte Zusammenhang der kognitiven Akzeptanz im Mediationsmodell negativ. Da der spezifische indirekte Zusammenhang der affektiven Akzeptanz positiv ist, könnte dies zum Ausbleiben des direkten Zusammenhangs zwischen der Nutzung der Promptkarten und der Selbstwirksamkeit beitragen.

Überraschend ist, dass der Zusammenhang zwischen der kognitiven Akzeptanz und der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention negativ ausfällt. Zwar ist der Zusammenhang nicht signifikant, es könnte jedoch darauf hindeuten, dass Schüler*innen, die die Intervention als besonders wirksam wahrnehmen (kognitive Akzeptanz), geringere Ausprägungen der Selbstwirksamkeit am Ende der Intervention aufweisen. Zur weiteren Erklärung dieser Zusammenhänge können die Ergebnisse der zweiten Fragestellung herangezogen werden, in der individuelle Unterschiede in der Nutzung der Promptkarten und Akzeptanzdimensionen durch Lernvoraussetzungen der Schüler*innen untersucht wurden. Entgegen den Erwartungen sagte das Leseverstehen zu Beginn der Intervention beide Akzeptanzdimensionen negativ vorher. Schüler*innen, die im Leseverstehen niedrigere Ergebnisse erzielen, bewerten die Intervention demnach auf affektiver und kognitiver Ebene positiver. Die Nutzung der Promptkarten wurde durch das Leseverstehen nicht signifikant vorhergesagt, allerdings ist auch hier der Zusammenhang wider Erwarten negativ. Die Methode des Reziproken Lehrens hat sich insbesondere für schwächere Schüler*innen als wirksam erwiesen, die zwar Prozesse des Dekodierens schon beherrschen, aber noch Defizite im Textverstehen aufweisen (Palinscar & Brown, 1984). Im Gegensatz zur Untersuchung der Lernwirksamkeit kann die vorliegende Untersuchung die Datenlage dahingehend ergänzen, dass auch die soziale Validität im Rahmen der Akzeptanz bei schwächeren Leser*innen für diese Intervention hoch ausfällt und damit viel Potenzial für positive Lernerlebnisse und die Weiternutzung der Methode gegeben ist. Kritisch muss jedoch auch angemerkt werden, dass entsprechend davon auszugehen ist, dass für Schüler*innen mit höheren Ausprägungen im Leseverstehen eine weniger gute Passung des Schwierigkeitsniveaus der Intervention gegeben sein könnte.

Die Akzeptanzdimensionen werden weiterhin durch die Lesefreude positiv vorhergesagt. Schüler*innen müssen demnach ein gewisses Interesse am Interventionsgegenstand aufweisen, um eine hohe Akzeptanz zu zeigen. Die Lesefreude stellt eine intrinsische Wertzuschreibung des übergeordneten Interventionsgegenstandes Lesen dar und kann daher aus motivationspsychologischer Perspektive als zentraler Faktor für leistungsbezogene Entscheidungen und Handlungen gesehen werden (Wigfield & Eccles, 2000). Schrader et al. (2020) stellen bereits heraus, dass die lerngegenstandsbezogene Motivation einen Prädiktor für den Implementationserfolg darstellt. Zudem wird die Passung zu individuellen Lernvoraussetzungen und eine adressatengerechte Gestaltung einer Intervention in der

Implementationsforschung schon lange als zentral angenommen (Rogers, 2003; Sanetti & Kratochwill, 2009). Die Ergebnisse schließen hieran an und zeigen, dass Schüler*innen mit niedrigerem Leseverstehen und einem hohen Maß an Lesefreude die Intervention besonders positiv bewerteten. Hieraus können wichtige Implikationen für die Nutzung des untersuchten Konzepts in der Praxis abgeleitet werden, auf die in Kapitel 7.3 näher eingegangen wird.

Die Nutzung der Promptkarten wird schließlich durch keine der angenommenen Prädiktoren signifikant vorhergesagt. Es zeigten sich lediglich marginal signifikante Zusammenhänge des Leseverstehens, der Lesefreude und der Einstellungen zu Mehrsprachigkeit. Methodisch ist hierbei zu bedenken, dass mit der dichotomisierten Erfassung eine eingeschränkte Varianz einhergeht, weswegen mögliche Zusammenhänge in der vorliegenden Untersuchung schwieriger aufzudecken sein könnten. Dies stellt ein häufiges Problem bei der Untersuchung der Wiedergabetreue dar, da zwar eine hohe Adhärenz angestrebt wird, dies jedoch auch zu weniger Varianz führt, um die angenommenen Effekte zu erfassen (Low et al., 2014). Weiterhin kann aufgrund der fehlenden Zusammenhänge inhaltlich hinterfragt werden, was mit der Nutzung der Promptkarten gemessen wird und wie dies zukünftig präzisiert werden könnte. Die Einhaltung der Vorgaben im Sinne der Wiedergabetreue könnte auch auf ein gewissenhaftes Lernverhalten zurückzuführen sein (Brisson et al., 2020). Zudem sind interventionsspezifische Kernelemente und Aktivitäten, die hier als Wiedergabetreue erfasst werden, nah an inhaltlichen oder übergeordneten Konstrukten angesiedelt, die zusätzlich erfasst werden könnten, wie z. B. die konkrete Anwendung der Lesestrategien oder die Interaktionen in der Gruppenarbeit. Die Erhebung der Nutzung der Promptkarten allein könnte hier nicht ausgereicht haben, um Effekte aufzudecken.

7.2 Bedeutung der Mehrsprachigkeit für den Implementationserfolg

Im zweiten Teil der Diskussion wird die Rolle der Mehrsprachigkeit für den Implementationserfolg fokussiert. Hierbei wird zum einen der Frage nachgegangen, ob die Mehrsprachigkeit der Schüler*innen und ihre Einstellungen zu Mehrsprachigkeit prädiktiv für die untersuchten Facetten des Implementationserfolgs sind. Zum anderen wird aus einer differenziellen Perspektive für mehrsprachige Schüler*innen analysiert, ob die Aktivierung der Mehrsprachigkeit über die Nutzung der bilingualen Promptkarten in einer anderen Sprache gelungen ist, wie dies mit dem Lernerfolg der Schüler*innen zusammenhängt und welche individuellen Merkmale die Nutzung der Promptkarten in einer anderen Sprache vorhersagen.

7.2.1 Mehrsprachigkeit als Prädiktor für die Facetten des Implementationserfolgs

Eine zentrale Frage der Arbeit war, wie mehrsprachige und monolinguale Schüler*innen die Intervention wahrnehmen und ob mehrsprachige Schüler*innen durch die Intervention besonders angesprochen werden. Aus den Ergebnissen der zweiten Fragestellung geht hervor, dass die Mehrsprachigkeit keinen Zusammenhang zur Nutzung der Promptkarten aufweist. Mehrsprachige Schüler*innen zeigen jedoch ein höheres Maß an affektiver und kognitiver Akzeptanz als monolinguale Schüler*innen. Diese Zusammenhänge werden (teilweise) durch lerngegenstandsbezogene Voraussetzungen der Schüler*innen vermittelt. Der Zusammenhang zwischen der Mehrsprachigkeit und affektiven Akzeptanz wird dabei vollständig durch das Leseverstehen und die Einstellungen zu Mehrsprachigkeit mediiert. Der Zusammenhang zwischen der Mehrsprachigkeit und kognitiven Akzeptanz wird im Rahmen einer partiellen Mediation teilweise durch das Leseverstehen und die Einstellungen zu Mehrsprachigkeit mediiert.

Für die affektive Akzeptanz ist abzuleiten, dass Mehrsprachigkeit für diese nicht direkt prädiktiv ist. Zentral sind vielmehr das Leseverstehen und die Einstellungen zu Mehrsprachigkeit, die sich, wie vermutet, als vermittelnde Variablen zeigen. Aus den Mediationsmodellen geht hervor, dass Mehrsprachigkeit mit einem niedrigeren Leseverstehen und positiveren Einstellungen zu Mehrsprachigkeit einhergeht. Dies knüpft an bestehende Forschung an, nach der Schüler*innen mit Migrationshintergrund häufig schlechtere schulische Leistungen erzielen als ihre monolingualen Mitschüler*innen (Edele et al., 2020). Für diese Schüler*innen scheint die Intervention hinsichtlich der sozialen Validität besonders geeignet zu sein. Die gleichen vermittelnden Zusammenhänge sind auch zwischen der Mehrsprachigkeit und der kognitiven Akzeptanz zu beobachten, hier bleibt der direkte Zusammenhang jedoch unter Kontrolle der Mediatoren bestehen. Mehrsprachige Schüler*innen bewerten die Intervention demnach auch unabhängig ihres Leseverstehens und ihrer Einstellungen zu Mehrsprachigkeit noch als wirksamer für die Verbesserung der Lesekompetenz als monolinguale Schüler*innen. Aufgrund dieser partiellen Mediation schließt sich die Frage nach weiteren mediierten Mechanismen an, die in der vorliegenden Untersuchung nicht erfasst wurden. Schüler*innen mit Migrationshintergrund verfügen häufig über niedrigere sprachliche Kompetenzen im Deutschen, was sich auch in niedrigeren Leistungen zeigt. Sprachliche Kompetenzen könnten eine weitere vermittelnde Variable darstellen und mehrsprachige Schüler*innen, die niedrigere sprachliche Kompetenzen im Deutschen aufweisen, dürften die Intervention als besonders wirksam und nützlich empfinden, da sie zum einen von dem strukturierten Training profitieren und zum anderen die Möglichkeit haben, ihre mehrsprachigen Ressourcen im Lernprozess zu integrieren. Die Nutzung der Familiensprachen könnte dabei zusätzlich moderierend wirken.

Lehrkräfte weisen häufig Ängste und Unsicherheiten in Bezug auf die Integration von Mehrsprachigkeit im Unterricht auf. Ein Aspekt stellt dabei auch die Angst vor möglicher Benachteiligung oder Ausgrenzung monolingualer Schüler*innen dar (vgl. Kapitel 2.1.3). Die Befunde der Arbeit zeigen jedoch, dass Unterschiede in den Ausprägungen der Akzeptanz nicht ausschließlich auf die Mehrsprachigkeit zurückzuführen sind, sondern, wie vermutet, durch proximale Merkmale, die mit dem Lerngegenstand verbunden sind, vermittelt werden. Entsprechend kann vorsichtig abgeleitet werden, dass für monolinguale und mehrsprachige Schüler*innen die Intervention insbesondere auf affektiver Ebene gleichermaßen positiv erlebt werden kann, vielmehr lerngegenstandsbezogene Voraussetzungen und eine Passung zum Interventionsgegenstand von Relevanz sind. Mehrsprachige Schüler*innen zeigen zwar durchschnittlich positivere Einstellungen zu Mehrsprachigkeit, jedoch ist zu beachten, dass mehrsprachige Schüler*innen einen individuellen Umgang mit der eigenen Mehrsprachigkeit aufweisen (Ritterfeld et al., 2014), sodass auch für mehrsprachige Schüler*innen die Passung von Relevanz ist. Die Ergebnisse sind zudem anschlussfähig an erste Ergebnisse zur Durchführung mehrsprachigkeitssensibler Unterrichtskonzepte, in denen eine hohe Beteiligung und positive Bewertung aller Schüler*innen (monolingual und mehrsprachig) beobachtet werden konnte (Schnitzer, 2020). Die Befunde dieser Arbeit ergänzen die Ergebnisse jedoch dahingehend, dass dies systematisch an einer größeren Stichprobe und quantitativ untersucht wurde und sich Unterschiede in der affektiven und kognitiven Akzeptanz finden ließen. Die Akzeptanz fungiert hierbei zusätzlich als systematisches Konstrukt der Implementationsforschung, wodurch zur Vergleichbarkeit weiterer Ergebnisse beigetragen werden kann.

Die Einstellungen zu Mehrsprachigkeit als vermittelnde Variable zeigen weiterhin, dass die Etablierung eines positiven und wertschätzenden Klimas gegenüber Mehrsprachigkeit ein wichtiger Baustein für den Implementationserfolg des mehrsprachigkeitssensiblen Unterrichtskonzepts darstellt und knüpft damit an aktuelle Forschungsarbeiten zum pädagogischen Translanguaging an (Lorenz et al., 2021). Negative Einstellungen zu Mehrsprachigkeit können Ausdruck des *monolingualen Habitus* (Gogolin, 1994) sein, dem durch die Förderung eines wertschätzenden Klimas für Mehrsprachigkeit und dem Aufgreifen sprachlicher Vielfalt in den Klassen entgegengewirkt werden kann. Die deskriptiven Ergebnisse der Untersuchung zeigen, dass sich die Einstellungen der Schüler*innen zu Mehrsprachigkeit im Verlauf der Intervention verbessern. Dies wurde in der vorliegenden Untersuchung nicht im Vergleich zu einer Kontrollgruppe untersucht, weswegen ein kausaler Wirkungszusammenhang nicht bestätigt werden kann. Es kann allerdings angenommen werden, dass das mehrsprachigkeitssensible Reziproke Lehren zur Förderung positiver Einstellungen beitragen kann, die den Implementationserfolg zukünftiger mehrsprachigkeitssensibler Unterrichtskonzepte positiv beeinflussen. Dies gilt es in weiteren Untersuchung zu überprüfen.

7.2.2 Aktivierung der Mehrsprachigkeit durch die Nutzung der bilingualen Promptkarten in einer anderen Sprache

Die Nutzung der bilingualen Promptkarten in einer anderen Sprache ist eines der mehrsprachigkeitssensiblen Elemente der Intervention, mit dem das gesamte sprachliche Repertoire mehrsprachiger Schüler*innen aktiviert werden soll. Die Promptkarten standen den Schüler*innen während der Unterrichtseinheit in ihren Familiensprachen zu Verfügung (vgl. Kapitel 5.1.1 und Kapitel 5.3.1). In Fragestellung 1 und 2 wurde die Nutzung der Promptkarten als Facette der Wiedergabetreue (*participant adherence*) in der Gesamtstichprobe untersucht. Hierbei wurde allgemein nach der Nutzung der Promptkarten während der Intervention gefragt. Für eine differenzielle Perspektive auf die Aktivierung der Mehrsprachigkeit wurden mehrsprachige Schüler*innen zusätzlich gefragt, ob sie die Promptkarten *auf einer anderen Sprache* genutzt haben. In die Auswertungen gingen nur mehrsprachige Schüler*innen ein, die die Promptkarten in (mindestens) einer ihrer Familiensprachen vorliegen hatten (vgl. Kapitel 5.3.1).

Die deskriptiven Ergebnisse zeigen, dass mehrsprachige Schüler*innen mehrheitlich (64,5 %) angaben, die bilingualen Promptkarten in einer anderen Sprache genutzt zu haben. Damit zeigt sich zunächst, dass das mehrsprachigkeitssensible Angebot von vielen mehrsprachigen Schüler*innen angenommen wurde und eine Aktivierung und Einbindung mehrsprachiger Ressourcen hierüber erreicht werden konnte. Bisherige Studien in unterschiedlichen Kontexten zeigen, dass die Aktivierung mehrsprachiger Ressourcen in kooperativen Settings gelingen kann, jedoch oft nur ein geringes Maß an mehrsprachiger Kommunikation zu beobachten ist (Decristan et al., 2021; Meyer & Prediger, 2011; Schastak, 2020). Die bilingualen Promptkarten könnten ein niedrigschwelliges Angebot darstellen, um die Mehrsprachigkeit aktiv im Lernprozess zu integrieren. Es muss von den genannten Studien allerdings abgegrenzt werden, dass die Art der Nutzung der Promptkarten in der vorliegenden Untersuchung nicht weiter aufgeschlüsselt werden kann. Hierbei ist neben einer aktiven Sprachnutzung in der Kleingruppe (z. B. Vorlesen der Strategien oder Diskutieren über die Strategien) auch eine ausschließlich rezeptive Nutzung im Sinne des Lesens denkbar. Zudem wurde in der Intervention ein Sprachlernstift eingesetzt, mit dem die Strategien vorgelesen werden konnten (vgl. Kapitel 5.1.1). Auch hierbei ist denkbar, dass Schüler*innen, die den Sprachlernstift verwendet haben, um die Strategien in den Familiensprachen anzuhören, die Nutzung in einer anderen Sprache angegeben haben. Auch können keine Aussagen über Häufigkeit oder Qualität der Nutzung der Promptkarten in einer anderen Sprache abgeleitet werden. Inwiefern die Promptkarten im Sinne der theoretischen Annahmen auf lernförderliche Art und Weise genutzt wurden, z. B. zur kognitiven Entlastung, Anknüpfung an Vorwissen in der Familiensprache oder als Anregung zur effektiven Kommunikation in der Kleingruppe (vgl. Kapitel 2.1.4), kann daher nicht untersucht werden.

Im Mediationsmodell zeigte sich jedoch, dass die Nutzung der Promptkarten in einer anderen Sprache signifikant mit beiden Akzeptanzdimensionen zusammenhängt (vgl. Abbildung 7). Mehrsprachige Schüler*innen, die die Promptkarten in einer anderen Sprache genutzt haben, bewerten die Intervention demnach positiver und für das Lernen wirksamer, als mehrsprachige Schüler*innen ohne diese Nutzung. Dies deutet darauf hin, dass der Einbezug aller sprachlichen Ressourcen von Schüler*innen als lernförderlich erachtet wird und kann Hinweise auf eine Nutzung im Sinne lernförderlicher Funktionen geben (z. B. García, 2009; Grosjean, 2013; vgl. Kapitel 2.1.4). Dies schließt sich auch an die Untersuchung von Decristan et al. (2022) an, die auf Basis der gleichen Datengrundlage die mehrsprachigkeitssensible Intervention hinsichtlich ihrer Wirksamkeit im Vergleich zu einer Kontrollgruppe untersuchen. Die Ergebnisse zeigen, dass mehrsprachige Schüler*innen, die während der Intervention auch eine andere Sprache genutzt haben, höhere Lernzuwächse im Leseverstehen aufweisen als mehrsprachige Schüler*innen, die nur Deutsch verwendet haben. Die Ergebnisse der vorliegenden Arbeit zeigen daran anknüpfend, dass Schüler*innen die Lernförderlichkeit auch selbst wahrnehmen.

Die Nutzung der Promptkarten in einer anderen Sprache sagt bei mehrsprachigen Schüler*innen die Lesefreude am Ende der Intervention in geringem Maße positiv vorher. Unter Berücksichtigung des Eingangsniveaus weisen demnach mehrsprachige Schüler*innen, die die Promptkarten während der Intervention in einer anderen Sprache genutzt haben, ein höheres Maß an Lesefreude am Ende der Intervention auf, als mehrsprachige Schüler*innen, die keine mehrsprachige Nutzung angegeben haben. Da sich dieser signifikante Zusammenhang mit der allgemeinen Nutzung der Promptkarten in der Gesamtstichprobe nicht zeigte, kann dies darauf hindeuten, dass das zusätzliche mehrsprachigkeitssensible Angebot Potenziale zur Förderung der Lesemotivation mit sich bringt. Dies ist anschlussfähig an Ergebnisse von Schastak (2020) und Reitenbach et al. (2023), bei denen die Aktivierung sprachlicher Ressourcen bei mehrsprachigen Schüler*innen mit dem Erleben positiver Emotionen begründet wird. Es ist zudem denkbar, dass der Zusammenhang durch weitere motivationspsychologische Variablen vermittelt wird, die mit der Lesefreude in Zusammenhang stehen, wie bspw. die Erfüllung von Grundbedürfnissen nach der *Self-Determination-Theory* (Deci & Ryan, 1985). Mehrsprachige Schüler*innen könnten durch die Nutzung der Promptkarten in einer anderen Sprache ein höheres Maß an Autonomie- und Kompetenzerleben aufweisen, da sie ihre Sprachwahl im Lernprozess individuell steuern und ihre sprachlichen Ressourcen einbringen können. Dies könnte die Entwicklung der Lesefreude positiv beeinflussen. Die Zuordnung von Sprachpartnerschaften könnte sich hierbei über die soziale Eingebundenheit ebenso positiv auf das Erleben auswirken.

Unterschiede in der Nutzung der Promptkarten in einer anderen Sprache können für mehrsprachige Schüler*innen schließlich durch die Einstellungen zu Mehrsprachigkeit erklärt werden. Alle anderen untersuchten lerngegenstands-

bezogenen Voraussetzungen zeigen keinen signifikanten Zusammenhang. Die Ergebnisse unterstreichen nochmals die zentrale Bedeutung der Einstellungen zu Mehrsprachigkeit, die sich bereits in der Gesamtstichprobe zeigte. Auch zur Aktivierung der mehrsprachigen Ressourcen scheinen die Einstellungen eine wichtige Voraussetzung zu sein. Zudem wird hierbei verdeutlicht, dass mehrsprachige Schüler*innen mit ihrer Mehrsprachigkeit individuell umgehen und die Mehrsprachigkeit alleine nicht prädiktiv dafür ist, ob die Schüler*innen ihre Sprachen im Unterrichtskontext einbringen.

7.3 Praktische Implikationen

Aus den Ergebnissen der vorliegenden Arbeit können praktische Implikationen für die Implementationsforschung, die Weiterentwicklung des mehrsprachigkeitssensiblen Reziproken Lehrens und den Umgang mit Mehrsprachigkeit im Unterricht abgeleitet werden. Diese werden im Folgenden beschrieben.

Die Untersuchung der Schüler*innenperspektive ist bei der Evaluation von innovativen didaktischen Konzepten besonders praxisrelevant, da eine Intervention auch mit der höchsten Wiedergabetreue durch die Lehrkräfte erfolglos sein kann, wenn Schüler*innen diese im Unterricht nicht akzeptieren oder wie vorgesehen umsetzen. Aufgrund der gefundenen Zusammenhänge zur Selbstwirksamkeit der Lesestrategieanwendung unterstreichen die Ergebnisse die Bedeutung der Facetten des Implementationserfolgs auf Schüler*innenebene für wünschenswerte Veränderungen, die mit der Innovation angestrebt werden. Für zukünftige Forschungsvorhaben kann abgeleitet werden, dass auch die Schüler*innenperspektive in die Evaluation der Implementation und zur Erklärung der Wirksamkeit von Interventionen einfließen sollte. Interventionen sollten so gestaltet sein, dass auch Schüler*innen eine hohe Wiedergabetreue und Akzeptanz aufweisen, als Voraussetzung für eine erfolgreiche Etablierung der Innovationen in der Praxis.

Die Ergebnisse verdeutlichen weiterhin die Relevanz der Passung zwischen Interventionsgegenstand und Schüler*innenvoraussetzungen für eine gelingende Implementation. Verknüpfend mit Untersuchungen der Wirksamkeit kann die Untersuchung der Implementation durch Schüler*innen wichtige Impulse bieten, die Intervention noch besser auf individuelle Bedürfnisse und Lernvoraussetzungen abzustimmen. Daraus können nützliche Implikationen für die Weiterentwicklung der Unterrichtseinheit abgeleitet werden. In der vorliegenden Arbeit zeigt sich, dass eine niedrigere Lesekompetenz und positivere Einstellungen zu Mehrsprachigkeit zentrale Prädiktoren für Akzeptanz darstellen. Hieraus kann gefolgert werden, dass die Intervention für diese Schüler*innengruppe besonders gut geeignet ist. Umgekehrt lässt sich aus den Ergebnissen ableiten, dass Schüler*innen mit höherer Lesekompetenz die Intervention weniger akzeptieren. Als Weiterentwicklung könnten verstärkt differenzierte Lesetexte und individualisierte Möglichkeiten für unterschiedliche Kompetenzstufen im Leseverstehen angeboten werden, um

auch Schüler*innen abzuholen, die im Leseverstehen und in der Strategieranwendung schon kompetenter sind. Praktische Umsetzungsmöglichkeiten könnten hier neben weiteren Abstufungen der Schwierigkeit der Lesetexte auch in einer differenzierteren Rollenverteilung innerhalb der Kleingruppenarbeit gesehen werden. Stärkere Schüler*innen könnten in der Tandemarbeit zusätzlich als Peer-Tutor*in für schwächere Schüler*innen fungieren und konkrete Aufgaben zur Unterstützung schwächerer Schüler*innen übernehmen. Auf diese Weise kann das Kompetenzerleben der bereits kompetenteren Schüler*innen zusätzlich gestärkt werden, was eine gesteigerte intrinsische Motivation und positive Gesamtevaluation der Intervention zur Folge haben könnte (vgl. Deci & Ryan, 1985).

Die Ergebnisse der Arbeit liefern schließlich praktische Implikationen für den Umgang mit Mehrsprachigkeit im Unterricht. Die Einstellungen zu Mehrsprachigkeit zeigen sich in den Analysen als zentrale Variable zur Vorhersage der Akzeptanzdimensionen in der Gesamtstichprobe sowie für die Aktivierung der Mehrsprachigkeit im Rahmen der bilingualen Promptkarten. Die Ergebnisse bestätigen damit die zentrale Rolle eines wertschätzenden Klimas für sprachliche Vielfalt, welches in vielen Ansätzen mehrsprachigkeitssensiblen Unterrichts als wichtige Grundlage gesehen wird (Lorenz et al., 2021; vgl. Kapitel 2.2.1). Eine Hürde für den Einbezug von Mehrsprachigkeit in den Unterricht stellen oftmals negative Einstellungen sowie Ängste und Unsicherheiten bei Lehrkräften dar (vgl. Kapitel 2.1.3). Die Ergebnisse machen jedoch deutlich, dass auch, wenn mehrsprachigkeitssensible Konzepte von Lehrkräften umgesetzt werden, die Einstellungen der Schüler*innen eine weitere Hürde für den Implementationserfolg darstellen können. Damit sollte in der Literatur nicht nur die Förderung positiver Einstellungen der Lehrkräfte, sondern auch der Schüler*innen gefordert werden. An den deskriptiven Ergebnissen ist zu erkennen, dass sich die Einstellungen zu Mehrsprachigkeit über den Verlauf der Intervention bei Schüler*innen verbessert haben. Zwar kann hier nicht im Vergleich zu einer Kontrollgruppe eine Wirkung der Intervention bestätigt werden, die Veränderung liefert jedoch Hinweise darauf, dass die Intervention geeignet sein könnte, um positive Einstellungen von Schüler*innen zu fördern und damit langfristig zum Erfolg weiterer mehrsprachigkeitssensibler Konzepte beizutragen. Lehrkräfte können weiterhin mit einer wertschätzenden Haltung gegenüber sprachlicher Vielfalt und durch Wertschätzung mehrsprachiger Äußerungen von Schüler*innen ihre Einstellungen positiv beeinflussen, da die Interaktion zwischen Lehrkräften und Schüler*innen eine bedeutende Rolle spielt (Sierens & van Avermaet, 2014). Zudem zeigen die Ergebnisse, dass auch monolinguale Schüler*innen die Intervention positiv bewerten und die Mehrsprachigkeit der Schüler*innen nicht direkt prädiktiv für die affektive Akzeptanz ist. Dies könnte Ängste bei Lehrkräften abbauen, monolinguale Schüler*innen durch den Einbezug von Mehrsprachigkeit im Unterricht zu benachteiligen oder gar auszuschließen.

Die Einstellungen zu Mehrsprachigkeit sagen bei mehrsprachigen Schüler*innen die Nutzung der Promptkarten in einer anderen Sprache positiv vorher. Ob-

wohl die Förderung eines wertschätzenden Klimas für Mehrsprachigkeit allgemein angestrebt werden sollte, muss dennoch beachtet werden, dass Schüler*innen einen individuellen Umgang mit der eigenen Mehrsprachigkeit aufweisen (Ritterfeld et al., 2014). Das Angebot zur Nutzung der Familiensprachen im Lernprozess sollte demnach als freiwilliger Baustein gestaltet und über wertschätzende Impulse unterstützt werden, nicht aber verpflichtend sein.

7.4 Limitationen und Ausblick

Bei der Interpretation und Diskussion der Ergebnisse müssen einige Limitationen berücksichtigt werden. Diese werden im Folgenden dargestellt und ausblickende Forschungsfragen abgeleitet, die in zukünftigen Forschungsvorhaben bearbeitet werden sollten.

Als methodische Einschränkung ist zunächst die Operationalisierung der Nutzung der Promptkarten zu nennen, die ausschließlich auf Selbstberichten der Schüler*innen basiert. Untersuchungen bei Lehrkräften zeigen, dass bei Selbstberichten zur Wiedergabetreue die Validität durch soziale Erwünschtheit eingeschränkt sein kann (Dusenbury et al., 2005; Humphrey et al., 2018). Eine Ergänzung um objektive Verfahren, wie z.B. Beobachtungen, wäre auch für die Schüler*innenperspektive gewinnbringend, um diesen Einschränkungen zu begegnen. Eine objektive Erfassung über Beobachtungen stellt jedoch insbesondere für die Wiedergabetreue auf Schüler*innenebene einen großen organisatorischen Aufwand dar, da viele Schüler*innen gleichzeitig beobachtet werden müssten. Videografierungen und Audioaufnahmen, die eine anschließende einzelne Auswertung für alle Schüler*innen ermöglichen, könnten hierbei Möglichkeiten zur Umsetzung darstellen.

Weiterhin ist zu nennen, dass mit der Nutzung der Promptkarten in der vorliegenden Arbeit nur eine Facette der Wiedergabetreue – die Adhärenz eines Kernelements – erfasst wurde. Dabei müssen verschiedene Einschränkungen berücksichtigt werden: Die Nutzung der Promptkarten wurde als dichotomes Item erhoben, welches keine Aussage über die Art, Häufigkeit oder Qualität der Nutzung durch die Schüler*innen zulässt. Zwar ist diese dichotome Erfassung zur Umsetzung von Kernelementen in der Implementationsforschung gängig, Einblicke in aktuelle Forschungsbefunde zeigen jedoch, dass insbesondere auch die qualitative Dimension der Wiedergabetreue für ein differenziertes Bild in den Blick genommen werden sollte, um die Mechanismen zwischen Implementation und Wirksamkeit besser zu verstehen (Dowling & Barry, 2020; Humphrey et al., 2018; Rojas-Andrade & Bahamondes, 2019). In der vorliegenden Untersuchung wird angenommen, dass die Nutzung der Promptkarten lernförderliche Aktivitäten während der Kleingruppenarbeit initiiert (vgl. Kapitel 2.3.3). Dies kann mit den zur Verfügung stehenden Daten allerdings nicht abschließend überprüft werden. Inwiefern die Nutzung der Promptkarten tatsächlich zu qualitativ vollen und lern-

förderlichen Interaktionen führt, sollte weiterführend untersucht werden. Insbesondere für einen differenziellen Blick auf die Aktivierung von Mehrsprachigkeit müsste zudem weiter differenziert werden, wie die Schüler*innen die Promptkarten in ihrer Familiensprache nutzen und ob diese, wie theoretisch angenommen, zur Erleichterung des Verständnisses und kognitiven Entlastung im Lernprozess verwendet werden oder eher einen motivationalen Charakter aufweisen.

Als weitere Einschränkung muss schließlich angeführt werden, dass die Nutzung der Promptkarten durch die Schüler*innen lediglich eins von verschiedenen Kernelementen der Intervention darstellt. Zwar wurde argumentiert, dass die Nutzung der Promptkarten ein besonders zentrales Element in der Gruppenarbeit ist, welches die Bearbeitung der Lesestrategien und den mehrsprachigkeitssensiblen Unterricht verknüpft. Für kooperative Lernsettings sind jedoch weitere Aspekte der Gruppenarbeit relevant, die hier nicht berücksichtigt wurden. Dies bezieht sich bspw. auf die Übernahme von Rollen oder die Bearbeitung der Arbeitsblätter und Lesetexte. Damit könnte ein differenzierteres Bild über die Umsetzung durch die Schüler*innen gezeichnet und weitere Mechanismen bezüglich der Wirksamkeit der Intervention aufgedeckt werden.

Weiterhin muss für die Interpretation der Mediationsmodelle berücksichtigt werden, dass sich die Wirkungsrichtung der Zusammenhänge nicht feststellen lässt und keine kausalen Schlüsse gezogen werden können. Zwar kann theoretisch angenommen werden, dass die Nutzung der Promptkarten *während* der Intervention das Akzeptanzurteil *am Ende* der Intervention beeinflusst, durch die querschnittliche Erhebung beider Variablen zum gleichen Erhebungszeitpunkt kann dies empirisch jedoch nicht geprüft werden. Vor dem Hintergrund konzeptueller Modelle der Implementationsforschung (Witt & Elliott, 1985) und der Zusammenhänge im Angebot-Nutzungsmodell der Wirkweise des Unterrichts (Vieluf et al., 2020) ist jedoch davon auszugehen, dass kein einseitiger Zusammenhang zwischen den Konstrukten besteht und diese sich vielmehr im Verlauf der Intervention wechselseitig beeinflussen. In bisherigen Untersuchungen konnten entsprechend auch unterschiedliche Wirkrichtungen bestätigt werden. Mautone et al. (2009) zeigen, dass die Wiedergabetreue von Lehrkräften während einer Leseintervention für Lernende mit ADHS mit ihrer Akzeptanz am Ende der Intervention zusammenhängt. Eine Untersuchung von Henninger (2010) stellt dagegen heraus, dass die Akzeptanz der Lehrkräfte zu Beginn der Intervention die Wiedergabetreue während der Intervention vorhersagt. Um die komplexen Zusammenhänge besser zu verstehen, könnte es für zukünftige Forschungsvorhaben sinnvoll sein, die Wiedergabetreue und Akzeptanz als dynamische Konstrukte zu mehreren Zeitpunkten während der Intervention zu erfassen (Eckert et al., 2021). Eine verlaufsorientierte Untersuchung kann zudem mehr Aufschluss darüber bieten, welche Elemente der Intervention von Schüler*innen besonders positiv und wirksam empfunden werden und bei welchen Elementen sich Hürden zeigen.

Auch die Selbstwirksamkeit der Lesestrategieanwendung wurde über Selbstberichte der Schüler*innen erfasst, da es sich hierbei um ein latentes Konstrukt

handelt. Jedoch muss beachtet werden, dass daraus nicht eindeutig abgeleitet werden kann, dass die Schüler*innen die Strategien auch tatsächlich gut anwenden können. Eine zusätzliche objektive Messung anhand eines Strategietests wäre sinnvoll, um das Argument zu stärken, dass die Facetten des Implementationserfolgs mit der Wirksamkeit der Intervention in Zusammenhang stehen. Bestehende Ergebnisse zeigen, dass die Selbstwirksamkeit geringe bis mittlere Korrelationen zu Strategietests aufweisen (z. B. Koch & Spörer, 2016).

In den Ergebnissen zeigt sich für die affektive und kognitive Akzeptanzdimension häufig ein ähnliches Muster, was auf die hohe Korrelation der beiden Konstrukte zurückzuführen sein könnte. Dies ist anknüpfbar an Ergebnisse der Einstellungsforschung, wonach Komponenten von Einstellungen hoch korreliert und empirisch nicht immer in unterschiedliche Dimensionen zu trennen sind (z. B. Schulze et al., 2019). Die Ergebnisse der Faktorenanalyse bestätigen jedoch, dass ein zweidimensionales Modell eine bessere Passung auf die Daten aufweist als ein gemeinsamer Faktor (vgl. Kapitel 5.3.2). Insgesamt geben die hohen Korrelationen jedoch Hinweise darauf, dass eine stärkere theoretische Abgrenzung notwendig sein könnte, um die Konstrukte trennscharf zu erfassen. Auch potenzielle Probleme der Multikollinearität könnten dadurch reduziert werden.

Aus den Ergebnissen der vorbereitenden Analysen geht hervor, dass ein substanzieller Anteil der Gesamtvarianz der Akzeptanzdimensionen und Nutzung der Promptkarten auf Klassenebene zu verorten ist. Damit wird die kontextuelle Rahmung der Implementation durch die Lehrkraft verdeutlicht. Zudem kann durch die untersuchten individuellen Merkmale jeweils nur ein geringer Anteil an Varianz in den Facetten des Implementationserfolgs erklärt werden. Dennoch wurden die Analysen in der vorliegenden Arbeit ausschließlich auf Individualebene durchgeführt und die hierarchische Struktur der Daten nur durch die Korrektur der Standardfehler berücksichtigt, da die Stichprobe auf der Klassenebene für eine mehrbenenanalytische Auswertung nicht ausreicht. Durch die Abhängigkeit zur Klassenebene zeigt sich jedoch, dass in weiteren Untersuchungen Merkmale der Lehrkraft, Klasse und des Kontextes integriert werden sollten, zumal die Bedeutung der Berücksichtigung von Kontextfaktoren in der Implementationsforschung bereits etabliert ist (Durlak & DuPre, 2008).

Die Wiedergabetreue der gesamten Intervention durch die Lehrkräfte wurde in der vorliegenden Arbeit nur zum Zweck der vorbereitenden Datenanalyse und auf Basis eines Umsetzungsindex aller Kernelemente untersucht. Unterschiede zwischen den Klassen, die sich auf die Implementation durch die Schüler*innen und Wirksamkeit der Intervention auswirken, sind hinsichtlich der Unterrichtsqualität und Anleitung durch die Lehrkraft in der Gruppenarbeit denkbar. Die Methode des Reziproken Lehrens ist eine anspruchsvolle Instruktionmethode, in der die Rolle der Lehrkraft von zentraler Bedeutung ist. Das Maß und die Qualität an Anleitung und Unterstützung bei der Strategieanwendung in der Gruppenarbeit könnten zwischen den Lehrkräften variieren. Aus der Unterrichtsforschung ist zudem bekannt, dass die Unterrichtsqualität von großer Bedeutung für den

Lernerfolg der Schüler*innen ist (Fauth et al., 2014). Es ist daher denkbar, dass die Unterrichtsqualität auch mit dem Implementationserfolg auf Schüler*innenebene zusammenhängt. Adl-Amini (2018) stellt heraus, dass die Wiedergabetreue auf Klassenebene beim tutoriellen Lernen in der Grundschule mit dem Klassenklima und kognitiven Voraussetzungen der Klasse zusammenhängt. Implementationserfolg kann demnach als interaktiver Prozess interpretiert werden, der sich auch in Abhängigkeit der Mitschüler*innen ausbildet. Die Akzeptanz könnte entsprechend mit der Klassenkomposition, bspw. dem Anteil an mehrsprachigen Schüler*innen oder den Einstellungen zu Mehrsprachigkeit auf Klassenebene zusammenhängen. Kombinierte Untersuchungen der Implementation durch die Lehrkraft und der daraus resultierenden Implementation auf Schüler*innenebene könnten schließlich mehr Aufschluss über diese Zusammenhänge geben.

Ebenso sind Abhängigkeiten vom kooperativen Setting denkbar, die in zukünftigen Studien adressiert werden könnten. In der vorliegenden Untersuchung kamen die Promptkarten vorrangig im Rahmen der Gruppenarbeit zum Einsatz. Je nachdem, ob und wie die Promptkarten hierbei genutzt werden, kann die Gestaltung des Lernangebots zwischen den Gruppen variieren. Auch für die Untersuchung der Mehrsprachigkeit könnte es relevant sein, die Gruppenebene zu betrachten. Schastak (2020) führt aus, dass bilinguale Interaktion beim außerunterrichtlichen Peer-Learning von der Dyade abhängig ist. Weiterhin zeigt sich in einer Untersuchung von Reitenbach et al. (2023), dass die Nutzung der Familiensprachen beim mehrsprachigkeitssensiblen Reziproken Lehren von mehrsprachigen Schüler*innen häufig mit der Kompetenz der Kleingruppe argumentiert wird. Die Nutzung der Promptkarten könnte demnach auch vom Vorhandensein einer Sprachpartnerschaft in der Gruppe abhängen. Schüler*innen können zwar einen individuellen Umgang mit der eigenen Mehrsprachigkeit aufweisen, mehrsprachige Sprecher*innen wählen jedoch unterschiedliche Sprachmodi in Abhängigkeit des Kontextes, der Gesprächspartner*innen und der Themen (Grosjean, 2020). Die Untersuchung der ICC (1) (vgl. Kapitel 5.4.1) zeigt zudem eine stärkere Abhängigkeit der Nutzung der Promptkarten in einer anderen Sprache zur Klassenebene als das Item zur allgemeinen Nutzung der Promptkarten. Dies könnte ein Hinweis darauf sein, dass die Aktivierung der Mehrsprachigkeit in besonderem Maße vom Kontext abhängig ist und somit kontextbezogene Faktoren in die weitere Untersuchung einbezogen werden sollten.

In die Analysen der dritten und vierten Fragestellung gingen nur Schüler*innen ein, die die Promptkarten in mindestens einer ihrer Familiensprachen vorliegen hatten. Für 17 Schüler*innen war dies nicht der Fall, da für diese seltenen Sprachen keine Übersetzungen umsetzbar waren (vgl. Kapitel 5.3.1). Diese Schüler*innen wurden aus der Substichprobe mehrsprachiger Schüler*innen ausgeschlossen. Ergänzende Analysen zeigen jedoch, dass etwa die Hälfte ($n = 9$) dieser Schüler*innen dennoch angab, die Promptkarten in einer anderen Sprache genutzt zu haben. Daran schließt sich die Frage an, ob für alle Schüler*innen angenommen werden kann, dass sie die Promptkarten in ihrer eigenen Familien-

sprache oder einer Sprache, die sie verstehen, genutzt haben. Eine Präzisierung der Itemformulierung könnte in zukünftigen Studien diese unklaren Arten der Nutzung besser abgrenzen. Zudem muss diskutiert werden, inwieweit die Items zur Nutzung der Promptkarten trennscharf zu unterscheiden sind. Die Nutzung der Promptkarten auf Deutsch wurde in der Itemformulierung nicht spezifiziert. Mehrsprachige Schüler*innen könnten demnach auch ausschließlich die Promptkarten in einer anderen Sprache genutzt haben, dies kann in der vorliegenden Untersuchung nicht differenziert werden. Ergänzende Analysen zeigen jedoch, dass die Items zur Nutzung der Promptkarten in der Substichprobe mehrsprachiger Schüler*innen in mittlerem Maße korrelieren ($r = .48$, $p < .001$). Es kann daher davon ausgegangen werden, dass mehrsprachige Schüler*innen zwischen einer deutschen und mehrsprachigen Nutzung der Promptkarten differenzieren.

Insgesamt weisen die Schüler*innen hohe Ausprägungen auf den beiden Akzeptanzdimensionen auf, was auf eine positive Evaluation der Intervention durch die Schüler*innen schließen lässt. Die hohen Ausprägungen deuten jedoch auf Deckeneffekte hin, die mit einer geringeren Varianz einhergehen und die Aufdeckung von Zusammenhängen erschweren können (Bracken, 2007). In zukünftigen Forschungsarbeiten sollten zudem auch nicht-lineare Zusammenhänge zu den Facetten des Implementationserfolgs berücksichtigt werden. Die Bedeutung der Passung der individuellen Schüler*innenvoraussetzungen zum Lernangebot wurde als ein zentrales Ergebnis herausgestellt. Es ist hierbei denkbar, dass es ein optimales Niveau der Passung gibt und bspw. sehr schwache Schüler*innen, die hierarchieniedrigere Prozesse des Dekodierens noch nicht beherrschen, kein hohes Maß an Akzeptanz aufweisen. Für diese Schüler*innen könnten die Aufgaben zu schwierig sein und überfordernd wirken, sodass sie keine Freude an der Intervention haben und diese nicht als wirksam wahrnehmen.

Die Messinvarianzanalysen geben schließlich Hinweise darauf, dass die Konstrukte der Selbstwirksamkeit und Einstellungen zu Mehrsprachigkeit zwischen den Gruppen mehrsprachiger und monolingualer Schüler*innen nicht invariant sind. Zwar wurden die Gruppen in der vorliegenden Arbeit nicht direkt miteinander verglichen, jedoch ist für die Interpretation der Ergebnisse aus der Gesamtstichprobe und der Substichprobe mehrsprachiger Schüler*innen zu berücksichtigen, dass hier unterschiedliche Konstrukte vorliegen könnten. Für die Einstellungen zu Mehrsprachigkeit kann hieraus auch abgeleitet werden, dass sich das verwendete Instrument nur bedingt eignet, um Einstellungen bei mehrsprachigen und monolingualen Schüler*innen einheitlich zu erfassen. Damit alle Schüler*innen die Items beantworten können, bezieht sich die Skala auf sprachliche Vielfalt im Allgemeinen und nicht auf die eigene Mehrsprachigkeit der Schüler*innen. Weitere Untersuchungen könnten sich mit der Frage beschäftigen, wie die Einstellungen für beide Gruppen präzise erfasst werden können.

Zuletzt muss berücksichtigt werden, dass die Ergebnisse der Untersuchung auf einer Teilpopulation aus der Grundschule basieren und nicht auf alle Schüler*innen übertragen werden können. Zum einen wurden durch die Rekrutie-

rung vorrangig Lehrkräfte angesprochen, die Interesse am Thema Mehrsprachigkeit haben und die Notwendigkeit zur Auseinandersetzung mit diesem Thema in ihren Klassen sehen. Dies kann die Ergebnisse dahingehend verzerren, dass auch die Schüler*innen in den Klassen bereits Anknüpfungspunkte mit Mehrsprachigkeit im Unterricht aufweisen oder durch ihre Lehrkräfte ein gesteigertes Interesse am Thema haben. Zudem wurden mögliche Unterschiede im allgemeinen Umgang mit Mehrsprachigkeit in den Klassen nicht erhoben, die sich auf die Aktivierung der Familiensprachen auswirken könnten.

7.5 Fazit

In der vorliegenden Arbeit wurden Facetten des Implementationserfolgs beim mehrsprachigkeitssensiblen Reziproken Lehren aus Schüler*innenperspektive untersucht. Hierbei wurden Fragestellungen der Implementationsforschung und zum Umgang mit sprachlicher Heterogenität im Unterricht miteinander verknüpft. Mehrsprachigkeitssensible Unterrichtskonzepte werden benötigt, um der strukturellen Benachteiligung, die mehrsprachige Schüler*innen durch einen monolingualdeutsch ausgerichteten Unterricht erfahren, entgegenzuwirken. Bislang ist allerdings nur wenig darüber bekannt, wie diese in der Unterrichtspraxis implementiert werden können und welche förderlichen und hemmenden Faktoren beachtet werden müssen. In der Implementationsforschung wird häufig die Perspektive der Lehrkraft und das von ihr gestaltete Unterrichtsangebot fokussiert. Hierbei werden jedoch die Prozessebene des Unterrichts und die komplexen Wahrnehmungs- und Nutzungsprozesse auf Schüler*innenebene vernachlässigt, die eine besonders hohe Relevanz für den Lernerfolg aufweisen.

Die Ergebnisse der Untersuchung machen deutlich, dass Schüler*innen eine zentrale Rolle im Implementationsprozess einnehmen. Wenn auch nicht alle untersuchten Schüler*innenoutcomes Zusammenhänge zu den untersuchten Facetten des Implementationserfolgs gezeigt haben, ist an der Selbstwirksamkeit der Lesestrategieanwendung zu erkennen, dass ein höheres Maß an Akzeptanz mit einem höheren Lernerfolg einhergeht. Interventionen sollten entsprechend so gestaltet sein, dass nicht nur Lehrkräfte diese positiv und implementierbar bewerten, sondern auch auf Ebene der Schüler*innen Implementationserfolg zu beobachten ist. Für eine hohe Akzeptanz auf Schüler*innenebene zeigte sich zudem die Passung kognitiver und motivational-affektiver Merkmale als bedeutsam. Die Ergebnisse können entsprechend dazu beitragen, die zentrale Rolle der Schüler*innen im Implementationsprozess besser zu verstehen und in zukünftiger Forschung sowie der Konzeption von didaktischen Konzepten zu berücksichtigen.

Weiterhin können an den Ergebnissen der vorliegenden Untersuchung wichtige Implikationen für den Umgang mit Mehrsprachigkeit in sprachlich heterogenen Klassen abgeleitet werden. Untersuchungen zu mehrsprachigkeitssensiblen Unterricht stellen dar, dass die Einstellungen der Lehrkräfte entscheidend dafür sind,

ob Mehrsprachigkeit in den Unterricht integriert wird und Konzepte überhaupt umgesetzt werden. Die Ergebnisse der vorliegenden Untersuchung gehen einen Schritt weiter und zeigen, dass Merkmale der Schüler*innen, wie ihre Einstellungen zu Mehrsprachigkeit, nach der Implementation durch die Lehrkraft weiterhin bedeutend dafür sind, wie die Umsetzung durch die Schüler*innen gelingt und sich mehrsprachige Ressourcen aktivieren lassen. Die Implementation ist folglich mit dem Handeln der Lehrkräfte nicht abgeschlossen und der Implementationserfolg kann durch Schüler*innen beeinflusst werden.

Die Integration von Mehrsprachigkeit im Unterricht ist bei Lehrkräften häufig mit der Angst verbunden, an Autorität und Kontrolle zu verlieren oder als Reaktion auf die Heterogenität im Unterricht weitere Ungleichheiten zu produzieren und bspw. monolinguale Schüler*innen zu benachteiligen. Die Ergebnisse der Untersuchung können diesen Ängsten vorsichtig begegnen, da die Mehrsprachigkeit der Schüler*innen nicht allein prädiktiv dafür ist, wie sie das mehrsprachigkeits-sensible Lesetraining im Sinne der Akzeptanz bewerten. Vielmehr wurde deutlich, dass lerngegenstandsbezogene Voraussetzungen, wie das Leseverstehen und die Einstellungen zu Mehrsprachigkeit von Relevanz sind. Die Förderung einer wertschätzenden Atmosphäre gegenüber Mehrsprachigkeit kann demnach zum Gelingen mehrsprachigkeitssensibler Unterrichtsangebote beitragen. Zudem stellen die Ergebnisse heraus, dass die Nutzung mehrsprachigkeitssensibler Impulse, wie die bilingualen Promptkarten in einer anderen Sprache, motivationale Potenziale im Lernprozess entfalten können und die Aktivierung des sprachlichen Repertoires als lernförderlich wahrgenommen wird. Die Integration und Förderung von Mehrsprachigkeit sollte demnach weiter angestrebt werden, um fachliche und überfachliche Potenziale auszuschöpfen. Insgesamt wurde deutlich, dass die Untersuchung des Implementationserfolgs über Effektivitätsuntersuchungen hinaus wichtige Hinweise für die Weiterentwicklung didaktischer Konzepte bieten kann.

In der vorliegenden Arbeit wurden gleichwohl auch methodische Hürden deutlich, die sich bei der Untersuchung von Facetten des Implementationserfolgs ergeben können. Zukünftige Untersuchungen könnten von einer stärkeren Verzahnung der Implementations- und Lehr-Lernforschung profitieren, um Konzeptualisierungen und Operationalisierungen von Implementationsprozessen aus dem klinischen Bereich für die komplexen Prozesse im Unterricht zu adaptieren.

Literatur

- Abry, T., Hulleman, C. S. & Rimm-Kaufman, S. E. (2015). Using Indices of Fidelity to Intervention Core Components to Identify Program Active Ingredients. *American Journal of Evaluation*, 36(3), 320–338. <https://doi.org/10.1177/1098214014557009>
- Adesope, O. O., Lavin, T., Thompson, T. & Ungerleider, C. (2010). A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism. *Review of Educational Research*, 80(2), 207–245. <https://doi.org/10.3102/0034654310368803>
- Adl-Amini, K. (2018). *Tutorielles Lernen im naturwissenschaftlichen Sachunterricht der Grundschule: Umsetzung und Wirkung*. Waxmann.
- Ahrenholz, B. (2020). Erstsprache – Zweitsprache – Fremdsprache – Mehrsprachigkeit. In B. Ahrenholz & I. Oomen-Welke (Hrsg.), *Deutschunterricht in Theorie und Praxis: Band 9. Deutsch als Zweitsprache* (5. Aufl., S. 3–20). Schneider Verlag Hohengehren.
- Alexander, P. A. & The Disciplined Reading and Learning Research Laboratory (2012). Reading into the Future: Competence for the 21st Century. *Educational Psychologist*, 47(4), 259–280. <https://doi.org/10.1080/00461520.2012.722511>
- Allinder, R. M. & Oats, R. G. (1997). Effects of Acceptability on Teachers' Implementation of Curriculum-Based Measurement and Student Achievement in Mathematics Computation. *Remedial and Special Education*, 18(2), 113–120. <https://doi.org/10.1177/074193259701800205>
- Altrichter, H. & Wiesinger, S. (2004). Der Beitrag der Innovationsforschung im Bildungswesen zum Implementierungsproblem. In G. Reinmann & H. Mandl (Hrsg.), *Psychologie des Wissensmanagements: Perspektiven, Theorien und Methoden* (S. 220–233). Hogrefe.
- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., Schneider, W., Stanat, P., Ostermeier, C., Schiefele, U., Valtin, R. & Ring, K. (2007). *Förderung von Lesekompetenz – Expertise*. Bildungsforschung Band 17. Bundesministerium für Bildung und Forschung.
- Artelt, C., Naumann, J. & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 73–112). Waxmann.
- Artelt, C., Schiefele, U., Schneider, W. & Stanat, P. (2002). Leseleistungen deutscher Schülerinnen und Schüler im internationalen Vergleich (PISA): Ergebnisse und Erklärungen. *Zeitschrift für Erziehungswissenschaft*, 5(1), 6–27. <https://doi.org/10.1007/s11618-002-0002-1>
- Baker, C. (2011). *The Foundations of Bilingual Education and Bilingualism* (3rd Edition). Multilingual Matters.
- Bandura, A. (1977). Self-Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295x.84.2.191>
- Baron, R. M. & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.
- Becker-Mrotzek, M., Lindauer, T., Pfost, M., Weis, M., Strohmaier, A. & Reiss, K. (2019). Lesekompetenz heute – eine Schlüsselqualifikation im Wandel. In K. Reiss, M. Weis, E. Klieme & O. Köller (Hrsg.), *PISA 2018: Grundbildung im internationalen Vergleich* (S. 21–46). Waxmann.

- Best, H. & Wolf, C. (2012). Modellvergleich und Ergebnisinterpretation in Logit- und Probit-Regressionen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 64(2), 377–395. <https://doi.org/10.1007/s11577-012-0167-4>
- Binanzer, A. & Jessen, S. (2020). Mehrsprachigkeit in der Schule – aus der Sicht migrationsbedingt mehrsprachiger Jugendlicher. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 25(1), 221–252.
- Blase, K. A. & Fixsen, D. L. (2013). Core Intervention Components: Identifying and Operationalizing What Makes Programs Work. *ASPE Research Brief*, 1–22.
- Bloomfield, L. (1984). *Language*. The University of Chicago Press.
- Borsch, F. (2018). *Kooperatives Lernen: Theorie – Anwendung – Wirksamkeit*. Kohlhammer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Auflage). Springer.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R., Voss, A. & Walther, G. (Hrsg.). (2005). *IGLU: Skalenhandbuch zur Dokumentation der Erhebungsinstrumente*. Waxmann.
- Bos, W., Strietholt, R., Goy, M., Stubbe, T. C., Tarelli, I. & Hornberg, S. (Hrsg.). (2010). *IGLU 2006: Dokumentation der Erhebungsinstrumente*. Waxmann. <https://doi.org/10.25656/01:15670>
- Böse, S., Neumann, M., Becker, M., Maaz, K. & Baumert, J. (2019). Empirische Arbeit: Bereit für Veränderungen? Zum Zusammenhang von Akzeptanz und handlungsbezogener Auseinandersetzung von Schulleiterinnen und Schulleitern in Reformprozessen. *Psychologie in Erziehung und Unterricht*, 66, 51–68. <https://doi.org/10.2378/peu2019.art04d>
- Bosma, E., Bakker, A., Zenger, L. & Blom, E. (2022). Supporting the Development of the Bilingual Lexicon through Translanguaging: A Realist Review Integrating Psycholinguistics with Educational Sciences. *European Journal of Psychology of Education*, 38, 1–23. <https://doi.org/10.1007/s10212-021-00586-6>
- Bracken, B. A. (2007). Clinical Observation of Preschool Assessment Behavior. In B. A. Bracken & R. J. Nagle (Eds.), *Psychoeducational Assessment of Preschool Children* (S. 95–110). Lawrence Erlbaum Associates.
- Braun-Thürmann, H. (2005). *Innovation*. transcript.
- Bredthauer, S. & Engfer, H. (2018). *Natürlich ist Mehrsprachigkeit toll! Aber was hat das mit meinem Unterricht zu tun?* <https://kups.ub.uni-koeln.de/8092/>
- Brehmer, B. & Mehlhorn, G. (2018). *Herkunftssprachen*. Narr Francke Attempo.
- Bremerich-Vos, A., Wendt, H. & Bos, W. (2017). Lesekompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79–142). Waxmann.
- Brisson, B. M., Hulleman, C. S., Häfner, I., Gaspard, H., Flunger, B., Dicke, A.-L., Trautwein, U. & Nagengast, B. (2020). Who Sticks to the Instructions – and Does it Matter? Antecedents and Effects of Students' Responsiveness to a Classroom-Based Motivation Intervention. *Zeitschrift für Erziehungswissenschaft*, 23, 121–144. <https://doi.org/10.1007/s11618-019-00922-z>
- Brown, T. (2006). *Confirmatory Factor Analysis for Applied Research*. The Guilford Press.
- Browne, M. W. & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, 21(2), 230–258. <https://doi.org/10.1177/0049124192021002005>

- Brühwiler, C. & Blatchford, P. (2011). Effects of Class Size and Adaptive Teaching Competency on Classroom Processes and Academic Outcome. *Learning and Instruction*, 21, 95–108. <https://doi.org/10.1016/j.learninstruc.2009.11.004>
- Bruner, J. (1985). Vygotsky: A Historical and Conceptual Perspective. In J. V. Wertsch (Eds.), *Culture, Communication and Cognition: Vygotskian Perspectives* (S. 21–33). Cambridge University Press.
- Budde, M. (2012). *Über Sprache reflektieren: Unterricht in sprachheterogenen Lerngruppen*. Kassel University Press.
- Bührig, K. & Duarte, J. (2013). Zur Rolle lebensweltlicher Mehrsprachigkeit für das Lernen im Fachunterricht – ein Beispiel aus einer Videostudie der Sekundarstufe II. *Gruppendynamik & Organisationsberatung*, 44, 245–275. <https://doi.org/10.1007/s11612-013-0217-z>
- Büttner, G., Warwas, J. & Adl-Amini, K. (2012). Kooperatives Lernen und Peer Tutoring im inklusiven Unterricht. *Zeitschrift für Inklusion* (1-2), 1–14. <https://doi.org/10.25656/01:5877>
- Camenzuli, R., Lundberg, A. & Gauci, P. (2022). Collective Teachers' Beliefs about Multilingualism in Maltese Primary Education. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2022.2114787>
- Cenoz, J. (2013). Defining Multilingualism. *Annual Review of Applied Linguistics*, 33, 3–18. <https://doi.org/10.1017/S026719051300007X>
- Cenoz, J. & Gorter, D. (2021). *Pedagogical Translanguaging*. Cambridge University Press. <https://doi.org/10.1017/9781009029384>
- Cenoz, J., Santos, A. & Gorter, D. (2022). Pedagogical Translanguaging and Teachers' Perceptions of Anxiety. *International Journal of Bilingual Education and Bilingualism*, 1–12. <https://doi.org/10.1080/13670050.2021.2021387>
- Century, J. & Cassata, A. (2016). Implementation Research: Finding Common Ground on What, How, Why, Where, and Who. *Review of Research in Education*, 40(1), 169–215. <https://doi.org/10.3102/0091732X1666533>
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Cheung, G. W. & Rensvold, R. B. (1999). Testing Factorial Invariance Across Groups: A Reconceptualization and Proposed New Method. *Journal of Management*, 25(1), 1–27. <https://doi.org/10.1177/014920639902500101>
- Chi, M. T. H. & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Conteh, J. & Meier, G. (Eds.). (2014). *The Multilingual Turn in Languages Education: Opportunities and Challenges*. Multilingual Matters.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Press Syndicate of the University of Cambridge.
- Crawford, L., Carpenter, D. M., Wilson, M. T., Schmeister, M. & McDonsals, M. (2012). Testing the Relation Between Fidelity of Implementation and Student Outcomes in Math. *Assessment for Effective Intervention*, 37(4). <https://doi.org/10.1177/1534508411436111>
- Creese, A. & Blackledge, A. (2010). Translanguaging in the Bilingual Classroom: A Pedagogy for Learning and Teaching? *The Modern Language Journal*, 94(1), 103–115.

- Cummins, J. (2000). *Language, Power and Pedagogy: Bilingual Children in the Crossfire. Bilingual education and bilingualism: Bd. 23. Multilingual Matters.*
- Dane, A. V. & Schneider, B. H. (1998). Program Integrity in Primary and Early Secondary Prevention: Are Implementation Effects out of Control? *Clinical Psychology Review, 18*(1), 23–45.
- Deci, E. L. & Ryan, R. M. (1985). Conceptualizations of Intrinsic Motivation and Self-Determination. In E. L. Deci & R. M. Ryan (Eds.), *Intrinsic Motivation and Self-Determination in Human Behavior* (S. 11–40). Springer-Verlag US. https://doi.org/10.1007/978-1-4899-2271-7_2
- Decristan, J., Bertram, V., Reitenbach, V., Schneider, K. M., Köhler, C. & Rauch, D. (2022). Linguistically Responsive Reciprocal Teaching in Primary School: Effectiveness of an Intervention Study on Students' Reading Competence. *Journal of Multilingual and Multicultural Development, 1–19*. <https://doi.org/10.1080/01434632.2022.2141757>
- Decristan, J., Schastak, M., Reitenbach, V. & Rauch, D. (2021). Außerunterrichtliches Peer Tutoring mit deutsch-türkischsprachigen Grundschulkindern: Umsetzungsgenauigkeit und Umfang von bilingualer Kommunikation. *Zeitschrift für Erziehungswissenschaft, 24*, 841–860. <https://doi.org/10.1007/s1618-021-01023-6>
- Dirim, I. (2010). „Wenn man mit Akzent spricht, denken die Leute, dass man auch mit Akzent denkt oder so.“ Zur Frage des (Neo-)Linguizismus in den Diskursen über Sprache(n) der Migrationsgesellschaft. In P. Mecheril (Hrsg.), *Spannungsverhältnisse. Assimiliationsdiskurse und interkulturelle pädagogische Forschung* (S. 91–114). Waxmann.
- Dirim, I. & Wegener, A. (2016). Bildungsgerechtigkeit und der Umgang mit migrationsbedingter Mehrsprachigkeit in Schule und Unterricht. In A. Wegener & I. Dirim (Hrsg.), *Mehrsprachigkeit und Bildungsgerechtigkeit. Erkundungen einer didaktischen Perspektive* (S. 198–223). Verlag Barbara Budrich.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht: Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Zeitschrift für Pädagogik: Beiheft 41. Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule* (S. 73–92). Beltz. <https://doi.org/10.25656/01:8486>
- Donker, A. S., Boer, H. de, Kostons, D., Dignath van Ewijk, C. C. & van der Werf, M. P. C. (2014). Effectiveness of Learning Strategy Instruction on Academic Performance: A Meta-Analysis. *Educational Research Review, 11*, 1–26. <https://doi.org/10.1016/j.edurev.2013.11.002>
- Dowling, K. & Barry, M. M. (2020). The Effects of Implementation Quality of a School-Based Social and Emotional Well-Being Program on Students' Outcomes. *European Journal of Investigation in Health, Psychology and Education, 10*(2), 595–614. <https://doi.org/10.3390/ejihpe10020044>
- Dumas, J. E., Lynch, Anne M., Laughlin, James E., Phillips Smith, E. & Prinz, R. J. (2001). Promoting Intervention Fidelity: Conceptual Issues, Methods, and Preliminary Results from the Early Alliance Prevention Trial. *American Journal of Preventive Medicine, 20*(1), 38–47.
- Durlak, J. A. & DuPre, E. P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology, 41*(3), 327–350.

- Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J. & Falco, M. (2005). Quality of Implementation: Developing Measures Crucial to Understanding the Diffusion of Preventive Interventions. *Health Education Research*, 20(3), 308–313. <https://doi.org/10.1093/her/cygl34>
- Ebe, A. E. & Chapman-Santiago, C. (2016). Student Voices Shining Through: Exploring Translanguaging as a Literary Device. In O. García & T. Kley (Eds.), *Translanguaging with Multilingual Students: Learning from Classroom Moments* (S. 57–82). Routledge.
- Eckert, T. L., Eggleston, B. N., Hamsho, N. F. & Williams, N. M. (2021). Assessing Children's Perceptions of Intervention Acceptability Ratings. *Psychology in the Schools*, 58(10), 1962–1979. <https://doi.org/10.1002/pits.22564>
- Eckert, T. L., Hier, B. O., Hamsho, N. F. & Malandrino, R. D. (2017). Assessing Children's Perceptions of Academic Interventions: The Kids Intervention Profile. *School Psychology Quarterly*, 32(2), 268–281. <https://doi.org/10.1037/spq0000200>
- Eckert, T. L. & Hintze, J. M. (2000). Behavioral Conceptions and Applications of Acceptability: Issues Related to Service Delivery and Research Methodology. *School Psychology Quarterly*, 15(2), 123–148. <https://doi.org/10.1037/h0088782>
- Edele, A., Kempert, S. & Stanat, P. (2020). Mehrsprachigkeit und Bildungserfolg. In I. Gogolin, A. Hansen, S. McMonagle & D. Rauch (Hrsg.), *Handbuch Mehrsprachigkeit und Bildung* (S. 151–155). Springer Fachmedien.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2010). *Statistik und Forschungsmethoden: Lehrbuch*. Beltz.
- Eisenchlas, S. A. & Schalley, A. C. (2020). Making Sense of “Home Language” and Related Concepts. In A. C. Schalley & S. A. Eisenchlas (Eds.), *Handbook of Home Language Maintenance and Development: Social and Affective Factors* (S. 17–37). Walter de Gruyter.
- Esser, H. (2009). Der Streit um die Zweisprachigkeit: Was bringt die Bilingualität? In I. Gogolin & U. Neumann (Hrsg.), *Streifzug Zweisprachigkeit – The Bilingual Controversy* (S. 69–88). VS Verlag für Sozialwissenschaften.
- Euler, D. & Sloane, P. F. E. (1998). Implementation als Problem der Modellversuchsforschung. *Unterrichtswissenschaft*, 26(4), 312–326. <https://doi.org/10.25656/01:7777>
- Evans, M., Kelley, J., Sikora, J. & Treiman, D. J. (2010). Family Scholarly Culture and Educational Success: Books and Schooling in 27 Nations. *Research in Social Stratification and Mobility*, 28(2), 171–197. <https://doi.org/10.1016/j.rssm.2010.01.002>
- Faber, G. (2010). *Ganztagsangebote im Projekt „Schule im Stadtteil“ der Stadt Hannover: Eine empirische Bestandsaufnahme sowie Analysen zu ausgewählten Schüler- und Kontextmerkmalen der dritten und vierten Grundschulklassen*. Leibniz Universität Hannover.
- Fauth, B., Altay, C., Dumont, H. & Decristan, J. (2021). Does What you get Depend on who you are With? Effects of Student Composition on Teaching Quality. *Learning and Instruction*, 71(101355). <https://doi.org/10.1016/j.learninstruc.2020.101355>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive: Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28(3). <https://doi.org/10.1026/2191-9186/a000145>
- Fend, H. (1998). *Qualität im Bildungswesen. Schulforschung zu Systembildungen, Schulprofilen und Lehrerleistungen*. Juventa.
- Field, A. (2009). *Discovering Statistics Using SPSS*. Sage.

- Fixsen, D. L. & Blase, K. A. (2009). Implementation: The Missing Link Between Research and Practice. *NIRN Implementation Brief, 1*, 1–2.
- Fogarty, M., Oslund, E., Simmons, D., Davis, J., Simmons, L., Anderson, L., Clemens, N. & Robers, G. (2014). Examining the Effectiveness of a Multicomponent Reading Comprehension Intervention in Middle Schools: A Focus on Treatment Fidelity. *Educational Psychology Review, 26*(3), 425–449. <https://doi.org/10.1007/s10648-014-9270-6>
- Franceschini, R. (2015). Sollen wir noch am Begriff Mehrsprachigkeit feshalten? In R. Schmidlin, H. Behrens & H. Bickel (Hrsg.), *Sprachgebrauch und Sprachbewusstsein: Implikationen für die Sprachtheorie*. De Gruyter.
- Franken, R. & Franken, S. (2020). *Wissen, Lernen und Innovation im digitalen Unternehmen*. Springer Gabler.
- Fullan, M. (2007). *The NEW Meaning of Educational Change*. Routledge.
- Fung, I. Y., Wilkinson, I. A. & Moore, D. W. (2003). LI-Assisted Reciprocal Teaching to Improve ESL Students' Comprehension of English Expository Text. *Learning and Instruction, 13*(1), 1–31. [https://doi.org/10.1016/S0959-4752\(01\)00033-0](https://doi.org/10.1016/S0959-4752(01)00033-0)
- Fürstenau, S. (2011). Mehrsprachigkeit als Voraussetzung und Ziel schulischer Bildung. In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Mehrsprachigkeit* (S. 25–50). Springer.
- Fürstenau, S. (2016). Multilingualism and School Development in Transnational Educational Spaces. Insights from an Intervention Study at German Elementary Schools. In A. Küppers, B. Pusch & P. U. Semerci (Hrsg.), *Bildung in transnationalen Räumen: Education in Transnational Spaces* (S. 71–90). Springer Fachmedien. https://doi.org/10.1007/978-3-658-09642-7_4
- Fürstenau, S. (2017). Migrationsbedingte Mehrsprachigkeit als Gegenstand der Grundschulforschung. *Zeitschrift für Grundschulforschung, 10*(2), 9–22.
- Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P. R., Shavelson, R. J. & Yin, Y. (2008). On the Fidelity of Implementing Embedded Formative Assessments and its Relation to Student Learning. *Applied Measurement in Education, 21*(4), 360–389. <https://doi.org/10.1080/08957340802347852>
- Gantefort, C. (2020). Nutzung von Mehrsprachigkeit in jedem Unterricht: Das Beispiel „Translanguaging“. In I. Gogolin, A. Hansen, S. McMonagle & D. Rauch (Hrsg.), *Handbuch Mehrsprachigkeit und Bildung* (S. 201–206). Springer Fachmedien. https://doi.org/10.1007/978-3-658-20285-9_29
- García, O. (2009). *Bilingual Education in the 21st Century: A Global Perspective*. Wiley-Blackwell.
- García, O. & Lin, A. M. (2017). Translanguaging in Bilingual Education. In O. García, A. M. Lin & S. May (Eds.), *Encyclopedia of Language and Education. Bilingual and Multilingual Education* (S. 117–130). Springer. https://doi.org/10.1007/978-3-319-02258-1_9
- García, O. & Sylvan, C. E. (2011). Pedagogies and Practices in Multilingual Classrooms: Singularities in Pluralities. *The Modern Language Journal, 95*(3), 385–400. <https://doi.org/10.1111/j.1540-4781.2011.01208.x>
- García, O. & Wei, L. (2014). *Translanguaging: Language, Bilingualism and Education*. Palgrave Macmillan. <https://doi.org/10.1057/9781137385765>
- Geiser, C. (2011). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung* (2. Aufl.). VS Verlag für Sozialwissenschaften.
- Gillwald, K. (2000). *Konzepte sozialer Innovation: WZB Discussion Paper, Np P 00-519*. Wissenschaftszentrum Berlin für Sozialforschung (WZB).

- Ginsburg-Block, M. D., Rohrbeck, C. A. & Fantuzzo, J. W. (2006). A Meta-Analytic Review of Social, Self-Concept, and Behavioral Outcomes of Peer-Assisted Learning. *Journal of Educational Psychology*, 98(4), 732–749. <https://doi.org/10.1037/0022-0663.98.4.732>
- Gogolin, I. (1994). *Der monolinguale Habitus der multilingualen Schule*. Waxmann.
- Gogolin, I., Hansen, A., McMonagle, S. & Rauch, D. (2020). Mehrsprachigkeit und Bildung – zur Konzeption des Handbuchs. In I. Gogolin, A. Hansen, S. McMonagle & D. Rauch (Hrsg.), *Handbuch Mehrsprachigkeit und Bildung* (S. 1–12). Springer Fachmedien.
- Gogolin, I. & Lange, I. (2011). Bildungssprache und Durchgängige Sprachbildung. In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Mehrsprachigkeit* (S. 107–127). Springer. https://doi.org/10.1007/978-3-531-92659-9_6
- Gogolin, I. & Neumann, U. (1991). Sprachliches Handeln in der Grundschule. *Die Grundschulzeitschrift*, 43(5), 6–13.
- Goldenbaum, A. (2012). *Innovationsmanagement in Schulen: Eine empirische Untersuchung zur Implementation eines Sozialen Lernprogramms*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-19425-7>
- Götz, K., Hendel, R., Kutzelmann, S. & Massler, U. (2017). Mehrsprachiges Lesetheater: Ergebnisse zur Akzeptanz dieser neuen mehrsprachigen Lesefördermethode. In S. Kutzelmann, U. Massler, K. Peter, K. Götz & A. Ilg (Hrsg.), *Mehrsprachiges Lesetheater: Handbuch zu Theorie und Praxis* (S. 42–56). Verlag Barbara Budrich.
- Goy, M., Valtin, R. & Hußmann, A. (2017). Leseselbstkonzept, Lesemotivation, Leseverhalten und Lesekompetenz. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 143–175). Waxmann. <https://doi.org/10.25656/01:20755>
- Gräsel, C. (2010). Stichwort: Transfer und Transferforschung im Bildungsbereich. *Zeitschrift für Erziehungswissenschaft*, 13(1), 7–20. <https://doi.org/10.1007/s11618-010-0109-8>
- Gräsel, C. & Parchmann, I. (2004). Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 32(3), 196–214. <https://doi.org/10.25656/01:5813>
- Gresham, F. M. (2009). Evolution of the Treatment Integrity Concept: Current Status and Future Directions. *School Psychology Review*, 38(4), 533–540.
- Gröschner, A. (2011). *Innovation als Lernaufgabe: Eine quantitativ-qualitative Studie zur Erfassung und Umsetzung von Innovationskompetenz in der Lehrerbildung*. Waxmann.
- Grosjean, F. (2013). Bilingualism: A Short Introduction. In F. Grosjean & P. Li (Eds.), *The Psycholinguistics of Bilingualism* (1. Aufl., S. 5–26). Wiley-Blackwell.
- Grosjean, F. (2020). Individuelle Zwei- und Mehrsprachigkeit. In I. Gogolin, A. Hansen, S. McMonagle & D. Rauch (Hrsg.), *Handbuch Mehrsprachigkeit und Bildung* (S. 13–22). Springer Fachmedien.
- Guo, Y., Dynia, J. M., Logan, J. A., Justice, L. M., Breit-Smith, A. & Kaderavek, J. N. (2016). Fidelity of Implementation for an Early-Literacy Intervention: Dimensionality and Contribution to Children's Intervention Outcomes. *Early Childhood Research Quarterly*, 37(4), 165–174. <https://doi.org/10.1016/j.ecresq.2016.06.001>
- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Toboada, A., Davis, M. H., Scalfiddi, N. T. & Tonks, S. (2004). Increasing Reading Comprehension and Engagement Through Concept-Oriented Reading Instruction. *Journal of Educational Psychology*, 96(3), 403–423. <https://doi.org/10.1037/0022-0663.96.3.403>

- Hacker, D. J. & Tenent, A. (2002). Implementing Reciprocal Teaching in the Classroom: Overcoming Obstacles and Making Modifications. *Journal of Educational Psychology*, 94(4), 699–718. <https://doi.org/10.1037//0022-0663.94.4.699>
- Haddock, G. & Maio, G. R. (2014). Einstellungen. In K. Jonas, W. Stroebe & M. Hewstone (Hrsg.), *Sozialpsychologie* (6. Aufl., S. 197–230). Springer.
- Haddock, G. & Maio, G. R. (2019). Chapter Two – Inter-Individual Differences in Attitude Content: Cognition, Affect, and Attitudes. *Advances in Experimental Social Psychology*, 59, 53–102. <https://doi.org/10.1016/bs.aesp.2018.10.002>
- Halai, A. (2009). Politics and Practice of Learning Mathematics in Multilingual Classrooms: Lessons from Pakistan. In R. Barwell (Eds.), *Multilingualism in Mathematics Classrooms: Global Perspectives* (S. 47–62). Multilingual Matters.
- Halliday, A. J., Kern, M. L., Garrett, D. K. & Turnbull, D. A. (2019). Understanding Factors Affecting Positive Education in Practice: An Australian Case Study. *Contemporary School Psychology*, 24, 128–145. <https://doi.org/10.1007/s40688-019-00229-0>
- Hartmann, U., Ehm, Jan-Henning, Höltge, Lea & Hasselhorn, M. (2020). Unter welchen Bedingungen gelingt die Implementation eines Modellvorhabens zur Förderung der Schulbereitschaft in Kindertagesstätten? *Zeitschrift für Erziehungswissenschaft*, 23, 61–82. <https://doi.org/10.1007/s11618-019-00921-0>
- Hasselhorn, M. & Gold, A. (2017). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren* (4. aktualisierte Auflage). Kohlhammer.
- Hasselhorn, M., Köller, O., Maaz, K. & Zimmer, K. (2014). Implementation wirksamer Handlungskonzepte im Bildungsbereich als Forschungsaufgabe. *Psychologische Rundschau*, 65(3), 140–149. <https://doi.org/10.1026/0033-3042/a000216>
- Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression Based Approach* (2nd Edition). The Guilford Press.
- Hebbecke, K., Meudt, S.-I., Schütze, B. & Souvignier, E. (2020). Der Transfer von formativem Assessment in den Leseunterricht der Grundschule: Die vermittelnde Rolle von Akzeptanz, Machbarkeit und Kooperation. *Zeitschrift für Pädagogische Psychologie*, 36(4), 248–262. <https://doi.org/10.1024/1010-0652/a000279>
- Hebbecke, K. & Souvignier, E. (2018). Formatives Assessment im Leseunterricht der Grundschule – Implementation und Wirksamkeit eines modularen, materialgestützten Konzepts. *Zeitschrift für Erziehungswissenschaft*, 21(4), 735–765. <https://doi.org/10.1007/s11618-018-0834-y>
- Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Klett Kallmeyer.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie: Band 3. Psychologie des Unterrichts und der Schule* (S. 71–176). Hogrefe.
- Henninger, K. (2010). *Exploring The Relationship Between Factors of Implementation, Treatment Integrity and Reading Fluency*. Dissertation. University of Massachusetts Amherst.
- Henschel, S., Heppt, B., Weirich, S., Edele, A., Schipolowski, S. & Stanat, P. (2019). Zuwanderungsbezogene Disparitäten. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Hrsg.), *IQB Bildungstrend 2017: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der Sekundarstufe I im zweiten Ländervergleich* (S. 295–336). Waxmann.

- Hesse, H.-G., Göbel, K. & Hartig, J. (2008). Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 208–230). Beltz. <https://doi.org/10.25656/01:3155>
- Hilbe, R. & Kutzelmann, S. (2017). Ohne Akzeptanz keine Innovation? Theoretische Grundlagen für die Erhebung der Akzeptanz des mehrsprachigen Vorlesens. In R. Hilbe, S. Kutzelmann, U. Massler & K. Peter (Hrsg.), *Mehrsprachiges Vorlesen durch die Lehrperson: Handbuch zu Theorie und Praxis* (S. 33–44). Verlag Barbara Budrich.
- Hilbe, R., Kutzelmann, S., Massler, U. & Peter, K. (2017a). Ergebnisse zur Akzeptanz des mehrsprachigen Vorlesens. In R. Hilbe, S. Kutzelmann, U. Massler & K. Peter (Hrsg.), *Mehrsprachiges Vorlesen durch die Lehrperson: Handbuch zu Theorie und Praxis* (S. 147–166). Verlag Barbara Budrich.
- Hilbe, R., Kutzelmann, S., Massler, U. & Peter, K. (Hrsg.). (2017b). *Mehrsprachiges Vorlesen durch die Lehrperson: Handbuch zu Theorie und Praxis*. Verlag Barbara Budrich.
- Hochweber, J. & Hartig, J. (2012). Mehrebenenanalyse. *EEO Enzyklopädie Erziehungswissenschaft Online*, 1–33. <https://doi.org/10.3262/EEO07120217>
- Holtappels, H. G. (2013). Innovation in Schulen – Theorieansätze und Forschungsbefunde. In M. Rürup & I. Bormann (Hrsg.), *Innovationen im Bildungswesen: Analytische Zugänge und empirische Befunde* (S. 45–70). Springer. https://doi.org/10.1007/978-3-531-19701-2_2
- Homburg, C. & Baumgartner, H. (1995). Beurteilung von Kausalmodellen. Bestandsaufnahme und Anwendungsempfehlungen. *Journal of Research and Management*, 17(3), 162–176.
- Homburg, C. & Giering, A. (1996). Konzeptualisierung und Operationalisierung komplexer Konstrukte. Ein Leitfaden für die Marketingforschung. *Journal of Research and Management*, 18(1), 3–24.
- Honneth, A. (1992). *Kampf um Anerkennung: Zur moralischen Grammatik sozialer Konflikte*. Suhrkamp.
- Hu, A. (2003). *Schulischer Fremdsprachenunterricht und migrationsbedingte Mehrsprachigkeit*. Gunter Narr Verlag.
- Hu, L. & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Humphrey, N., Barlow, A. & Lendrum, A. (2018). Quality Matters: Implementation Moderates Student Outcomes in the PATHS Curriculum. *Prevention Science*, 19(2), 197–208. <https://doi.org/10.1007/s11121-017-0802-4>
- Hunneshagen, H. (2005). *Innovationen in Schulen: Identifizierung implementationsfördernder und -hemmender Bedingungen des Einsatzes neuer Medien*. Waxmann.
- IBM Corp. (2021). *IBM SPSS Statistics for Windows* (Version Version 28.0). IBM Corp. Armonk, NY.
- Jakisch, J. (2015). Mehrsprachigkeitsförderung über die 1. Fremdsprache: Der Beitrag des Faches Englisch. *Fremdsprachen Lehren und Lernen*, 44(2), 20–33.
- Jessner, U. & Allgäuer-Hackl, E. (2015). Mehrsprachigkeit aus einer dynamisch-komplexen Sicht oder warum sind Mehrsprachige nicht einsprachig in mehrfacher Ausführung? In E. Allgäuer-Hackl, K. Brogan, U. Henning, B. Hufeisen & J. Schlabach (Hrsg.), *Mehrsprachigkeit und multiples Sprachenlernen: Bd. II. MehrSprachen? – PlurCur! Be-*

- richte aus Forschung und Praxis zu Gesamtsprachencurricula (S. 209–229). Schneider Verlag Hohengehren.
- John, R. (2013). Innovation als soziales Phänomen. In M. Rürup & I. Bormann (Hrsg.), *Innovationen im Bildungswesen: Analytische Zugänge und empirische Befunde* (S. 71–88). Springer.
- Johnson, D. W. & Johnson, R. T. (1998). *Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning* (5th Edition). Allyn and Bacon.
- Johnson, D. W., Johnson, R. T. & Stanne, M. B. (2000). *Cooperative Learning Methods: A Meta-Analysis*. University of Minnesota Press.
- Kaendler, C., Wiedmann, M., Rummel, N. & Spada, H. (2015). Teacher Competencies for the Implementation of Collaborative Learning in the Classroom: A Framework and Research Review. *Educational Psychology Review*, 27, 505–536. <https://doi.org/10.1007/s10648-014-9288-9>
- Kazdin, A. E. (1977). Assessing the Clinical or Applied Importance of Behavior Change Through Social Validation. *Behavior Modification*, 1(4), 427–452.
- Kazdin, A. E. (1980). Acceptability of Alternative Treatments for Deviant Child Behavior. *Journal of Applied Behavior Analysis*, 13(2), 259–273. <https://doi.org/10.1901/jaba.1980.13-259>
- Keim, I. (2012). *Mehrsprachige Lebenswelten: Sprechen und Schreiben der türkischstämmigen Kinder und Jugendlichen*. Narr Francke Attempo. <https://doi.org/10.1515/in-fodaf-2014-2-348>
- Kemp, C. (2009). Defining Multilingualism. In L. Aronin & B. Hufeisen (Eds.), *Exploration of Multilingualism: Development of Research on L3, Multilingualism and Multiple Language Acquisition* (S. 11–26). John Benjamins Publishing Company.
- Kenny, D. A., Kaniskan, B. & McCoach, D. B. (2015). The Performance of RMSEA in Models with Small Degrees of Freedom. *Sociological Methods & Research*, 44(3), 486–507. <https://doi.org/10.1177/0049124114543236>
- Keyserlingk, L. von, Becker, M., Jansen, M. & Maaz, K. (2020). Supplemental Material: Leaving the Pond – Choosing an Ocean: Effects of Student Composition on STEM Major Choices at University. *Journal of Educational Psychology*, 112(4), 751–764. <https://doi.org/10.1037/edu0000378>
- Khan, J. (2018). *Mehrsprachigkeit, Sprachkompetenz und Schulerfolg*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-20413-6>
- Kleidat, C. P. (2011). *Bedingungen der Akzeptanz von Reform*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-93088-6>
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht: Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. Prenzel & L. Allilio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule: Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127–146). Waxmann.
- Klingner, J. K. & Vaughn, S. (2000). The Helping Behaviors of Fifth Graders While Using Collaborative Strategic Reading During ESL Content Classes. *TESOL Quarterly*, 34(1), 69–98. <https://doi.org/10.2307/3588097>
- KMK. (2015). *Empfehlungen zur Arbeit in der Grundschule. Beschluss der Kultusministerkonferenz vom 02.07.1970 i. d. F. vom 11.06.2015*. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2015/Empfehlung_350_KMK_Arbeit_Grundschule_01.pdf

- Koch, H. & Spörer, N. (2016). Förderung der Lesekompetenz mittels reziproken Lehrens: Implementation und Wirksamkeit im Regelunterricht. *Zeitschrift für Pädagogische Psychologie*, 30(4), 213–225. <https://doi.org/10.1024/1010-0652/a000176>
- Kohler, B. & Wacker, A. (2013). Das Angebots-Nutzungs-Modell: Überlegungen zu Chancen und Grenzen des derzeit prominentesten Wirkmodells der Schul- und Unterrichtsforschung. *DDS – Die Deutsche Schule*, 105(3), 242–258.
- Krapp, A., Geyer, C. & Lewalter, D. (2014). Motivation und Emotion. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie* (6. Aufl., S. 193–222). Beltz.
- Kropp, A. (2017). (Herkunftsbedingte) Mehrsprachigkeit als Ressource? Ressourcenorientierung und -management im schulischen FSU. In T. Ambrosch-Baroua, A. Kropp & J. Müller-Lancé (Hrsg.), *Mehrsprachigkeit und Ökonomie* (S. 107–130). Universitätsbibliothek der Ludwig-Maximilians-Universität. <https://doi.org/10.5282/ubm/epub.40521>
- Krumm, H.-J. (2016). Mehrsprachigkeit als Kapital – eine Einsicht und ihre Konsequenzen für ein bildungsgerechtes Schulwesen. In A. Wegener & I. Dirim (Hrsg.), *Mehrsprachigkeit und Bildungsgerechtigkeit. Erkundungen einer didaktischen Perspektive* (S. 59–68). Verlag Barbara Budrich.
- Kuhn, D. (2018). *Implementationsgrad und Wirksamkeit von Unterrichtsinterventionen zur Förderung der selbstbestimmten Lernmotivation*. Dissertation. Pädagogische Hochschule Ludwigsburg.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Ferdinand Schöningh.
- Kutzlmann, S., Massler, U., Peter, K., Götz, K. & Ilg, A. (Hrsg.). (2017). *Mehrsprachiges Lesetheater: Handbuch zu Theorie und Praxis*. Verlag Barbara Budrich. <https://doi.org/10.2307/j.ctvddznv3>
- Lai, K. (2018). Estimating Standardized SEM Parameters Given Nonnormal Data and Incorrect Model: Methods and Comparison. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 600–620. <https://doi.org/10.1080/10705511.2017.1392248>
- Lauer, A. (2006). *Möglichkeiten und Grenzen von Innovationen im Lehrplan – evaluiert am Beispiel der Jahrgangsstufe 6 im Fach Mathematik*. Dissertation.
- LeBreton, J. & Senter, J. L. (2008). Answer to 20 Questions About Interrater Reliability and Interrater Agreement. *Organizational Research Methods*, 11(4), 815–852. <https://doi.org/10.1177/1094428106296642>
- Leeper, T. J. (2021). *Interpreting Regression Results using Average Marginal Effects with R's margins*. <https://cran.r-project.org/web/packages/margins/vignettes/TechnicalDetails.pdf>
- Lendrum, A. & Humphrey, N. (2012). The Importance of Studying the Implementation of Interventions in School Settings. *Oxford Review of Education*, 38(5), 635–652.
- Lengyel, D. (2017). Stichwort: Mehrsprachigkeitsforschung. *Zeitschrift für Erziehungswissenschaft*, 20, 153–174. <https://doi.org/10.1007/s11618-017-0734-6>
- Lenhard, W. (2019). *Leseverständnis und Lesekompetenz: Grundlagen – Diagnostik – Förderung* (2., aktualisierte Auflage). Kohlhammer.
- Lenhard, W. & Artelt, C. (2009). Komponenten des Leseverständnisses. In W. Lenhard & W. Schneider (Hrsg.), *Diagnostik und Förderung des Leseverständnisses* (S. 1–18). Hogrefe.
- Lenhard, W., Lenhard, A. & Schneider, W. (2017). *ELFE II: Ein Lesverständnistest für Erst- bis Siebtklässler – Version II*. Hogrefe.
- Leonet, O., Cenoz, J. & Gorter, D. (2020). Developing Morphological Awareness Across Languages: Translanguaging Pedagogies in Third Language Acquisition. *Language Awareness*, 29(1), 41–59. <https://doi.org/10.1080/09658416.2019.1688338>

- Leutner, D. (2010). Perspektiven pädagogischer Interventionsforschung. In T. Hascher & B. Schmitz (Hrsg.), *Pädagogische Interventionsforschung. Theoretische Grundlagen und empirisches Handlungswissen* (S. 63–72). Juventa.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–105). Springer-Verlag. https://doi.org/10.1007/978-3-642-41291-2_4
- Lipowsky, F., Böhnert, A., Nemeth, L., Hirstein, A. & Pätzold, F. (2022). Kooperatives Lernen in Unterrichtsvideos sichtbar machen – Ein Seminarkonzept zur Weiterentwicklung der professionellen Wahrnehmung von Lehramtsstudierenden. In J. Klusmeyer & D. Bosse (Hrsg.), *Konzepte reflexiver Praxisstudien in der Lehrer*innenbildung* (S. 115–152). Springer VS. https://doi.org/10.1007/978-3-658-35483-1_5
- Lorenz, E., Krulatz, A. & Torgersen, E. N. (2021). Embracing Linguistic and Cultural Diversity in Multilingual EAL Classrooms: The Impact of Professional Development on Teacher Beliefs and Practice. *Teaching and Teacher Education*, 105, 103428. <https://doi.org/10.1016/j.tate.2021.103428>
- Low, S., van Ryzin, M. J., Brown, E. C., Smith, B. H. & Haggerty, K. P. (2014). Engagement Matters: Lessons From Assessing Classroom Implementation of Steps to Respect: A Bullying Prevention Program over a One-Year Period. *Prevention Science*, 15(2), 165–176. <https://doi.org/10.1007/s11121-012-0359-1>
- Lucas, T., Villegas, A. M. & Freedson-Gonzalez, M. (2008). Linguistically Responsive Teacher Education: Preparing Classroom Teachers to Teach English Language Learners. *Journal of Teacher Education*, 59(4), 361–373. <https://doi.org/10.1177/0022487108322110>
- Lucke, D. (1995). Akzeptanz: Legitimität in der Abstimmungsgesellschaft. Leske & Budrich.
- Lüdtke, O. & Robitzsch, A. (2011). Umgang mit fehlenden Daten in der empirischen Bildungsforschung. *EEO Enzyklopädie Erziehungswissenschaft Online*, 1–42. <https://doi.org/10.3262/EEO07100150>
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2006). Analyse von Lernumwelten: Ansätze zur Bestimmung der Reliabilität und Übereinstimmung von Schülerwahrnehmungen. *Zeitschrift für Pädagogische Psychologie*, 20(1/2), 85–96. <https://doi.org/10.1024/1010-0652.20.1.85>
- Lui, Q., Çolak, F. Z. & Agirdag, O. (2021). Teachers' Beliefs and Practices in Culturally Diverse Schools: An Empirical Study in Southwest China. *Teachers and Teaching*, 27(7), 672–695. <https://doi.org/10.1080/13540602.2021.1987878>
- MacKinnon, D. P., Fairchild, A. J. & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology*, 58, 593–614. <https://doi.org/10.1146/annurev.psych.58.110405.085542>
- MacKinnon, D. P., Lockwood, C. M. & Williams, J. (2004). Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods. *Multivariate Behavioral Research*, 39(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4
- Mautone, J. A., DuPaul, G. J., Jitendra, A. K., Tresco, K. E., Junod, R. V. & Volpe, R. J. (2009). The Relationship Between Treatment Integrity and Acceptability of Reading Interventions for Children with Attention-Deficit/Hyperactivity Disorder. *Psychology in the Schools*, 46(10), 919–931. <https://doi.org/10.1002/pits.20434>
- McElvany, N., Becker, M. & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(3), 121–131. <https://doi.org/10.1026/0049-8637.41.3.121>

- McElvany, N., Kortenbruck, M. & Becker, M. (2008). Lesekompetenz und Lesemotivation: Entwicklung und Mediation des Zusammenhangs durch Leseverhalten. *Zeitschrift für Pädagogische Psychologie*, 22(3-4), 207–219. <https://doi.org/10.1024/1010-0652.22.34.207>
- McElvany, N., Ohle, A., El-Khechen, W., Hardy, I. & Cinar, M. (2017). Förderung sprachlicher Kompetenzen – Das Potenzial der Familiensprache für den Wortschatzerwerb aus Texten. *Zeitschrift für Pädagogische Psychologie*, 31(1), 13–25. <https://doi.org/10.1024/1010-0652/a000189>
- Mecheril, P. & Quehl, T. (2006). Sprache und Macht: Theoretische Facetten eines (migrations-)pädagogischen Zusammenhangs. In P. Mecheril & T. Quehl (Hrsg.), *Die Macht der Sprachen* (S. 355–382). Waxmann.
- Mehlhorn, G. (2020). Herkunftssprachen und ihre Sprecher/innen. In I. Gogolin, A. Hansen, S. McMonagle & D. Rauch (Hrsg.), *Handbuch Mehrsprachigkeit und Bildung* (S. 23–30). Springer Fachmedien.
- Meudt, S.-I., Zeuch, N., Neuber, L. & Souvignier, E. (2020). Kurzskalen zur Erfassung des Transfererfolgs von Leseförderkonzepten in der Schulpraxis. In K. Mackowiak, C. Beckerle, S. Gentrup & C. Titz (Hrsg.), *Forschungsinstrumente im Kontext institutioneller (schrift-)sprachlicher Bildung* (S. 141–159). Verlag Julius Klinkhardt. https://doi.org/10.35468/5801_08
- Meuleman, B. & Billiet, J. (2009). A Monte Carlo Sample Size Study: How Many Countries are Needed for Accurate Multilevel SEM? *Survey Research Methods*, 3(1), 45–58. <https://doi.org/10.18148/srm/2009.v3i1.666>
- Meyer, M. & Prediger, S. (2011). The Use of First Language Turkish as a Resource. A German Case Study on Chances and Limits for Building Conceptual Understanding. In M. Setati, T. Nkambule & L. Goosen (Eds.), *Proceedings of the ICMI Study 21 Mathematics and language diversity* (S. 225–234). University Press.
- Milfont, T. L. & Fischer, R. (2010). Testing Measurement Invariance Across Groups: Applications in Cross-Cultural Research. *International Journal of Psychological Research*, 3(1), 111–121. <https://doi.org/10.21500/20112084.857>
- Moosbrugger, H. & Kelava, A. (Hrsg.). (2020). *Testtheorie und Fragebogenkonstruktion* (3. Aufl.). Springer. <https://doi.org/10.1007/978-3-662-61532-4>
- Moschkovic, J. (2007). Using Two Languages when Learning Mathematics. *Educational Studies in Mathematics*, 64, 121–144. <https://doi.org/10.1007/s10649-005-9005-1>
- Müller, B. (2015). *Förderung des Leseverständnisses durch Lesestrategien: Eine Interventionsstudie in der Grundschule*. Dissertation.
- Müller-Böling, D. & Müller, M. (1986). *Akzeptanzfaktoren der Bürokommunikation. Fachberichte und Referate/Lectures and Tutorials: Bd. 17*. Oldenbourg.
- Muñoz-Sandoval, A. F., Cummins, J., Alvarado, C. G. & Rued, M. L. (1998). *Bilingual Verbal Ability Test: Comprehensive Manual*. Riverside Publishing.
- Munser-Kiefer, M. (2014). *Leseförderung im Leseteam in der Grundschule: eine Interventionsstudie zur Förderung von basaler Lesefertigkeit und (meta-)kognitiven Lesestrategien*. Waxmann.
- Murrah, W. M., Kosovich, J. & Hulleman, C. S. (2016). A Framework for Incorporating Intervention Fidelity in Educational Evaluation Studies. In G. Roberts, S. Vaughn, S. N. Berevts & V. Wong (Eds.), *Treatment Fidelity in Studies of Educational Intervention* (S. 39–60). Routledge.
- Muthén, B. O., Muthén, L. K. & Asparouhov, T. (2016). *Regression and Mediation Analysis Using Mplus*. Muthén & Muthén.

- Naumann, A., Musow, S., Aichele, C., Hochweber, J. & Hartig, J. (2019). Instruktionssensitivität von Tests und Items. *Zeitschrift für Erziehungswissenschaft*, 22(1), 181–202. <https://doi.org/10.1007/s11618-018-0832-0>
- O'Donnell, C. L. (2008). Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K–12 Curriculum Intervention Research. *Review of Educational Research*, 78(1), 33–84.
- Oksaar, E. (2003). *Zweitspracherwerb: Wege zur Mehrsprachigkeit und zur interkulturellen Verständigung*. Kohlhammer.
- Palinscar, A. S. & Brown, A. L. (1984). Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. *Cognition and Instruction*, 1(2), 117–175.
- Pauli, C. & Reusser, K. (2000). Zur Rolle der Lehrperson beim kooperativen Lernen. *Schweizerische Zeitschrift für Bildungswissenschaften*, 22(3421-442).
- Perepletchikova, F. & Kazdin, A. E. (2005). Treatment Integrity and Therapeutic Change: Issues and Research Recommendations. *Clinical Psychology: Science and Practice*, 12(4), 365–383. <https://doi.org/10.1093/clipsy.bpi045>
- Petermann, F. (2014). Implementationsforschung: Grundbegriffe und Konzepte. *Psychologische Rundschau*, 65(3), 122–128. <https://doi.org/10.1026/0033-3042/a000214>
- Pillemer, D. P. (1991). One- Versus Two-Tailed Hypothesis Tests in Contemporary Educational Research. *Educational Researcher*, 20(9), 13–17.
- Planas, N. & Setati, M. (2009). Bilingual Students Using Their Languages in the Learning of Mathematics. *Mathematics Education Research Journal*, 21(3), 36–59. <https://doi.org/10.1007/BF03217552>
- Pohlmann-Rother, S. & Lange, S. D. (2020). Mehrsprachigkeit in der Lehrkräftebildung: Empirische Ergebnisse zu den Überzeugungen von Grundschullehrkräften und Implikationen für die Kooperation von KiTa und Grundschule. In S. Pohlmann-Rother, S. D. Lange & U. Franz (Hrsg.), *Kooperation von KiTa und Grundschule: Band 2: Digitalisierung, Inklusion und Mehrsprachigkeit – Aktuelle Herausforderungen beim Übergang bewältigen* (S. 5–44). Wolters Kluwer.
- Pötschke, M. (2019). Mehrebenenanalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 1493–1509). Springer VS. https://doi.org/10.1007/978-3-658-21308-4_111
- Preacher, K. J. & Hayes, A. F. (2004). SPSS and SAS Procedures for Estimating Indirect Effects in Simple Mediation Models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717–731. <https://doi.org/10.3758/BF03206553>
- Prediger, S. & Neugebauer, P. (2023). Can Students with Different Language Backgrounds Profit Equally From a Language-Responsive Instructional Approach for Percentages? Differential Effectiveness in a Field Trial. *Mathematical Thinking and Learning*, 25(1), 2–22. <https://doi.org/10.1080/10986065.2021.1919817>
- Prediger, S., Uribe, Á., Wagner, J., Krause, A. & Redder, A. (2021). Sieben Sprachen im Sachfachunterricht – Ansätze zum Einbezug nicht-geteilter mehrsprachiger Ressourcen zum Aufbau von Konzeptverständnis. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 26(2), 165–194.
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R. & Hensley, M. (2011). Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. *Administration and policy in mental health*, 38, 65–76. <https://doi.org/10.1007/s10488-010-0319-7>

- Pulinx, R., van Avermaet, P. & Agirdag, O. (2017). Silencing Linguistic Diversity: The Extent, the Determinants and Consequences of the Monolingual Beliefs of Flemish Teachers. *International Journal of Bilingual Education and Bilingualism*, 20(5), 542–556. <https://doi.org/10.1080/13670050.2015.1102860>
- Quiring, O. (2006). Methodische Aspekte der Akzeptanzforschung bei interaktiven Medientechnologien. *Münchener Beiträge zur Kommunikationswissenschaft*(6), 1–29. <https://doi.org/10.5282/ubm/epub.1348>
- Rauch, D., Jurecka, A. & Hesse, H.-G. (2010). Für den Drittspracherwerb zählt auch die Lesekompetenz in der Herkunftssprache: Untersuchung der Türkisch-, Deutsch- und Englisch-Lesekompetenz bei Deutsch-Türkisch bilingualen Schülern. In C. Allemann-Ghionda, P. Stanat, K. Göbel & C. Röhner (Hrsg.), *Zeitschrift für Pädagogik: Beiheft 55. Migration, Identität, Sprache und Bildungserfolg* (S. 78–100). Beltz. <https://doi.org/10.25656/01:6946>
- Reich, H. H. (2009). Entwicklungswege türkisch-deutscher Zweisprachigkeit. In U. Neumann & H. H. Reich (Hrsg.), *Erwerb des Türkischen in einsprachigen und mehrsprachigen Situationen* (S. 63–90). Waxmann.
- Reich, H. H. & Krumm, H.-J. (2013). *Sprachbildung und Mehrsprachigkeit: Ein Curriculum zur Wahrnehmung und Bewältigung sprachlicher Vielfalt im Unterricht*. Waxmann. <https://elibrary.utb.de/doi/book/10.31244/9783830979241>
- Reinecker, H. (2009). Therapieforschung. In Margraf, J., Schneider, S. (Hrsg.), *Lehrbuch der Verhaltenstherapie: Grundlagen, Diagnostik, Verfahren, Rahmenbedingungen* (S. 84–99). Springer.
- Reinmann-Rothmeier, G. (2003). *Didaktische Innovation durch Blended Learning: Leitlinien anhand eines Beispiels aus der Hochschule*. Verlag Hans Huber.
- Reitenbach, V., Decristan, J., Rauch, D., Bertram, V. & Schneider, K. M. (2023). Selbsterberichtete Gründe für die (Nicht)Nutzung von Familiensprachen beim mehrsprachigkeitssensiblen Reziproken Lehren. *Unterrichtswissenschaft*, 51, 221–243. <https://doi.org/10.1007/s42010-023-00171-4>
- Reljić, G., Ferring, D. & Martin, R. (2015). A Meta-Analysis on the Effectiveness of Bilingual Programs in Europe. *Review of Educational Research*, 8(1), 92–128. <https://doi.org/10.3102/0034654314548514>
- Rhemtulla, M., Brosseau-Liard, P. É. & Savalei, V. (2012). When Can Categorical Variables Be Treated as Continuous? A Comparison of Robust Continuous and Categorical SEM Estimation Methods Under Suboptimal Conditions. *Psychological Methods*, 17(3), 354–373. <https://doi.org/10.1037/a0029315>
- Riehl, C. (2014). *Mehrsprachigkeit: Eine Einführung*. WBG.
- Ritterfeld, U., Lüdke, C. & Dürkoop, A.-L. (2014). Differentieller Umgang mit Mehrsprachigkeit: Eine Untersuchung in der Sekundarstufe I. *German as a Foreign Language*(1), 45–67.
- Rogers, E. M. (2003). *Diffusion of Innovations*. Free Press.
- Rojas-Andrade, R. & Bahamondes, L. L. (2019). Is Implementation Fidelity Important? A Systematic Review on School-Based Mental Health Programs. *Contemporary Educational Psychology*, 23(4), 339–350. <https://doi.org/10.1007/s40688-018-0175-0>
- Rosebrock, C. (2018). Leseförderung im Kontext der Mehrsprachigkeit aus der Perspektive der Deutschdidaktik. In S. Kutzelnann & U. Massler (Hrsg.), *Mehrsprachige Leseförderung: Grundlagen und Konzepte* (S. 15–28). Narr Francke Attempo.

- Rosenberg, M. J. & Hovland, C. I. (1960). Cognitive, Affective and Behavioral Components of Attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude Organization and Change. An Analysis of Consistency among Attitude Components* (Bd. 3, S. 1–14). Yale University Press.
- Rosenshine, B. & Meister, C. (1994). Reciprocal Teaching: A Review of the Research. *Review of Educational Research*, 64(4), 479–530. <https://doi.org/10.3102/00346543064004479>
- Roth, H.-J., Uçan, Y., Sieger, S. & Gollan, C. (2021). Stichwort: Implementationsforschung zwischen Intervention und Transfer im Kontext von Mehrsprachigkeit und sprachlicher Bildung. *Zeitschrift für Erziehungswissenschaft*, 24(4), 775–818. <https://doi.org/10.1007/s11618-021-01024-5>
- Sanetti, L. M. & Kratochwill, T. R. (2009). Toward Developing a Science of Treatment Integrity: Introduction to the Special Series. *School Psychology Review*, 38(4), 445–459.
- Sauerwein, M. (2017). *Qualität in Bildungssettings der Ganztagschule: Über Unterrichtsforschung und Sozialpädagogik*. Beltz Juventa.
- Schastak, M. (2020). *Bilinguale Interaktion beim Peer-Learning: Eine Mixed-Methods Studie mit bilingual türkisch-deutschsprachig aufwachsenden Schülerinnen*. Verlag Barbara Budrich. <https://doi.org/10.3224/84742378>
- Schastak, M., Reitenbach, V., Rauch, D. & Decristan, J. (2017). Türkisch-deutsch bilinguale Interaktion beim Peer-Learning in der Grundschule: Selbstberichtete Gründe für die Annahme oder Ablehnung bilingualer Interaktionsangebote. *Zeitschrift für Erziehungswissenschaft*, 20(2), 213–235. <https://doi.org/10.1007/s11618-017-0735-5>
- Scheele, A. F., Lesemann, P. P. & Mayo, A. Y. (2010). The Home Language Environment of Monolingual and Bilingual Children and their Language Proficiency. *Applied Psycholinguistics*, 31(1), 117–140. <https://doi.org/10.1017/S0142716409990191>
- Schiefele, U. (1996). *Motivation und Lernen mit Texten*. Hogrefe.
- Schmieder, L. (2019). *Dynamiken der Einstellungsentwicklung: Konzeptuelle und empirische Untersuchungen am Beispiel zweier schulischer Innovationen*. Universitätsverlag Hildesheim. <https://doi.org/10.18442/030>
- Schnitzer, K. (2020). *Mehrsprachigkeit als Ressource: Zur Praxis des Sprachunterrichts in der Sekundarstufe I*. Waxmann.
- Schrader, J. & Hasselhorn, M. (2020). Implementationsforschung im Bildungsbereich. *Zeitschrift für Erziehungswissenschaft*, 23(1), 1–8. <https://doi.org/10.1007/s11618-020-00929-x>
- Schrader, J., Hasselhorn, M., Hetfleisch, P. & Goeze, A. (2020). Stichwortbeitrag Implementationsforschung: Wie Wissenschaft zu Verbesserungen im Bildungssystem beitragen kann. *Zeitschrift für Erziehungswissenschaft*, 23(1), 9–59. <https://doi.org/10.1007/s11618-020-00927-z>
- Schüler-Meyer, A., Prediger, S., Kuzu, T., Wessel, L. & Redder, A. (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, 317–339. <https://doi.org/10.1007/s10763-017-9857-8>
- Schulze, S., Lüke, T., Schröter, A., Krause, K. & Kuhl, J. (2019). Replikationsstudie zur Testgüte der Professionsunabhängigen Einstellungsskala zum Inklusiven Schulsystem – ein Beitrag zur Validierung. *Unterrichtswissenschaft*, 47(2), 201–219. <https://doi.org/10.1007/s42010-018-00034-3>

- Schwartz, I. S. (1996). Expanding the Zone: Thoughts About Social Validity and Training. *Journal of Early Intervention, 20*(3). <https://doi.org/10.1177/105381519602000302>
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), *Zeitschrift für Pädagogik: Beiheft 44. Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (S. 28–53). Beltz. <https://doi.org/10.25656/01:3930>
- Seegerer, R., Marx, A., Stanat, P., Schneider, W., Roick, T. & Marx, P. (2013). Determinanten der Lesekompetenz bei Jugendlichen nicht deutscher Herkunftssprache: Zur Bedeutung der Spracherwerbsreihenfolge im Falle von Mehrsprachigkeit. In N. Jude & E. Klieme (Hrsg.), *Zeitschrift für Pädagogik: 59. Beiheft. PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (S. 111–131). Beltz Juventa. <https://doi.org/10.25656/01:7832>
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik, 60*(6), 850–866.
- Shrout, P. E. & Bolger, N. (2002). Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations. *Psychological Methods, 7*(4), 422–445.
- Sierens, S. & van Avermaet, P. (2014). Language Diversity in Education: Evolving from Multilingual Education to Functional Multilingual Learning. In D. Little, C. Leung & P. van Avermaet (Eds.), *New Perspectives on Language and Education: Vol. 33. Managing Diversity in Education: Languages, Policies, Pedagogies* (S. 204–222). Multilingual Matters.
- Silva, M. R., Collier-Meek, M. A., Coddig, R. S. & DeFouw, E. R. (2019). Acceptability Assessment of School Psychology Interventions From 2005 to 2017. *Psychology in the Schools, 1*–16. <https://doi.org/10.1002/pits.22306>
- Simon, B. (2001). *Wissensmedien im Bildungssektor: Eine Akzeptanzuntersuchung an Hochschulen*. Dissertation.
- Slavin, R. E., Hurley, E. A. & Chamberlain, A. (2003). Cooperative Learning and Achievement: Theory and Research. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of Psychology: Volume 7. Educational Psychology* (S. 177–198). Wiley.
- Slembrouck, S., van Avermaet, P. & van Gorp, K. (2018). Strategies of Multilingualism in Education for Minority Children. In P. van Avermaet, S. Slembrouck, K. van Gorp, S. Sierens & K. Maryns (Eds.), *The Multilingual Edge of Education* (S. 9–39). Palgrave Macmillan.
- Snijders, T. A. B. & Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd Edition). Sage.
- Souvignier, E. (2014). Lesestrategien zur Förderung des Leseverständnisses und der Ansatz einer diagnosbasierten, individualisierten Leseförderung. In R. Valtin & I. Tarelli (Hrsg.), *Lesekompetenz nachhaltig stärken: Evidenzbasierte Maßnahmen und Programme* (S. 167–182). Deutsche Gesellschaft für Lesen und Schreiben.
- Souvignier, E. (2020). Interventionsforschung im Kontext Schule. In T. Hascher, T.-S. Idel & W. Helsper (Hrsg.), *Handbuch Schulforschung* (S. 1–17). Springer VS. https://doi.org/10.1007/978-3-658-24734-8_9-1
- Souvignier, E. & Philipp, M. (2016). Implementation – Begrifflichkeiten, Befunde und Herausforderungen. In M. Philipp & E. Souvignier (Hrsg.), *Implementation von Lesefördermaßnahmen: Perspektiven auf Gelingensbedingungen und Hindernisse*. Waxmann.
- Souvignier, E., Zeuch, N., Jost, J., Karstens, F., Meudt, S.-I. & Schmitz, A. (2021). Evaluation der Implementation konzeptuell unterschiedlicher Maßnahmen zur Leseförderung

- in der Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 24, 883–908. <https://doi.org/10.1007/s11618-021-01033-4>
- Spinath, B. (2011). Lernmotivation. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung: Gegenstandsbereiche* (S. 45–56). VS Verlag für Sozialwissenschaften.
- Spörer, N., Brunstein, J. C. & Kieschke, U. (2009). Improving Students' Reading Comprehension Skills: Effects of Strategy Instruction and Reciprocal Teaching. *Learning and Instruction*, 19(3), 272–282. <https://doi.org/10.1016/j.learninstruc.2008.05.003>
- Spörer, N., Koch, H., Schünemann, N. & Völlinger, V. (2016). *Das Lesetraining mit Käpt'n Carlo für 4. und 5. Klassen: Ein Lehrermanual mit Unterrichtsmaterialien zur Förderung des verstehenden und motivierten Lesens*. Hogrefe.
- Stanat, P. & Schneider, W. (2004). Schwache Leser unter 15-jährigen Schülerinnen und Schülern in Deutschland: Beschreibung einer Risikogruppe. In U. Schiefele, C. Artelt & W. Schneider (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz: Vertiefende Analysen im Rahmen von PISA 2000* (S. 243–274). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-81031-1_10
- Stevens, R. J. & Slavin, R. E. (1995). Effects of a Cooperative Learning Approach in Reading and Writing on Academically Handicapped and Nonhandicapped Students. *The Elementary School Journal*, 95(3), 241–262. <https://doi.org/10.1086/461801>
- Tanol, G. (2010). *Treatment Fidelity: Relation to Treatment Acceptability and Change Over Time*. Dissertation. University of Minnesota.
- Ticheloven, A., Blom, E., Leseman, P. & McMonagle, S. (2021). Translanguaging Challenges in Multilingual Classrooms: Scholar, Teacher and Student Perspectives. *International Journal of Multilingualism*, 18(3), 491–514. <https://doi.org/10.1080/14790718.2019.1686002>
- Ullrich, C. G. (2008). *Die Akzeptanz des Wohlfahrtsstaates: Präferenzen, Konflikte, Deutungsmuster*. VS Verlag für Sozialwissenschaften.
- Unrau, N. J., Rueda, R., Son, E., Polanin, J. R., Lundeen, R. J. & Muraszewski, A. K. (2018). Can Reading Self-Efficacy be Modified? A Meta-Analysis of the Impact of Interventions on Reading Self-Efficacy. *Review of Educational Research*, 88(2), 167–204. <https://doi.org/10.3102/0034654317743199>
- van Ryzin, M. J. & Roseth, C. J. (2018). Cooperative Learning in Middle School: A Means to Improve Peer Relations and Reduce Victimization, Bullying, and Related Outcomes. *Journal of Educational Psychology*, 110(8), 1192–1201. <https://doi.org/10.1037/edu0000265>
- Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- Vertovec, S. (2007). Super-Diversity and its Implications. *Ethnic and Racial Studies*, 30(6), 1024–1054. <https://doi.org/10.1080/01419870701599465>
- Vieluf, S., Kunter, M. & van de Vijver, F. J. (2013). Teacher Self-Efficacy in Cross-National Perspective. *Teaching and Teacher Education*, 35, 92–103. <https://doi.org/10.1016/j.tate.2013.05.006>
- Vieluf, S., Praetorius, A.-K., Rakoczy, K., Kleinknecht, M. & Pietsch, M. (2020). Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts: Ein kritischer Vergleich verschiedener Modellvarianten. In A.-K. Praetorius, J. Grünkern & E. Klieme (Hrsg.), *Zeitschrift für Pädagogik: Beiheft 66. Empirische Forschung zu Unterrichtsqualität. Theore-*

- tische Grundfragen und quantitative Modellierungen (63-80). Beltz Juventa. <https://doi.org/10.3262/ZPB2001063>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Wamhoff, D., Maahs, I.-M. & Dewitz, N. von (2022). „Wichtiger find’ ich jetzt nicht, aber manche sind halt einfach nützlicher“ – Sprachliche Hierarchisierungen aus Sicht mehrsprachig aufwachsender Schüler*innen. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 27(1), 345–368.
- Webb, N. M. (1989). Peer Interaction and Learning in Small Groups. *International Journal of Educational Research*, 13(1), 21–39. [https://doi.org/10.1016/0883-0355\(89\)90014-1](https://doi.org/10.1016/0883-0355(89)90014-1)
- Weiber, R. & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung: Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS* (2. Aufl.). Springer-Verlag. <https://doi.org/10.1007/978-3-642-35012-2>
- Weiß, R. H. (2006). *CFT 20-R: Grundintelligenztest Skala 2 – Manual*. Hogrefe.
- Wendt, H. & Schwippert, K. (2017). Lesekompetenzen von Schülerinnen und Schülern mit und ohne Migrationshintergrund. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 219–234). Waxmann.
- Widaman, K. F. & Reise, S. P. (1997). Exploring the Measurement Invariance of Psychological Instruments: Applications in the Substance Use Domain. In K. J. Bryant, M. Windle & S. G. West (Eds.), *The Science of Prevention: Methodological Advances From Alcohol and Substance Abuse Research* (S. 281–324). American Psychological Association.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wildemann, A. & Fornol, S. (2016). *Sprachsensibel unterrichten in der Grundschule: Anregungen für den Deutsch-, Mathematik- und Sachunterricht*. Klett Kallmeyer.
- Witt, J. C. & Elliott, S. N. (1985). Acceptability of Classroom Intervention Strategies. In T. R. Kratochwill (Eds.), *Advances in School Psychology: Volume 4* (S. 251–288). Routledge.
- Witt, J. C. & Martens, B. K. (1983). Assessing the Acceptability of Behavioral Interventions Used in Classrooms. *Psychology in the Schools*, 20(4), 510–517. [https://doi.org/10.1002/1520-6807\(198310\)20:4<510::AID-PITS2310200420>3.0.CO;2-1](https://doi.org/10.1002/1520-6807(198310)20:4<510::AID-PITS2310200420>3.0.CO;2-1)
- Wolf, M. M. (1978). Social Validity: The Case For Subjective Measurement or How Applied Behavior Analysis is Finding its Heart. *Journal of Applied Behavior Analysis*, 11(2), 203–214. <https://doi.org/10.1901/jaba.1978.11-203>
- Zhao, X., Lynch, J. G., JR. & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>
- Ziegelbauer, S. & Gläser-Zikuda, M. (2016). Portfolio als pädagogisch-didaktische Innovation in Schule, LehrerInnenbildung und Hochschuldidaktik. In S. Ziegelbauer & M. Gläser-Zikuda (Hrsg.), *Das Portfolio als Innovation in Schule, Hochschule und LehrerInnenbildung: Perspektiven aus Sicht von Praxis, Forschung und Lehre* (S. 9–12). Julius Klinkhardt.

Abbildungsverzeichnis

Abbildung 1:	<i>Model of Treatment Acceptability</i> nach Witt & Elliott (1985)	53
Abbildung 2:	Bilinguale Promptkarten zu den Lesestrategien „Fragen stellen“ und „Zusammenfassen“ (Beispiel: Deutsch-Türkisch)	74
Abbildung 3:	Messmodell der affektiven und kognitiven Akzeptanz	84
Abbildung 4:	Mediationsmodell nach Baron und Kenny (1986)	102
Abbildung 5:	Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung durch die Nutzung der Promptkarten, mediiert über die Akzeptanzdimensionen und unter Kontrolle des Eingangswerts der Selbstwirksamkeit	116
Abbildung 6:	Mediationsmodell zur schematischen Darstellung der signifikanten Zusammenhänge zwischen Mehrsprachigkeit und Akzeptanz, unter Berücksichtigung der Kontrollvariablen	122
Abbildung 7:	Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung durch die Nutzung der Promptkarten in einer anderen Sprache, mediiert über die Akzeptanz- dimensionen und unter Kontrolle des Eingangswerts der Selbstwirksamkeit	129

Tabellenverzeichnis

Tabelle 1:	Interventions- und Erhebungsdesign der Datengrundlage	77
Tabelle 2:	Prozentualer Implementationsindex der Kernelemente	79
Tabelle 3:	Kreuztabelle zur Nutzung der Promptkarten und Nutzung der Promptkarten in einer anderen Sprache	82
Tabelle 4:	Deskriptive Kennwerte der Items zu den Akzeptanzdimensionen	83
Tabelle 5:	Modellstatistiken zum Vergleich der Faktorenstruktur der Akzeptanz	85
Tabelle 6:	Modellstatistiken zur Überprüfung der Messinvarianz des Zweifaktorenmodells der Akzeptanz zwischen monolingualen und mehrsprachigen Schüler*innen	86
Tabelle 7:	Deskriptive Kennwerte der Items zur Skala „Lesefreude“	88
Tabelle 8:	Deskriptive Kennwerte der Items zur Skala „Selbstwirksamkeit der Lesestrategieanwendung“	89
Tabelle 9:	Deskriptive Kennwerte der Items zur Skala „Einstellungen zu Mehrsprachigkeit“	90
Tabelle 10:	Varianzkomponentenzerlegung der untersuchten Konstrukte	94
Tabelle 11:	Häufigkeitsverteilungen zur Nutzung der Promptkarten	106
Tabelle 12:	Deskriptive Kennwerte der Akzeptanzdimensionen	107
Tabelle 13:	Häufigkeiten der dichotomen individuellen Merkmale	108
Tabelle 14:	Deskriptive Kennwerte der Leistungstests und eingesetzten Skalen zur Erfassung individueller Merkmale	109
Tabelle 15:	Bivariate Korrelationen der Akzeptanzdimensionen und Nutzung der Promptkarten	109
Tabelle 16:	Bivariate Korrelationen der untersuchten individuellen Merkmale und der Facetten des Implementationserfolgs	110
Tabelle 17:	Lineare Regressionsmodelle zur Vorhersage des Leseverstehens am Ende der Intervention	113
Tabelle 18:	Lineare Regressionsmodelle zur Vorhersage der Lesefreude am Ende der Intervention	113
Tabelle 19:	Lineare Regressionsmodelle zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention	114
Tabelle 20:	Lineare Regressionsmodelle zur Vorhersage der Einstellungen zu Mehrsprachigkeit am Ende der Intervention	115
Tabelle 21:	Indirekte Zusammenhänge im Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention	116
Tabelle 22:	Regressionsmodelle zur schrittweisen Vorhersage der Nutzung der Promptkarten durch lerngegenstandsbezogene Voraussetzungen	118
Tabelle 23:	Multivariate Regressionsmodelle zur schrittweisen Vorhersage der Akzeptanzdimensionen durch lerngegenstandsbezogene Voraussetzungen	120
Tabelle 24:	Spezifische indirekte Zusammenhänge lerngegenstandsbezogener Voraussetzungen im Mediationsmodell zur Vorhersage der Akzeptanzdimensionen durch das Merkmal Mehrsprachigkeit	123

Tabelle 25:	Lineare Regressionsmodelle zur Vorhersage des Leseverstehens am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen	126
Tabelle 26:	Lineare Regressionsmodelle zur Vorhersage der Lesefreude am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen	126
Tabelle 27:	Lineare Regressionsmodelle zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen	127
Tabelle 28:	Lineare Regressionsmodelle zur Vorhersage der Einstellungen zu Mehrsprachigkeit am Ende der Intervention in der Substichprobe mehrsprachiger Schüler*innen	128
Tabelle 29:	Regressionsmodelle zur schrittweisen Vorhersage der Nutzung der Promptkarten in einer anderen Sprache durch lerngegenstandsbezogene Voraussetzungen in der Substichprobe mehrsprachiger Schüler*innen	131
Tabelle A1:	Modellstatistiken des eindimensionalen Modells der Lesefreude	175
Tabelle A2:	Modellstatistiken zur Überprüfung der Messinvarianz des eindimensionalen Modells der Lesefreude zwischen monolingualen und mehrsprachigen Schüler*innen	175
Tabelle A3:	Modellstatistiken des eindimensionalen Modells der Selbstwirksamkeit der Lesestrategieanwendung.	176
Tabelle A4:	Modellstatistiken zur Überprüfung der Messinvarianz des eindimensionalen Modells der Selbstwirksamkeit der Lesestrategieanwendung zwischen monolingualen und mehrsprachigen Schüler*innen	176
Tabelle A5:	Modellstatistiken des eindimensionalen Modells der Einstellungen zu Mehrsprachigkeit.	177
Tabelle A6:	Modellstatistiken zur Überprüfung der Messinvarianz des eindimensionalen Modells der Einstellungen zu Mehrsprachigkeit zwischen monolingualen und mehrsprachigen Schüler*innen	177
Tabelle A7:	Korrelationen der individuellen Lernvoraussetzungen in der Gesamtstichprobe.	178
Tabelle A8:	Korrelationen der individuellen Lernvoraussetzungen in der Substichprobe mehrsprachiger Schüler*innen.	178
Tabelle A9:	Mediationsmodell zur vermittelnden Rolle der Akzeptanz in der Gesamtstichprobe (Fragestellung 1)	179
Tabelle A10:	Mediationsmodell zum Zusammenhang zwischen der Mehrsprachigkeit und Akzeptanz, vermittelt durch lerngegenstandsbezogene Voraussetzung in der Gesamtstichprobe (Fragestellung 2)	180
Tabelle A11:	Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung durch die Nutzung der Promptkarten in einer anderen Sprache, mediiert über die Akzeptanzdimensionen in der Substichprobe mehrsprachiger Schüler*innen (Fragestellung 3)	181

Abkürzungen

AME	Average Marginal Effect
CFI	Comparative Fit Index
CIRP	Childrens' Intervention Rating Profile
FIML	Full-Information-Maximum-Likelihood
FT	Fokustisch
ICC	Intraklassenkorrelationskoeffizient
IRP-20	Intervention Rating Profile
KIP	Kids Intervention Profile
meRLe	Förderung der Deutsch-Lesekompetenz durch mehrsprachigkeits-sensibles Reziprokes Lehren im Grundschulunterricht
MGCFA	Mehrgruppen-Konfirmatorische Faktorenanalyse
ML	Maximum-Likelihood
MLR	Robust-Maximum-Likelihood
RMSEA	Root-Mean-Square-Error of Approximation
TG	Treatmentgruppe
VIF	Variance Inflation Factor
W-KG	Warte-Kontrollgruppe
WLSMV	Weighted-Least-Square-Mean-Variance

Anhang

Tabelle A1: Modellstatistiken des eindimensionalen Modells der Lesefreude

Modell	1. Messzeitpunkt (Prä)		2. Messzeitpunkt (Post)	
	Monolingual	Mehrsprachig	Monolingual	Mehrsprachig
χ^2 (<i>df</i>)	6.849 (2)	1.733 (2)	10.126 (2)	4.981 (2)
χ^2/df	3.425	0.867	5.063	2.491
$p(\chi^2)$.033	.420	.006	.083
CFI	.998	1.000	.997	.999
RMSEA	.120	.000	.155	.060
<i>N</i>	169	414	170	419

Tabelle A2: Modellstatistiken zur Überprüfung der Messinvarianz des eindimensionalen Modells der Lesefreunde zwischen monolingualen und mehrsprachigen Schüler*innen

Modell	1. Messzeitpunkt (Prä)			2. Messzeitpunkt (Post)		
	Konfigural	Metrisch	Skalar	Konfigural	Metrisch	Skalar
χ^2 (<i>df</i>)	7.892 (4)	13.155 (7)	22.286 (14)	15.876 (4)	11.368 (7)	14.938 (14)
χ^2/df	1.973	1.879	1.592	3.969	1.624	1.067
$p(\chi^2)$.096	.068	.073	.003	.123	.382
CFI	.999	.999	.999	.998	.999	1.00
RMSEA	.058	.055	.045	.100	.046	.015
<i>N</i>	583			589		
$\Delta\chi^2(df)$		5.773 (3); $p = .123$	10.493 (7); $p = .162$		1.218 (3); $p = .749$	4.923 (7); $p = .669$

Tabelle A3: Modellstatistiken des eindimensionalen Modells der Selbstwirksamkeit der Lese-strategianwendung

Modell	1. Messzeitpunkt (Prä)		2. Messzeitpunkt (Post)	
	Monolingual	Mehrsprachig	Monolingual	Mehrsprachig
$\chi^2 (df)$	40.084 (2)	35.757 (2)	14.372 (2)	31.859 (2)
χ^2/df	20.042	17.879	7.186	15.930
$p(\chi^2)$	< .001	< .001	.001	< .001
CFI	.941	.976	.988	.985
RMSEA	.335	.202	.191	.189
<i>N</i>	170	414	170	419

Tabelle A4: Modellstatistiken zur Überprüfung der Messinvarianz des eindimensionalen Modells der Selbstwirksamkeit der Lesestrategianwendung zwischen monolingualen und mehrsprachigen Schüler*innen

Modell	1. Messzeitpunkt (Prä)			2. Messzeitpunkt (Post)		
	Konfigural	Metrisch	Skalar	Konfigural	Metrisch	Skalar
$\chi^2 (df)$	81.239 (4)	90.604 (7)	116.408 (14)	47.531 (4)	49.148 (7)	66.417 (14)
χ^2/df	20.310	12.943	8.315	11.883	7.021	4.725
$p(\chi^2)$	< .001	< .001	< .001	< .001	< .001	< .001
CFI	.965	.962	.953	.986	.987	.983
RMSEA	.257	.202	.158	.192	.143	.113
<i>N</i>	584			589		
$\Delta\chi^2(df)$		12.505 (3); $p = .006$	20.126 (7) $p = .005$		6.876 (3); $p = .076$	15.252 (7); $p = .033$

Tabelle A5: Modellstatistiken des eindimensionalen Modells der Einstellungen zu Mehrsprachigkeit

Modell	1. Messzeitpunkt (Prä)		2. Messzeitpunkt (Post)	
	Monolingual	Mehrsprachig	Monolingual	Mehrsprachig
χ^2 (df)	11.474 (2)	1.955 (2)	3.530 (2)	2.275 (2)
χ^2/df	5.737	0.978	1.765	1.138
$p(\chi^2)$.003	.376	.171	.321
CFI	.976	1.000	.997	.999
RMSEA	.167	.000	.067	.018
<i>N</i>	170	411	169	418

Tabelle A6: Modellstatistiken zur Überprüfung der Messinvarianz des eindimensionalen Modells der Einstellungen zu Mehrsprachigkeit zwischen monolingualen und mehrsprachigen Schüler*innen

Modell	1. Messzeitpunkt (Prä)			2. Messzeitpunkt (Post)		
	Konfigural	Metrisch	Skalar	Konfigural	Metrisch	Skalar
χ^2 (df)	10.086 (4)	11.655 (7)	18.321 (14)	6.059 (4)	14.821 (7)	33.271 (14)
χ^2/df	2.522	1.665	1.309	1.515	2.117	2.377
$p(\chi^2)$.039	.113	.193	.195	.038	.003
CFI	.993	.995	.995	.998	.993	.982
RMSEA	.072	.048	.033	.042	.062	.068
<i>N</i>	581			587		
$\Delta\chi^2$ (df)		1.762 (3); $p = .623$	7.260 (7) $p = .402$		8.435 (3); $p = .038$	20.661 (7); $p = .004$

Tabelle A7: Korrelationen der individuellen Lernvoraussetzungen in der Gesamtstichprobe.

	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Mehrsprachigkeit	-.01	-.19*	-.19*	-.21*	.001	-.03	.20*
(2) Geschlecht (Jungen)		-.10*	-.01	-.06	-.20*	-.05	-.05
(3) Kognitive Grundfähigkeiten			-.20*	.30*	.10*	.15*	-.11*
(4) Mehr als 25 Bücher im Haushalt				.27*	.25*	.28*	.01
(5) Leseverstehen (Prä)					.27*	.31*	-.04
(6) Lesefreude (Prä)						.44*	.08
(7) Selbstwirksamkeit Lesestrategien (Prä)							.17*
(8) Einstellungen zu Mehrsprachigkeit (Prä)							

Anmerkungen. $N_{\text{Min}} = 554$; $N_{\text{Max}} = 600$; * $p < .05$.

Tabelle A8: Korrelationen der individuellen Lernvoraussetzungen in der Substichprobe mehrsprachiger Schüler*innen.

	(2)	(3)	(4)	(5)	(6)	(7)
(1) Geschlecht (Jungen)	.06	.003	-.07	-.19*	.10	-.04
(2) Kognitive Grundfähigkeiten		-.18*	.21*	.04	.11*	-.09
(3) Mehr als 25 Bücher im Haushalt			.21*	.23*	.28*	.08
(4) Leseverstehen (Prä)				.19*	.28*	.04
(5) Lesefreude (Prä)					.44*	.08
(6) Selbstwirksamkeit Lesestrategien (Prä)						.21*
(7) Einstellungen zu Mehrsprachigkeit (Prä)						

Anmerkungen. $N_{\text{Min}} = 375$; $N_{\text{Max}} = 430$; * $p < .05$.

Tabelle A9: Mediationsmodell zur vermittelnden Rolle der Akzeptanz in der Gesamtstichprobe (Fragestellung 1)

Prädiktoren	Abhängige Variablen								
	Selbstwirksamkeit Lesestrategieanwen- dung (Post)			Affektive Akzeptanz			Kognitive Akzeptanz		
	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>	β	(SE)	<i>p</i>
Selbstwirksamkeit Lesestrategieanwendung (Prä)	.54	(.03)	<.001	.11	(.05)	.027	.04	(.05)	.398
Nutzung der Prompt- karten	-.03	(.04)	.431	.27	(.05)	<.001	.24	(.04)	<.001
Affektive Akzeptanz	.26	(.05)	<.001						
Kognitive Akzeptanz	-.08	(.05)	.140						
R ²	.36			.09			.06		

Anmerkung. *N* = 628.

Tabelle A10: Mediationsmodell zum Zusammenhang zwischen der Mehrsprachigkeit und Akzeptanz, vermittelt durch lerngegenstandsbezogene Voraussetzung in der Gesamtstichprobe (Fragestellung 2)

Prädiktoren	Abhängige Variable																	
	Leseverstehen (Prä)			Lese Freude (Prä)			Selbstwirksamkeit Lesestrategieanwendung (Prä)			Einstellungen zu Mehrsprachigkeit (Prä)			Affektive Akzeptanz			Kognitive Akzeptanz		
	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p
Mehrsprachigkeit	-.13	(.03)	<.001	.05	(.04)	.155	.04	(.04)	.333	.17	(.04)	<.001	.04	(.05)	.366	.09	(.04)	.038
Geschlecht	-.04	(.04)	.340	-.19	(.04)	<.001	-.03	(.04)	.423	-.06	(.04)	.203	-.04	(.04)	.280	-.05	(.04)	.214
Kognitive Grundfähigkeiten	.23	(.05)	<.001	.03	(.04)	.483	.10	(.04)	.014	-.09	(.04)	.022	.01	(.04)	.735	-.07	(.04)	.078
Mehr als 25 Bücher im Haushalt	.20	(.04)	<.001	.25	(.05)	<.001	.27	(.05)	<.001	.06	(.05)	.229	-.08	(.04)	.023	-.06	(.04)	.122
Leseverstehen (Prä)													-.15	(.06)	.008	-.16	(.05)	.003
Lese Freude (Prä)													.27	(.05)	<.001	.14	(.05)	.007
Selbstwirksamkeit Lesestrategieanwendung (Prä)													.03	(.05)	.589	.02	(.05)	.661
Einstellungen zu Mehrsprachigkeit (Prä)													.17	(.05)	.001	.20	(.05)	<.001

Anmerkung. N = 634.

Tabelle A11: Mediationsmodell zur Vorhersage der Selbstwirksamkeit der Lesestrategieanwendung durch die Nutzung der Promptkarten in einer anderen Sprache, mediiert über die Akzeptanzdimensionen in der Substichprobe mehrsprachiger Schüler*innen (Fragestellung 3)

Prädiktoren	Abhängige Variablen								
	Selbstwirksamkeit Lesestrategie- anwendung (Post)			Affektive Akzeptanz			Kognitive Akzeptanz		
	β	(SE)	p	β	(SE)	p	β	(SE)	p
Selbstwirksamkeit Lesestrategieanwendung (Prä)	.52	(.04)	<.001	.10	(.06)	.060	.03	(.05)	.545
Nutzung der Prompt- karten in einer anderen Sprachen	-.02	(.04)	.616	.13	(.07)	.059	.15	(.06)	.016
Affektive Akzeptanz	.23	(.07)	<.001						
Kognitive Akzeptanz	-.01	(.06)	.903						
R ²	.35			.03			.02		

Anmerkung. $N = 425$, Korrelation zwischen den Akzeptanzdimensionen: $r = .77$ ($p < .001$)