

McElvany, Nele [Hrsg.]; König, Sabrina [Hrsg.]; Schaufelberger, Rahim [Hrsg.]; Becker, Michael [Hrsg.]; Gaspard, Hanna [Hrsg.]; Heppt, Birgit [Hrsg.]; Naumann, Alexander [Hrsg.]

Bildungsprozesse und Kompetenzentwicklungen im Kontext sprachlicher und sozialer Heterogenität

Weinheim : Beltz Juventa 2025, 264 S. - (Jahrbuch der Schulentwicklung; 24)



Quellenangabe/ Reference:

McElvany, Nele [Hrsg.]; König, Sabrina [Hrsg.]; Schaufelberger, Rahim [Hrsg.]; Becker, Michael [Hrsg.]; Gaspard, Hanna [Hrsg.]; Heppt, Birgit [Hrsg.]; Naumann, Alexander [Hrsg.]: Bildungsprozesse und Kompetenzentwicklungen im Kontext sprachlicher und sozialer Heterogenität. Weinheim : Beltz Juventa 2025, 264 S. - (Jahrbuch der Schulentwicklung; 24) - URN: urn:nbn:de:0111-pedocs-341678 - DOI: 10.25656/01:34167; 10.3262/978-3-7799-9323-0

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-341678>

<https://doi.org/10.25656/01:34167>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

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-nc-nd/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

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

Nele McElvany | Sabrina König |
Rahim Schaufelberger | Michael Becker |
Hanna Gaspard | Birgit Heppt |
Alexander Naumann (Hrsg.)

Jahrbuch der Schulentwicklung Band 24

Bildungsprozesse und Kompetenzent-
wicklungen im Kontext sprachlicher
und sozialer Heterogenität

Nele McElvany | Sabrina König | Rahim Schaufelberger | Michael
Becker | Hanna Gaspard | Birgit Heppt | Alexander Naumann (Hrsg.)
Jahrbuch der Schulentwicklung. Band 24

Eine Veröffentlichung
des Instituts für Schulentwicklungsforschung
der Technischen Universität Dortmund

Nele McElvany | Sabrina König |
Rahim Schaufelberger | Michael Becker |
Hanna Gaspard | Birgit Heppt |
Alexander Naumann (Hrsg.)

Jahrbuch der Schulentwicklung. Band 24

Bildungsprozesse und
Kompetenzentwicklungen im Kontext
sprachlicher und sozialer Heterogenität

BELTZ JUVENTA

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Der Text dieser Publikation wird unter der Lizenz **Creative Commons Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0 International (CC BY-NC-ND 4.0)** veröffentlicht. Den vollständigen Lizenztext finden Sie unter: <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode.de>. Verwertung, die den Rahmen der **CC BY-NC-ND 4.0 Lizenz** überschreitet, ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für die Bearbeitung und Übersetzungen des Werkes. Die in diesem Werk enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Quellenangabe/Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



Dieses Buch ist erhältlich als:

ISBN 978-3-7799-9322-3 Print

ISBN 978-3-7799-9323-0 E-Book (PDF)

DOI 10.3262/978-3-7799-9323-0

1. Auflage 2025

© 2025 Beltz Juventa

Verlagsgruppe Beltz

Werderstraße 10, 69469 Weinheim

service@beltz.de

Einige Rechte vorbehalten

Satz: Helmut Rohde, Euskirchen

Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza

Beltz Grafische Betriebe ist ein Unternehmen mit finanziellem Klimabeitrag
(ID 15985-2104-1001)

Printed in Germany

Weitere Informationen zu unseren Autor:innen und Titeln finden Sie unter: www.beltz.de

Inhalt

Grußwort der Herausgeber*innen

[7](#)

Teil A:

Bestandsaufnahmen, Analysen und Forschungsbefunde

Reading Competence and Vocabulary of Students From Diverse Language Backgrounds: Employing a Lexical Distance Measure
Sebastian Nicolas Thomas Vogel & Justine Stang-Rabrig

[12](#)

Effectiveness of Professional Development for Language-Supportive Teaching: Insights From Employing Computational Linguistic Analysis Methods

Birgit Heppt, Denise Löfflad, Sofie Henschel, Katrin Gabler, Ilonca Hardy & Detmar Meurers

[35](#)

Enthusiastische Lehrkräfte, motivierte Lernende? Eine differenzielle Betrachtung der Verbindung von Unterrichtsenthusiasmus und intrinsischer Motivation in heterogenen Lerngruppen

Anne Kosubek, Hanna Gaspard, Hannes Schröter & Fani Lauermann

[61](#)

In welchem Maße ist die Gymnasialempfehlung von Lehrkräften heute noch vom familiären Hintergrund der Schüler*innen abhängig?

Ruben Kleinkorres, Ulrich Ludewig & Ramona Lorenz

[89](#)

Obtaining Abitur Without Teacher Recommendation – The Role of Social Background, Opportunity Costs, and Motivation

Katharina Molitor, Paul Fabian & Justine Stang-Rabrig

[128](#)

Addressing the Heterogeneity of Students' Attitudes Towards School: Uncovering School Alienation Profiles in German Secondary Schools

Alyssa Grecu, Sira Neuhaus, Sittipan Yotyodying & Nele McElvany

[158](#)

Kulturre sponsives Schulleitungshandeln als Bedingung erfolgreicher Bildungsprozesse von Schüler*innen aus marginalisierten Herkunftsmilieus – Ein erster Kartierungsversuch für den deutschen Sprachraum

Christine Becks, Susanne Julia Czaja & Esther Dominique Klein

[181](#)

Teil B:

Interventionskonzepte

Digitale Schreibförderung im Ganzttag an Grundschulen: Herausforderungen und Potenziale für Schulentwicklungsprozesse

Kathrin Gade & Vera Busse

[216](#)

Promoting Vocabulary and Grammar in German First-Grade Classrooms: Design and Structure of a Digital Language Intervention

Leonie Dargiewicz, Svenja Wehrhöfer, Annika Ohle-Peters,

Fani Lauermann & Nele McElvany

[235](#)

Autor*innen

[257](#)

Grußwort der Herausgeber*innen

Der 24. Band des Jahrbuchs der Schulentwicklung umfasst in bewährter Tradition der Reihe Bestandsaufnahmen, Analysen und Forschungsbefunde und stellt zudem innovative Interventionskonzepte zu aktuellen und bedeutsamen Themen der schul- und unterrichtsbezogenen Bildungsforschung vor. Thematischer Fokus dieses Jahrbuchs sind *Bildungsprozesse und Kompetenzentwicklungen im Kontext sprachlicher und sozialer Heterogenität*. Vor dem Hintergrund gesunkener Schüler*innenkompetenzen in zentralen Bereichen wie der deutschen Sprache, des Fortbestehens sozialer Disparitäten und der Relevanz von Bildungsgerechtigkeit steht das Schulsystem vor großen Herausforderungen, die es zu adressieren gilt. Zentrale Fragen schul- und unterrichtsbezogener Bildungsforschung sind daher unter anderem, wie Merkmale sprachlicher und sozialer Heterogenität mit Unterschieden in Kompetenzentwicklungen zusammenhängen und wie Bildungsprozesse zur Reduktion heterogenitätsbedingter Bildungsungerechtigkeit optimiert werden können. Zusätzlich werden zwei Interventionskonzepte zur systematischen Förderung im Rahmen einer digitalen Schreib- bzw. Wortschatz- und Grammatikförderung vorgestellt. Die Beiträge des vorliegenden Bandes des Jahrbuchs der Schulentwicklung nähern sich der Thematik aus unterschiedlichen Perspektiven. In dem ersten Bereich liegt der Fokus neben einer umfassenden Bestandsaufnahme auf aktuellen Forschungsergebnissen:

Vogel und Stang-Rabrig fokussieren auf ein Maß lexikalischer Distanz zwischen der Herkunftssprache multilingualer Viertklässler*innen und deren Unterrichtssprache Deutsch. Unter Berücksichtigung des Alters, genereller kognitiver Fähigkeiten und dem sozioökonomischen Status werden Zusammenhänge zwischen der lexikalischen Distanz und der Lesekompetenz sowie zwischen der lexikalischen Distanz und dem Wortschatz regressionsanalytisch untersucht.

Heppt, Löfflad, Henschel, Gabler, Hardy und Meurers erforschen das Potenzial computerlinguistischer Analyseverfahren zur Evaluation der Effekte von Lehrkräftefortbildungen. Es wird untersucht, ob Lehrkräfte, die für die Umsetzung fachintegrierter Sprachbildung im Sachunterricht der Grundschule fortgebildet wurden, ihren Schüler*innen einen variableren und anspruchsvolleren Sprachinput anboten als Lehrkräfte der Kontrollgruppe.

Kosubek, Gaspard, Schröter und Lauermann befassen sich in ihrem Beitrag mit Zusammenhängen zwischen dem Unterrichtsenthusiasmus von Lehrkräften und der intrinsischen Motivation von Lernenden. Die Zusammenhänge werden für sprachlich heterogene Gruppen von Lernenden in Vorbereitungsklassen an Regelschulen der Sekundarstufe I sowie in Integrationskursen für Erwachsene analysiert.

Kleinkorres, Ludewig und Lorenz gehen der Frage nach, inwiefern Gymnasialempfehlungen von Lehrkräften vom familiären Hintergrund von Schüler*innen abhängig sind. Unter Berücksichtigung leistungsbezogener Kriterien für den Übergang am Ende der vierten Klassenstufe wie Zeugnisnoten und Arbeitsverhalten werden in den Analysen der Zuwanderungshintergrund und der sozioökonomische Status fokussiert.

Molitor, Fabian und Stang-Rabrig untersuchen die Bedingungen eines erfolgreichen Gymnasialbesuchs von Schüler*innen mit und ohne Gymnasialempfehlung am Ende der Grundschulzeit. Dabei prüfen sie anhand von Strukturgleichungsmodellen, ob sich diese beiden Gruppen systematisch bezüglich des sozioökonomischen Status, der Opportunitätskosten und der Motivation unterscheiden.

Greco, Neuhaus, Yotyodying und McElvany befassen sich mit Schulentfremdung und nehmen dabei die Heterogenität der Einstellungen von Siebtklässler*innen sowie von Zehntklässler*innen zur Schule in den Blick. Sie stellen latente Schulentfremdungsprofile vor und analysieren, wie diese mit Merkmalen der Schüler*innen wie dem Geschlecht und einem Migrationshintergrund zusammenhängen.

Becks, Czaja und Klein beschäftigen sich in ihrem Beitrag mit kulturresponsivem Handeln als Bedingung erfolgreicher Bildungsprozesse von Schüler*innen aus marginalisierten Herkunftsmilieus. Auf Basis eines systematischen Literaturreviews bieten sie einen Überblick über die Forschung zu kulturresponsivem Schulleitungshandeln und diskutieren Adaptionismöglichkeiten für den deutschsprachigen Kontext.

In einem weiteren Bereich werden zwei aktuelle Interventionskonzepte vorgestellt:

Gade und Busse stellen eine zielgruppenspezifische Intervention vor, die im Ganztags an Grundschulen mit schreibschwachen Schüler*innen in der dritten und vierten Klassenstufe durchgeführt wurde, um prozessorientiertes Schreiben und kooperatives Lernen zu fördern. Sie diskutieren Herausforderungen und Potenziale einer digitalen Schreibförderung für Schulentwicklungsprozesse.

Dargiewicz, Wehrhöfer, Ohle-Peters, Lauermann und McElvany präsentieren die Konzeption eines Interventionsprojektes. Sie beschreiben, wie

im Rahmen des Projektes eine digitale Sprachförderung in den Bereichen Wortschatz und Grammatik für den Unterricht in der ersten Klassenstufe entwickelt wird, um schulisch relevanten Wortschatz zu fördern sowie das Genus- und Kasusystem zu trainieren.

Die Beiträge des 24. Bandes des Jahrbuchs der Schulentwicklung geben einen breiten Überblick über aktuelle schul- und unterrichtsbezogene Themen im Kontext sprachlicher und sozialer Heterogenität. Sie leisten damit einen wertvollen Beitrag für die Bildungsforschung.

Besonderer Dank gilt allen Gutachter*innen, die die interne und externe Begutachtung übernommen haben und mit ihrer Expertise zur Sicherung der Qualität der Beiträge des vorliegenden Jahrbuchs beigetragen haben. Neben den Kolleg*innen am IFS danken wir Birgit Becker, Michaela Gläser-Zikuda, Markus Hasselhorn, Jörg Jost, Harm Kuper, Katrin Lintorf, Elena Makarova, Camilla Rjosk, Nadine Spörer und Tobias C. Stubbe herzlich für ihre Mitwirkung.

Dortmund, im Juli 2025

Nele McElvany

Sabrina König

Rahim Schaufelberger

Michael Becker

Hanna Gaspard

Birgit Heppt

Alexander Naumann

Herausgeber*innen

Teil A:
Bestandsaufnahmen, Analysen
und Forschungsbefunde

Reading Competence and Vocabulary of Students From Diverse Language Backgrounds: Employing a Lexical Distance Measure

Sebastian Nicolas Thomas Vogel & Justine Stang-Rabrig

Zusammenfassung

Mit Blick auf die wachsende sprachliche Heterogenität fokussiert dieser Beitrag auf ein Maß lexikalischer Distanz zwischen der Herkunftssprache von multilingualen Lernenden und der Zielsprache Deutsch, welche Verkehrs- und Unterrichtssprache zugleich ist. Lexikalische Distanz wird als Prädiktor von Lesekompetenz und Wortschatz, zentrale Indikatoren des Spracherwerbs im Deutschen, genutzt. Darüber hinaus wird das Alter zugewanderter Lernender bei der Ankunft in Deutschland als möglicher Moderator von Effekten lexikalischer Distanz auf den Spracherwerb im Deutschen berücksichtigt. Anhand einer sprachlich heterogenen Stichprobe von $N = 193$ multilingualen Vierklässler*innen mit einem hohen Anteil zugewanderter Lernender zeigten Regressionsanalysen signifikant negative Zusammenhänge der lexikalischen Distanz mit Lesekompetenz, aber nicht Wortschatz, wenn für die Sprache, die die Lernenden primär zuhause sprechen (Deutsch vs. Herkunftssprache), die generellen kognitiven Fähigkeiten und den sozioökonomischen Status kontrolliert wurde. Das Alter zugewanderter Lernender bei der Ankunft war weder als Moderator der lexikalischen Distanz noch direkt mit Lesekompetenz oder Wortschatz verbunden. Die Ergebnisse konnten also zeigen, dass lexikalische Distanz unabhängig von Zuwanderungserfahrungen der Lernenden negativ mit der Entwicklung von Lesekompetenz im Deutschen assoziiert war, während es keinen Zusammenhang mit der Herausbildung des Wortschatzes über den Einfluss der primär gesprochenen Sprache und der allgemeinen kognitiven Fähigkeiten hinaus gab.

Schlagworte: Lexikalische Distanz, sprachliche Heterogenität, Lesekompetenz, Wortschatz, zugewanderte Lernende, Grundschule

Abstract

In light of growing linguistic heterogeneity, this contribution focuses on a measure of lexical distance between multilingual students' heritage language and German as a predictor of reading competence and vocabulary, indicators of the acquisition of language skills in the target language German, the common language and language of instruction in German schools. Additionally, first-generation immigrant students' age at arrival in Germany is considered as a potential moderator of lexical distance effects for language acquisition in the target language. Using a sample of $N = 193$ multilingual fourth-grade students with a large share of first-generation immigrant students and high linguistic diversity, regression analyses showed that lexical distance was significantly negatively related to reading competence, but not vocabulary when considering whether students primarily spoke German or the heritage language at home, general cognitive abilities, and socioeconomic status as important control variables. Including immigrant students' age at arrival and its interaction with lexical distance to check moderation effects revealed no additional significant associations. Therefore, independent of immigration experiences, lexical distance was detrimental to developing reading comprehension in German but did not affect vocabulary learning beyond the effects of general cognitive abilities and primarily speaking the heritage language rather than German at home.

Keywords: Lexical distance, linguistic diversity, reading competence, vocabulary, immigrant students, primary school

1. Introduction

Global migration movements cause classrooms to become increasingly heterogeneous in terms of cultures and languages. This especially applies to Germany as well, a popular destination for immigration, where 21.0% of fourth grade students now primarily speak another language than the common language German at home (Stubbe et al., 2023). Studies have shown that these students are often disadvantaged regarding the development of reading competence and acquisition of vocabulary in German (e.g., Henschel et al., 2022; Novita et al., 2022). However, most studies compared students based on

dichotomized measures of primarily speaking German at home or not, but did not consider the large heterogeneity of languages spoken by the students who do not exclusively speak German (e.g., Stubbe et al., 2023). Although such approaches have advantages, they artificially reduce variance in the group of multilingual students, despite theoretical frameworks suggesting that specific aspects of language learning might be helped or hindered when students' first language is more or less similar to the target language (e.g., Chung et al., 2019). Furthermore, another crucial individual aspect when learning a new language is age, which might be of particular importance for immigrant students learning the common language in their new country of residence as it directly relates to their exposure to the new language. Nevertheless, the role of immigrant students' age at arrival has rarely been studied in regard of its potential moderating role for language learning when considering the linguistic variety of heritage languages (e.g., Schepens et al., 2013a). To address these research desiderates, we employed a measurement of lexical distance to assess the heterogeneity in multilingual students' backgrounds and its role for reading competence and vocabulary in the common language German as two indicators of language learning. We investigated these questions in a multilingual sample surveyed in the fourth year of primary school which is not only an important time in the development of central language competences, especially reading (Chall, 1983), but a central point for shaping educational trajectories in the German school system as well (e.g., Maaz & Nagy, 2010). Finally, we also considered the role of immigrant students' age at arrival in Germany as a potential moderator of linguistic distance's effects on language learning.

2. Theory

2.1. Immigrant-Origin Students and Language Diversity in German Schools

As one of the most sought-out destinations for immigration worldwide (International Organization for Migration [IOM], 2021), society and by extension schools in Germany are continuously becoming more diverse. The share of immigrant-origin students in Grade 4 throughout Germany was 38.3% in 2021, marking an increase of more than 50% in ten years, with the proportion of first-generation immigrant students growing disproportionately, from making up less than one tenth to now representing more than one quarter of all immigrant-origin students (Henschel et al., 2022). Since immigrant fam-

ilies move to Germany from a large variety of different regions and countries of origin (Statistical Federal Office, 2022) and the vast majority of immigrant-origin students grows up speaking at least one heritage language next to German at home (Henschel et al., 2022), the language diversity among German students is at an unprecedented high, especially in primary schools.

2.2. Language-Related Outcomes in Primary School

Theoretical models regarding the stages of reading development (Chall, 1983) and central developmental tasks (McCormick et al., 2011) as well as international educational guidelines (Organization for Economic Co-operation and Development [OECD] et al., 2015) emphasize the importance of students developing adequate reading competence during the first years of primary school that allows them to use reading as a means to learn new information in later years of education and life. For this shift from learning to read to learning through reading (see Chall, 1983) to happen by the end of primary school, the development of good *reading competence*, describing students' ability to engage with texts in a way that allows them to generate meaning (McElvany et al., 2009), is crucial. Other closely related measures like the formation of an adequate *vocabulary*, which allows students to understand the lexical meaning of words and the concepts related to them (Aarnoutse et al., 2001), can also be understood to be indicators of achieving good language skills.

For multilingual students, who in Germany usually are of immigrant origin (Henschel et al., 2022), acquiring language skills in the country's primary language is of special importance as it represents a core acculturative task for these students (Suárez-Orozco et al., 2018). Accordingly, reading competence and vocabulary in the language of instruction have been understood and studied frequently as important indicators of language proficiency of multilingual and language-minority students. Studies focusing on these constructs tend to find disadvantages of language-minority students compared to students speaking German as their primary language in terms of both reading competence (e.g., Segerer et al., 2021; Seuring et al., 2020; Stubbe et al., 2023) and vocabulary (e.g., Kigel et al., 2015; Marx et al., 2015; Novita et al., 2022). However, empirically investigating language-minority students inevitably means that the large heterogeneity of language backgrounds has to be operationalized in a way that still allows for meaningful statistical analyses. In consequence, studies frequently do not distinguish between different language backgrounds at all (e.g., Kigel et al., 2015; Stubbe et al., 2023) or only include large language-minority groups on their own (in the German context, often

Turkish-language background students, e.g., Marx et al., 2015; Segerer et al., 2021) and subsume all other language-minority students in a single group. While the advantages of this method are apparent, it makes a compromise of sacrificing variance in students' language backgrounds to enable statistical comparison of the groups. Crucially, analyzing language effects in this way implicitly assumes that differences in multilingual students' reading competence or vocabulary when compared to other students will only be found if German is not their primary language, and that this is the only explanatory factor. That is not to say that meaningful results cannot still be attained with this method; however, it does by design largely or entirely negate the role of individual characteristics of multilingual students' heritage language for language learning and often excludes multilingual students who speak German as the primary language.

2.3. Language Diversity and Language Learning

Theories of cross-language transfer have recognized the role of the similarity of the target language that is to be learned and the language – or languages – that speakers have already achieved some degree of fluency in for language learning. This aspect of language similarity, reversely framed as language distance (i.e., the dissimilarity of two languages' linguistic features), is identified as one key factor of the interactive language transfer framework by Chung et al. (2019) which combines features of four influential cross-language transfer frameworks. While some skills, like phonological awareness, appear largely unaffected by language distance, the authors emphasize that shared linguistic features can help with the transfer of others, for example morphological and cognate awareness as well as orthographic processing to the second language. Similarly, the established model of Chiswick and Miller (1995, 2007), which focuses on the economics of language learning among immigrants and other language minorities, includes linguistic distance as part of the efficiency of language acquisition and one of the core determinants of language proficiency in second language learning.

Different studies have investigated the role of linguistic distance for language learning, and more specifically reading competence and vocabulary acquisition as central indicators of language proficiency. Common ways to operationalize linguistic distance measures utilize either morphological features, which are a result of the languages' systems of forming words and reflected in their internal structure (Aronoff & Fudeman, 2023), or lexical distance between languages, describing dissimilarities based on languages' basic

vocabularies and shared lexical forms (Schepens et al., 2016). Between these two attempts, lexical distance appears to be better suited to explain differences in language proficiency (e.g., Schepens et al., 2013b, 2016). Therefore, we will focus on lexical distance as an indicator for linguistic distance in this study.

Investigating lexical distance, a series of studies using large databases of second language learners' Dutch proficiency tests has shown negative associations with language proficiency for learners with Indo-European (Schepens et al., 2013a) as well as non-Indo-European first languages (Schepens et al., 2013b). Additionally, lexical distance of both the first and second language was found to explain differences in language proficiency when Dutch was learned as a third language (Schepens et al., 2016). Similarly, using a German dataset, a negative association of lexical distance with German proficiency could be shown in a sample of immigrants over the age of 17, even when controlling for multiple other relevant constructs (Isphording & Otten, 2011). Investigating reading proficiency specifically rather than language proficiency more broadly, Borgonovi and Ferrara (2020) showed a significant negative association with lexical distance in adolescent language-minority students even when multiple control variables were considered. Finally, a study focused on Dutch adolescents learning English as a second language showed that vocabulary gains were higher when the lexical distance of English target words and their Dutch counterparts was small (Mulder et al., 2019), implying that building a vocabulary should be relatively harder the more distant the target language is to other languages the learner is already speaking.

2.4. Language Learning and Age of Immigration

Another aspect that theoretical models deem crucial for learning a new language is exposure to that language (e.g., Chiswick & Miller, 2007; Esser, 2006). This factor can be considered especially important for first-generation immigrants. Due to the strong prevalence of German as a language in public spaces in Germany, native-born multilingual speakers will usually have some exposure to German in their every-day life even if it is not the primary language spoken in their family. In contrast, first-generation immigrants will usually have little to no exposure to German before immigration, as indicated by low average self-reported German skills upon immigration among different immigrant groups (e.g., Kristen & Seuring, 2021) and the low number of immigrants from German-speaking countries (Statistical Federal Office, 2022). Therefore, meaningful exposure to the German language often begins

upon the arrival in Germany, especially for young immigrants who lack the resources to begin learning German on their own prior to migration.

The role of lexical distance for language learning can arguably be more important for students who have less exposure to the language they are striving to learn. If language-learners have had sufficient opportunities to gain a core understanding of the language, lexical distance to their heritage language might be less relevant to learning the new language as they can draw from the knowledge about said language they already have developed through exposure. However, if they have had little exposure to the new language, they can only relate it to their heritage language in their learning process, meaning that students with a more dissimilar heritage language would be additionally disadvantaged in such scenarios. Therefore, theoretical models of second language learning include the efficiency of language learning, which is influenced by linguistic distance, and the exposure to the target language not only as relevant factors in their own right, but also their interaction as another important contributor (e.g., Esser, 2006). Extant research appears to back these theoretical considerations: Schepens et al. (2013a) found a significant negative interaction of lexical distance and age at arrival in the host country, implying that lexical distance hindered language acquisition more the older subjects were upon their immigration. A similar interaction was present among adolescent immigrant students who had participated in the PISA study between 2006 and 2015, where a positive interaction of lexical distance and early arrival (i.e., before the age of 12) emerged (Borgonovi & Ferrara, 2020). This again implied a stronger negative impact of lexical distance for language learning when students' later arrival in the host country had largely prevented earlier exposure to the new language.

2.5. Aims of the Study and Research Questions

The reported theoretical considerations and extant research clearly show that the individual features of students' heritage language, indicated by lexical distance, have an influence on the acquisition of a second language, and reading competence and vocabulary in that language specifically. The role of lexical distance might be even more pronounced for first-generation immigrant students with little exposure to the target language before migration. However, previous studies were focused on either adolescent students (Borgonovi & Ferrara, 2020; Mulder et al., 2019) or largely adult immigrants (Isphording & Otten, 2011; Schepens et al., 2013a, 2013b, 2016) whereas students in primary school have not been focused, despite its status as a critical phase in language

learning and the development of reading competence. Additionally, to our knowledge no study has attempted to investigate lexical distance in the context of German primary schools specifically, where students speaking a large variety of first languages face the task of becoming proficient in German. Therefore, we investigated the following research questions:

Research Question 1a: How is lexical distance of multilingual students' heritage language to German related to their reading competence and vocabulary in German?

Hypothesis 1a: The lexical distance of multilingual students' heritage language to German is significantly negatively related to students' reading competence and vocabulary in German.

Research Question 1b: Do the associations identified in Research Question 1a remain when controlling for which language (German vs. heritage language) students primarily speak at home?

Hypothesis 1b: The negative association between lexical distance of students' heritage language to German and students' reading competence and vocabulary in German is lowered but remains significant when controlling for which language students primarily speak at home.

Research Question 1c: Do the relations identified in Research Question 1b change when controlling for other significant contributors to students' reading competence and vocabulary (general cognitive abilities, socioeconomic status)?

Hypothesis 1c: The negative association between lexical distance of students' heritage language to German and students' reading competence and vocabulary in German remains stable when additionally controlling for general cognitive abilities and socioeconomic status.

Research Question 2: Is the relation of lexical distance with reading competence and vocabulary moderated by first-generation immigrant students' age at arrival in Germany?

Hypothesis 2: Age at arrival in Germany significantly negatively moderates the association of lexical distance with reading competence and vocabulary.

3. Methods

3.1. Participants and Procedure

The analysis sample consisted of $N = 193$ multilingual fourth-grade students from 47 classes in the Ruhr metropolitan area of North Rhine-Westphalia, Germany, after all students who exclusively spoke German at home ($n = 74$),

for whom insufficient information on the language spoken at home was available ($n = 10$), or whose data were unfit for analyses ($n = 2$) had been excluded. Students were 10.49 years old on average ($SD = 0.58$), and the sample was skewed towards female students (58.5% female, 39.9% male). Of the participants, 50.8% primarily spoke a heritage language and never or only sometimes German at home, whereas 49.2% mostly or almost always spoke German at home, and only sometimes a heritage language. Students reported a total of 42 different languages, detailed information on the languages is depicted in Table 1. The majority of the sample were immigrant-origin students with at least one parent born abroad (89.1%), including a comparably large amount of first-generation immigrant students born outside of Germany ($n = 85$), who on average were 10.63 years old ($SD = 0.63$) at the time of the study and had been 4.74 years old ($SD = 2.50$) upon their arrival in Germany. The share of students primarily speaking a heritage language at home was significantly higher in first-generation immigrant students (74.1%) than in the native-born student subsample (32.4%), $\chi^2(1) = 31.5, p < .001$. Data was collected towards the end of the 2021/2022 school year in the context of the MERCUR-funded study *School Integration of Newly Immigrated Children (SIGN)*, a joint research project by the University of Duisburg-Essen, Ruhr University Bochum, and the Center for Research on Education and School Development at TU Dortmund University. During the study, which took part in the classroom context or small groups during regular instruction time, students filled in well-established, standardized tests and self-report questionnaires in paper-pencil form. The study was conducted by two or three trained research administrators per class, depending on class size, and took 90 minutes. Participation in the study required parents' consent and was voluntary, with students who did not participate being given alternative tasks by their teachers. The study met ethical standards and was approved by the ethics committee of the Department of Psychology, University of Duisburg-Essen.

Table 1: Number of Times Languages and Language Families Other Than German Were Reported as Being Spoken at Home in the Sample

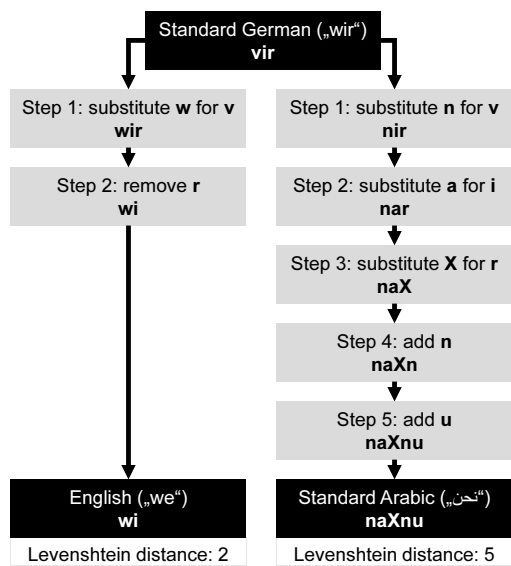
Language	<i>n</i>	%
Afro-Asiatic languages	59	30.6
Arabic (incl. different dialects)	53	27.5
<i>Other Afro-Asiatic languages</i>	6	3.1
Dravidian languages	2	1.0
Indo-European languages	123	63.7
English	28	14.5
Kurdish	14	7.3
Polish	17	8.8
Romanian	7	3.6
Russian	6	3.1
Spanish	8	4.1
<i>Indo-Aryan languages</i>	8	4.1
<i>Italo-Western Romance languages</i>	9	4.6
<i>Persian languages</i>	7	3.6
<i>South Slavic languages</i>	10	5.2
<i>Other Indo-European languages</i>	9	4.6
Niger-Congo languages	12	6.2
<i>Volta-Congo languages</i>	10	5.2
<i>Other Niger-Congo languages</i>	2	1.0
Turkic languages	35	18.1
Turkish	34	17.6
<i>Other Turkic languages</i>	1	0.5

Note. Students could name multiple languages spoken at home. Only languages reported by more than five children are depicted separately. All other languages were grouped into the next higher subgroup (indicated by *italics*) following classification by Eberhard et al. (2024) until a group comprising more than five students was reached.

3.2. Measures

An overview of descriptive information for all measures included in the study is given in Table 2. To assess *lexical distance* between students’ heritage languages and German, we used the Automated Similarity Judgment Program (ASJP; Wichmann et al., 2022), which encompasses information on 5590 distinct languages following the ISO 639-3 classification (International Organization for Standardization, 2023). ASJP makes use of wordlists of these different languages containing 40 words (Holman et al., 2008), a subset of a 100-word list developed by Swadesh (1955) for lexicostatistic dating. These words are transcribed into a standardized orthography featuring seven vowel and 34 consonant symbols (Holman et al., 2008). With these wordlists, a Levenshtein distance is calculated (Levenshtein, 1966), an indicator of the smallest possible number of substitutions, additions, or deletions of symbols necessary to transform the target word in one language into the word with the same meaning in the other (for an example, see Figure 1).

Figure 1: Example for the Calculation of a Levenshtein Distance for the Target Word “We” for Standard German With English and Standard Arabic, Respectively



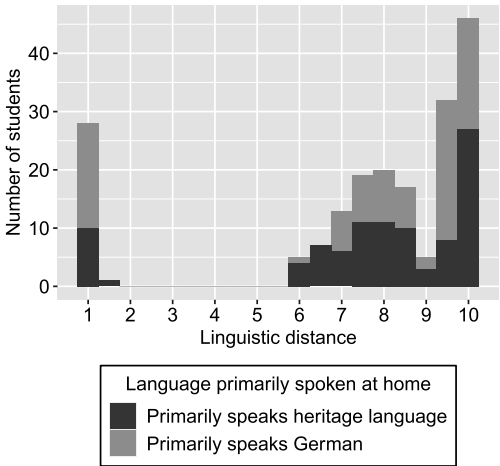
Note. ASJP standardized orthography transcriptions of the words that are used in the calculation of the Levenshtein distance are printed bold.

The distance index is then modified in two steps to take into account coincidental similarities, for example introduced by a general overlap of the phoneme inventories of two languages, and better reflect actual relation between languages (for a detailed description, see Wichmann et al., 2010). We calculated the lexical distance to German¹ for each language reported by students, then transformed values so they would fall on a scale ranging from 1 (lowest distance, in our sample: English) to 10 (highest lexical distance to German, in our sample: Fula). Students who reported speaking multiple languages other than German at home were appointed the value of the language with the smallest lexical distance to German. The distribution of lexical distance in the sample, differentiated by the language students primarily spoke at home, is depicted in Figure 2. The language families depicted in Table 1 are roughly reflected in the distribution of linguistic distance: While Germanic languages, mostly English in our sample, have a value equal or close to 1, other Indo-European languages tend to fall in the range between 6 and 9,

1 ASJP offers two wordlists for Standard German, the results reported here use wordlist “Standard German”. We also replicated analyses using wordlist “Standard German 2”, and results remained stable.

whereas languages from other language families mostly are assigned values between 9 and 10.

Figure 2: Distribution of Lexical Distance to German in the Sample, Differentiated by Language Primarily Spoken in the Student's Home



As dependent variables, we assessed *reading competence*, a measure of students' ability to understand information presented in short texts, with 26 items from the standardized ELFE II-instrument (Lenhard et al., 2017). Additionally, we utilized 16 items taken from the German PIRLS 2021 study to measure students' passive *vocabulary* (Schaufelberger et al., 2024). In this test, students had to pick a synonym for each target word among four options offered (e.g., "room" – "view/chamber/meal/music"). Both instruments showed good internal consistency, $\alpha = .86$ for reading competence and $\alpha = .68$ for vocabulary, respectively. As control variables, we used a dichotomized index of the frequency of speaking German at home as an indicator whether the primary language each student spoke at home was German or the heritage language (0 = *primarily speaks heritage language [never or sometimes German]*, 1 = *primarily speaks German [mostly or almost always German]*). We further included students' general *cognitive abilities*, assessed with a 25-item figural subtest of the well-established KFT 4-12+R test (Heller & Perleth, 2000), which showed excellent internal consistency, $\alpha = .90$. Regarding *socioeconomic status*, we utilized student-reported number of books at home (Schaufelberger et al., 2024) to form a dichotomous indicator of socioeconomic status (0 = *below or average number of books [0–10 books, 11–25 books]*, 1 = *above average number of books [26–100 books, 101–200 books, more than 200 books]*). Finally, for

Research Question 2, we also assessed immigrant students’ *age at arrival* in Germany, which was calculated as the difference between students’ age at the time of the study and their self-reported number of years of living in Germany. For students who had been born in Germany, age at arrival was set to zero.

Table 2: Correlations and Descriptive Information of All Measures

Measure	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1) Lexical distance	–	-.05	-.27*	-.13	-.03	-.16	.11*
2) Primary lang. ^a		–	.13	.29*	.13	.06	-.41*
3) Reading comp.			–	.51*	.18	.11	-.26*
4) Vocabulary				–	.24*	.15	-.30*
5) Cognitive ability					–	.10	-.06
6) SES ^a						–	-.21†
7) Age at arrival ^b							–
<i>M</i>	7.46	0.49	11.32	7.01	12.86	0.40	2.02
<i>SD</i>	2.92	–	4.63	3.12	6.46	–	2.85
% missing	0.00	0.00	1.00	0.00	0.52	5.18	2.59

^a Dummy-coded binary variable (primary language: 0 = *heritage language*, 1 = *German*. SES: 0 = *below-average number of books*, 1 = *above-average number of books*.).

^b Reported for the whole sample. Among the immigrant-student subsample, the mean age at arrival was *M* = 4.74 (*SD* = 2.50).

†*p* < .05. **p* < .10.

3.3. Analytic Approach

All analyses were conducted with R 4.3.1 (R Core Team, 2023), run in the RStudio environment (Posit team, 2023). For all research questions, we specified (multiple) regression models with reading competence and vocabulary, respectively, as dependent variables. To make results more easily comparable, reading competence and vocabulary were standardized for the analyses. For Research Question 1a, we included lexical distance as a predictor. To assess Research Question 1b, we added information about students’ primary language at home as a control variable. Similarly, for Research Question 1c, cognitive abilities and socioeconomic status were included in the models as additional control variables. Finally, to answer Research Question 2, we additionally included students’ age at arrival in Germany as well as the interaction of age at arrival and lexical distance as predictors.

Power analyses revealed that, assuming mid-sized effects, an α -error rate of 5%, and a power of 95%, a sample of *N* = 146 students was required, meaning that the sample of *N* = 193 students was sufficient for analyses.

4. Results

4.1. Descriptive Results

Intercorrelations among all variables included in the study are depicted in Table 2, alongside descriptive information. Correlations indicated a large positive association of the two dependent variables, reading competence and vocabulary. The positive correlation found between lexical distance and age at arrival must be interpreted with the coding of the variables in mind (i.e., native-born students' age at arrival being coded as 0). Among the subsample of first-generation immigrant students, the correlation between the two variables was no longer significant ($r = -.01$, not depicted). Due to these findings in the subsample and the coding of the variables, we can infer that the correlation of lexical distance with age at arrival indicates that multilingual students who were born in Germany on average spoke languages more similar to German at home than immigrant students.

4.2. Associations of Lexical Distance With Reading Competence and Vocabulary

Results of the regression models for Research Questions 1a, 1b, and 1c, respectively, are reported in Table 3. All specified regression models significantly predicted reading competence, $F_{\text{RQ1a}}(1, 189) = 14.70, p < .001$; $F_{\text{RQ1b}}(2, 188) = 8.76, p < .001$; $F_{\text{RQ1c}}(4, 175) = 5.63, p < .001$. Regarding vocabulary, all models were significant, although Model 1a only at the 10%-level, $F_{\text{RQ1a}}(1, 191) = 3.16, p = 0.077$; $F_{\text{RQ1b}}(2, 190) = 9.94, p < .001$; $F_{\text{RQ1c}}(4, 177) = 7.32, p < .001$.

Specifically, analyses showed that lexical distance was negatively related to both reading competence and vocabulary. For reading competence, these relations remained stable when information about students' primary language, cognitive abilities, and socioeconomic status were included as control variables. This indicates that even under control of other central predictors, multilingual children's average reading competence was lower when their heritage language's lexical distance to German was higher. For example, when control variables are included, a child speaking Arabic at home, receiving a value of 9.86 for lexical distance in our sample, would on average score 0.87 standard deviations lower on the reading competence test.

Table 3: Results of Regression Analyses for Research Questions 1a, 1b, 1c and 2

	Reading compet.		Vocabulary	
Predictors	<i>b</i>	<i>p</i>	<i>b</i>	<i>p</i>
<i>Research Question 1a</i>				
Lexical distance	-0.09*	<.001	-0.04†	.077
<i>R</i> ²	.07		.01	
<i>Research Question 1b</i>				
Lexical distance	-0.09*	<.001	-0.04	.103
Primary lang.	0.23	.103	0.56*	<.001
<i>R</i> ² _{adjusted}	.08		.09	
<i>Research Question 1c</i>				
Lexical distance	-0.09*	<.001	-0.03	.196
Primary lang.	0.15	.297	0.48*	<.001
Cognitive abilities	0.02*	.042	0.03*	.005
SES	0.10	.477	0.22	.133
<i>R</i> ² _{adjusted}	.09		.12	
<i>Research Question 2</i>				
Lexical distance	-0.08*	.006	-0.03	.320
Primary lang.	-0.04	.792	0.28†	.072
Cognitive abilities	0.02*	.037	0.03*	.005
SES	-0.01	.960	0.09	.540
Age at arrival	-0.07	.394	-0.10	.202
Lexical distance × Age at arrival	-0.00	.902	0.00	.739
<i>R</i> ² _{adjusted}	.12		.14	

Note. Regression coefficients are unstandardized, but dependent variables were standardized before analyses.

**p* < .05. [†]*p* < .10.

Regarding vocabulary, the small negative association between lexical distance and test score found in the model for Research Question 1a was no longer significant once information regarding students' primary language was included, as speaking German rather than the heritage language as the primary language at home was strongly positively associated with students' vocabulary. These findings remained stable when the additional control variables were included as well.

The amount of variance explained by these models, following conventions by Cohen (1988), was small to medium. Moreover, Hypotheses 1a, 1b, and 1c were supported by the analyses regarding reading competence, but only Hypothesis 1a was supported for vocabulary.

4.3. Moderation of the Influence of Lexical Distance by Age at Arrival

Results of the moderation analyses described in Research Question 2 are also depicted in Table 3, with the overall regression models again being significant for predicting reading competence, $F(6, 170) = 4.83, p < .001$, and vocabulary, $F(6, 171) = 5.72, p < .001$. Neither immigrant students' age at arrival nor its

interaction with lexical distance were significant predictors of students' reading competence or vocabulary in German. The models explained a medium amount of variance (Cohen, 1988). Hypothesis 2 could not be confirmed for either reading competence or vocabulary.

5. Discussion

5.1. Discussion of Findings

The aim of this study was to investigate the role of the lexical distance of multilingual elementary school students' heritage language to the target language German when regarding reading competence and vocabulary in German. Lexical distance was proposed as an alternative measure that reflects the heterogeneity of students' languages, in contrast to common approaches that subsume multilingual students from a variety of different language backgrounds in one group and, to our knowledge, implemented in the context of German elementary school for the first time.

Analyses regarding Research Question 1a showed significant negative associations of lexical distance with both reading competence and vocabulary. When information whether German or a heritage language was spoken as the primary language at home (Research Question 1b), general cognitive abilities, and socioeconomic status (Research Question 1c) were included as control variables in the analyses, lexical distance remained a significant predictor of reading competence, but not vocabulary. These findings imply that multilingual students on average had lower German-language reading competence the more dissimilar their heritage language was to German, whereas speaking German as primary language at home rather than the heritage language, but not lexical distance was significantly related to the formation of students' vocabulary. Confirming the assumptions made in Hypotheses 1a and, for reading competence, Hypotheses 1b and 1c, these results are in line with previous findings regarding the role of lexical distance for language competence in general (e.g., Isphording & Otten, 2011) as well as reading competence (e.g., Borgonovi & Ferrara, 2020), but not vocabulary (e.g., Mulder et al., 2019) in older students.

For Research Question 2, first-generation immigrant students' age at arrival and its interaction with lexical distance were additionally included in the analyses. Contrary to our expectations and previous findings, neither immigrant students' age of arrival itself nor its interaction with lexical distance were significantly related to either reading competence or vocabulary in

German (e.g., Borgonovi & Ferrara, 2020; Schepens et al., 2013a). This indicates that the effects of lexical distance on acquiring reading competence and vocabulary in German were similar for all students, independent of whether they were born in Germany and, if they were not, their age upon their arrival in Germany. Therefore, Hypothesis 2 could not be confirmed.

5.2. Limitations and Strengths

Some limitations must be considered when discussing the results of the study. As data were acquired in a cross-sectional design, we cannot draw causal inferences from the analyses. Therefore, we also could not include reciprocal effects of students' language abilities in German and their heritage language used at home. For our analyses, we assumed that students' heritage language usage at home and its lexical distance to German affected their reading competence and vocabulary in German. While this assumption is plausible, it is possible that effects in the opposite direction exist as well: If students' language ability in German are higher, indicated by better reading competence and a larger vocabulary, there might also be a higher probability that they speak German at home as a consequence. Some additional limitations arise from the way students' languages spoken at home were assessed: First, because students only provided concrete information about how often they spoke German at home and merely listed all additional languages, we could not determine which language was spoken most frequently in cases where more than one additional language was listed. Second, this also means that we could not consider the implication of earlier studies that, when learning German as a third language, the lexical distance to both the first and second language may influence students' language acquisition in German to varying degrees (Schepens et al., 2016). Third, while the free-text method of entry offered the possibility to do so, only few students in the largest heritage-language group, Arabic speakers, indicated the dialect they spoke at home (e.g., "Syrian (Arabic)"). Therefore, we classified all Arabic speaking students as if they spoke Modern Standard Arabic, meaning that the considerable variance between the different Arabic dialects (see e.g., Abu Kwaik et al., 2018) was not represented in full in our analyses. The measurement approach in the study also poses a limitation regarding socioeconomic status, as the number of books at home can be used as an informative indicator of socioeconomic status but should ideally be reported by parents rather than students and used in combination with other indicators like information regarding parents' occupation (see e.g., Heppt et al., 2022). Finally, the sample was comparably small and only small

shares of students from each class participated in the study on average which prohibited a multilevel analysis of the associations under consideration of the data clustering in classrooms. Additionally, due to the overrepresentation of multilingual and immigrant students, the sample was not representative for the German student body and results cannot necessarily be generalized for all students.

Nevertheless, the sample composition also poses a strength of the study. Since comparably diverse schools were recruited for the study, the sample included students from a wide variety of language backgrounds and many first-generation immigrant students, allowing a robust analysis of our research questions with a somewhat smaller sample. Another strength is the use of the ASJP database to assess lexical distance of the languages. The large number of languages already implemented in the ASJP database means that reliable information for all relevant languages in our study could be found, but also that the program can easily be used in future studies to assess lexical distance. Although not needed for our analyses, entering information for new languages into the program is easy on a technical level, meaning that the tool can flexibly be adapted as long as reliable information about the 40 words the distance measure is based on can be provided for all languages required. Moreover, including first-generation immigrant students' age at arrival in Germany as well as its interaction with lexical distance allowed us to separate age- and exposure-effects from the influence of lexical distance, but also consider their interplay when it comes to immigrant students' reading competence and vocabulary formation. Finally, in contrast to the majority of studies investigating lexical distance effects, rather than investigating adolescents or adult second-language learners, we focused on students in the last year of primary school, which marks the end of a crucially important period in language learning in general and the development of reading competence specifically.

5.3. Implications for Research and Practice

Several implications for future research can be drawn from the results of our study. First, we included a variable to distinguish whether German or a heritage language was the primary language spoken in students' home. While this reflects students' current language situation, future studies should also investigate how lexical distance affects language learning depending on which language, German or the heritage language, students learned first. Additionally, the role of proficiency in the heritage language should be included in future investigations of lexical distance effects as well to gain a better under-

standing which areas of language learning are affected by lexical distance specifically as opposed to more general spillover effects of a high proficiency in another language, but also how these two aspects may interact. Furthermore, by including other aspects of linguistic distance besides lexical distance, for example considering (dis-)similarities in morphology and orthography between languages, future research can investigate whether these can substantially improve the measurement of linguistic distance and help understand how its various aspects affect different indicators of language learning. In a similar way, other important factors for language learning as highlighted by theoretical models, for example incentives to learn the new language (e.g., Chiswick & Miller, 2007; Esser, 2006), can be included and related to lexical distance in future research. Finally, lexical distance has great potential to investigate associations with language learning beyond the individual level in terms of classroom composition, as the distance between the heritage languages of each pair of students within a classroom can be calculated with relative ease and in turn be used to aggregate a precise indicator of language heterogeneity in the classrooms.

Similarly, implications for educational practice can be deduced from the study at hand. In light of growing multilingualism and increasing linguistic heterogeneity in society in general and in school classes in particular (e.g., Henschel et al., 2022; Statistical Federal Office, 2022) and the repeatedly shown disadvantages of language-minority students regarding school performances such as reading competence (OECD, 2023; Stubbe et al., 2023), it is extremely relevant to provide pedagogical staff with sufficient training in dealing with multilingualism and to advertise existing extracurricular language promotion offers in the respective language of instruction. Corresponding offers in school could include, for example, providing lessons in foreign languages (Mediendienst Integration, 2024). However, this should also be viewed critically in light of the possible separation of these children from their majority-language classmates (see also e.g., Höckel & Schilling, 2022). In any case, it seems advisable to adapt the school system more effectively to multilingualism and to introduce established multilingualism concepts, which already exist in some federal states (Mediendienst Integration, 2024), more widely. Understanding how different aspects of linguistic distance influence the acquisition of language competence in the target language German can help with identifying means how students' existing skills in their heritage language can be used for language learning and areas where they might profit most from additional support as well. For newly immigrated students, a direct and intensive language support in German is essential so that they can catch up with the mainstream class as early as possible (Mediendienst

Integration, 2024). This is especially important considering that well-marked reading competence and vocabulary in German are immensely important not only for German lessons, but also for learning in other subjects and for societal participation. Ultimately, however, it is important to value the existing linguistic diversity and view it as a resource (e.g., Edele et al., 2023; Illman & Pietilä, 2018).

References

- Aarnoutse, C., van Leeuwe, J., Voeten, M., & Oud, H. (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. *Reading and Writing*, 14(1/2), 61–89. <https://doi.org/10.1023/A:1008128417862>
- Abu Kwaik, K., Saad, M., Chatzikyriakidis, S., & Dobnik, S. (2018). A lexical distance study of Arabic dialects. *Procedia Computer Science*, 142, 2–13. <https://doi.org/10.1016/j.procs.2018.10.456>
- Aronoff, M., & Fudeman, K. (2023). *What is morphology?* (3rd ed.). John Wiley & Sons Inc.
- Borgonovi, F., & Ferrara, A. (2020). Academic achievement and sense of belonging among non-native-speaking immigrant students: The role of linguistic distance. *Learning and Individual Differences*, 81, Article 101911. <https://doi.org/10.1016/j.lindif.2020.101911>
- Chall, J. S. (1983). *Stages of reading development*. McGraw-Hill Book Company.
- Chiswick, B. R., & Miller, P. W. (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics*, 13(2), 246–288. <https://doi.org/10.1086/298374>
- Chiswick, B. R., & Miller, P. W. (2007). *The economics of language: International analyses*. Routledge.
- Chung, S. C., Chen, X., & Geva, E. (2019). Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *Journal of Neurolinguistics*, 50, 149–161. <https://doi.org/10.1016/j.jneuroling.2018.01.003>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2024). *Ethnologue: Languages of the world* (27th ed.). SIL International. <https://www.ethnologue.com>
- Edele, A., Seuring, J., Schotte, K., Kristen, C., & Stanat, P. (2023). Is the first language a resource, an obstacle, or irrelevant for language minority students' education? In S. Weinert, G. J. Blossfeld, & H.-P. Blossfeld (Eds.), *Education, competence development and career trajectories: Analysing data of the National Educational Panel Study (NEPS)* (pp. 349–367). Springer International Publishing. https://doi.org/10.1007/978-3-031-27007-9_16
- Esser, H. (2006). *Migration, Sprache und Integration* [Migration, language, and integration] (AKI-Forschungsbilanz No. 4). Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration; Wissenschaftszentrum Berlin für Sozialforschung. <https://nbn-resolving.org/urn:nbn:de:0168-ssaoar-113493>
- Heller, K. A., & Perleth, C. (2000). *KFT 4-12+R: Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision* [KFT 4-12+R: Cognitive ability test for Grades 4 to 12, revision]. Beltz.
- Henschel, S., Heppt, B., Rjosk, C., & Weirich, S. (2022). Zuwanderungsbezogene Disparitäten [Immigration-related disparities]. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2021: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (pp. 181–219). Waxmann.

- Heppt, B., Olczyk, M., & Volodina, A. (2022). Number of books at home as an indicator of socioeconomic status: Examining its extensions and their incremental validity for academic achievement. *Social Psychology of Education*, 25(4), 903–928. <https://doi.org/10.1007/s11218-022-09704-8>
- Höckel, L. S., & Schilling, P. (2022). *Starting off on the right foot – Language learning classes and the educational success of immigrant children* (Ruhr Economics Paper No. 983). RWI – Leibniz-Institut für Wirtschaftsforschung. <https://doi.org/10.4419/96973148>
- Holman, E. W., Wichmann, S., Brown, C. H., Velupillai, V., Müller, A., & Bakker, D. (2008). Explorations in automated language classification. *Folia Linguistica*, 42(2), 331–354.
- Illman, V., & Pietilä, P. (2018). Multilingualism as a resource in the foreign language classroom. *ELT Journal*, 72(3), 237–248. <https://doi.org/10.1093/elt/ccx073>
- International Organization for Migration. (2021). *World migration report 2022*. United Nations. <https://www.un-ilibrary.org/content/books/9789292680763>
- International Organization for Standardization (2023). *Code for individual languages and language groups* (ISO Standard No. 639:2023). <https://www.iso.org/standard/74575.html>
- Ispording, I. E., & Otten, S. (2011). *Linguistic distance and the language fluency of immigrants* (Ruhr Economics Papers No. 274). Rheinisch-Westfälisches Institut für Wirtschaftsforschung. <https://hdl.handle.net/10419/61444>
- Kigel, R. M., McElvany, N., & Becker, M. (2015). Effects of immigrant background on text comprehension, vocabulary, and reading motivation: A longitudinal study. *Learning and Instruction*, 35, 73–84. <https://doi.org/10.1016/j.learninstruc.2014.10.001>
- Kristen, C., & Seuring, J. (2021). Destination-language acquisition of recently arrived immigrants: Do refugees differ from other immigrants? *Journal for Educational Research Online*, 13(1), 128–156. <https://doi.org/10.31244/jero.2021.01.05>
- Lenhard, W., Lenhard, A., & Schneider, W. (2017). *ELFE II: Ein Leseverständnistest für Erst- bis Siebtklässler* [ELFE II: A reading comprehension test for Grades 1 to 7]. Hogrefe.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics-Doklady*, 10(8), 707–710.
- Maaz, K., & Nagy, G. (2010). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems: Definitionen, Spezifikationen und Quantifizierung primärer und sekundärer Herkunftseffekte [The transition from elementary to secondary education in Germany: Definition, specification, and quantification of the primary and secondary effects of social background]. In K. Maaz, J. Baumert, C. Gresch, & N. McElvany (Eds.), *Der Übergang von der Grundschule in die weiterführende Schule: Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (pp. 151–180). Bundesministerium für Bildung und Forschung (BMBF) Referat Bildungsforschung.
- Marx, A., Stanat, P., Roick, T., Segerer, R., Marx, P., & Schneider, W. (2015). Components of reading comprehension in adolescent first-language and second-language students from low-track schools. *Reading and Writing*, 28(6), 891–914. <https://doi.org/10.1007/s11145-015-9554-3>
- McCormick, C. M., Kuo, S. I.-C., & Masten, A. S. (2011). Developmental tasks across the life span. In K. L. Fingerman, C. Berg, J. Smith, & T. C. Antonucci (Eds.), *Handbook of life-span development* (pp. 117–140). Springer Publishing Company.
- McElvany, N., Becker, M., & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten [The relevance of family characteristics for reading competence, vocabulary, reading motivation and reading behavior]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(3), 121–131. <https://doi.org/10.1026/0049-8637.41.3.121>
- Mediendienst Integration. (2024). *Mehrsprachigkeit* [Multilingualism]. Rat für Migration. <https://mediendienst-integration.de/integration/mehrsprachigkeit.html>

- Mulder, E., van de Ven, M., Segers, E., & Verhoeven, L. (2019). Context, word, and student predictors in second language vocabulary learning. *Applied Psycholinguistics*, 40(1), 137–166. <https://doi.org/10.1017/S0142716418000504>
- Novita, S., Lockl, K., & Gnambs, T. (2022). Reading comprehension of monolingual and bilingual children in primary school: The role of linguistic abilities and phonological processing skills. *European Journal of Psychology of Education*, 37(4), 993–1013. <https://doi.org/10.1007/s10212-021-00587-5>
- Organization for Economic Co-operation and Development. (2023). *Pisa 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Organization for Economic Co-operation and Development, European Union, & UNESCO Institute for Statistics. (2015). *ISCED 2011 Operational Manual: Guidelines for classifying national education programmes and related qualifications*. OECD Publishing. <https://doi.org/10.1787/9789264228368-en>
- Posit team. (2023). *RStudio: Integrated development environment for R* (Version 2023.6.1.524) [Computer software]. Posit Software, PBC. <https://www.posit.co>
- R – Core Team. (2023). *R: A language and environment for statistical computing* (Version 4.3.1) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org>
- Schaufelberger, R., Kleinkorres, R., Becher, L., Ludewig, U., Lorenz, R., & McElvany, N. (2024). *IGLU 2021. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen* [PIRLS 2021. Handbook for the documentation of measurement instruments and working with the datasets]. Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830998990>
- Schepens, J., van der Slik, F., & van Hout, R. (2013a). The effect of linguistic distance across Indo-European mother tongues on learning Dutch as a second language. In L. Borin & A. Saxena (Eds.), *Trends in Linguistics: Studies and monographs: Vol. 265. Approaches to measuring linguistic differences* (pp. 199–229). De Gruyter Mouton.
- Schepens, J., van der Slik, F., & van Hout, R. (2013b). Learning complex features: A morphological account of L2 learnability. *Language Dynamics and Change*, 3(2), 218–244. <https://doi.org/10.1163/22105832-13030203>
- Schepens, J., van der Slik, F., & van Hout, R. (2016). L1 and L2 distance effects in learning L3 Dutch. *Language Learning*, 66(1), 224–256. <https://doi.org/10.1111/lang.12150>
- Segerer, R., Niklas, F., Suggate, S., & Schneider, W. (2021). Young minority home-language students' biased reading self-concept and its consequences for reading development. *Reading Research Quarterly*, 56(1), 71–94. <https://doi.org/10.1002/rrq.300>
- Seuring, J., Rjosk, C., & Stanat, P. (2020). Ethnic classroom composition and minority language use among classmates: Do peers matter for students' language achievement? *European Sociological Review*, 36(6), 920–936. <https://doi.org/10.1093/esr/jcaa022>
- Statistical Federal Office (2022). *Bevölkerung und Erwerbstätigkeit: Bevölkerung mit Migrationshintergrund – Ergebnisse des Mikrozensus 2021* [Population and occupation: Population with migration background – Results of the Mikrozensus 2021]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Publicationen/Downloads-Migration/migrationshintergrund-endergebnisse-2010220217004.pdf?__blob=publicationFile
- Stubbe, T. C., Kleinkorres, R., Krieg, M., Schaufelberger, R., & Schlitter, T. (2023). Soziale und migrationsbedingte Disparitäten in der Lesekompetenz von Viertklässlerinnen und Viertklässlern [Social and migration-related disparities in fourth graders' reading competence]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. C. Stubbe (Eds.), *IGLU 2021: Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 151–177). Waxmann Verlag GmbH.

- Suárez-Orozco, C., Motti-Stefanidi, F., Marks, A., & Katsiaficas, D. (2018). An integrative risk and resilience model for understanding the adaptation of immigrant-origin children and youth. *American Psychologist*, 73(6), 781–796. <https://doi.org/10.1037/amp0000265>
- Swadesh, M. (1955). Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics*, 21(2), 121–137.
- Wichmann, S., Holman, E. W., Bakker, D., & Brown, C. H. (2010). Evaluating linguistic distance measures. *Physica a: Statistical Mechanics and Its Applications*, 389(17), 3632–3639. <https://doi.org/10.1016/j.physa.2010.05.011>
- Wichmann, S., Holman, E. W., & Brown, C. H. (2022). *The ASJP database* (Version 20) [Computer software]. <https://asjp.cild.org>

Effectiveness of Professional Development for Language-Supportive Teaching: Insights From Employing Computational Linguistic Analysis Methods

Birgit Heppt, Denise Löfflad, Sofie Henschel, Katrin Gabler, Ilonca Hardy & Detmar Meurers

Zusammenfassung

Die vorliegende Studie untersucht das Potenzial computerlinguistischer Analyseverfahren zur Evaluation der Effekte von Lehrkräftefortbildungen. Entsprechende Methoden ermöglichen die automatisierte Analyse einer Vielzahl sprachlicher Merkmale und dürften daher eine vertiefte Analyse der mündlichen Unterrichtssprache von Lehrkräften erleichtern. Auf Grundlage eines positiv evaluierten Professionalisierungsansatzes (z. B. Heppt et al., 2022) wurde untersucht, ob Lehrkräfte, die für die Umsetzung fachintegrierter Sprachbildung im Sachunterricht der Grundschule fortgebildet wurden (Interventionsgruppe [IG], $n = 9$), ihren Schüler*innen einen variableren und anspruchsvolleren Sprachinput anboten als Lehrkräfte der Kontrollgruppe (KG; $n = 19$). Anhand von Transkripten von zwei videografierten Sachunterrichtsstunden je Lehrkraft (Thema 1: Schwimmen und Sinken, Thema 2: Verdunstung und Kondensation) fanden wir nur wenige Merkmale, die für eine elaboriertere und komplexere Sprachverwendung der IG im Vergleich zur KG sprechen. Die Gruppenunterschiede waren überdies bei Thema 2 weniger stark ausgeprägt als bei Thema 1, was darauf hindeutet, dass sich die Fortbildungsteilnahme nur geringfügig auf den mündlichen Sprachgebrauch der Lehrkräfte während des Sachunterrichts auswirkte.

Schlagworte: Lehrkräftefortbildung, Sprachbildung, Grundschule, Sachunterricht, sprachliche Komplexität, computerlinguistische Analyseverfahren

Abstract

The present study aims at exploring the potential of using computational linguistic analysis methods for examining professional development (PD) effects. As respective methods allow for an automatic analysis of a wide range of linguistic features, they may facilitate an in-depth analysis of teachers' oral language use in classroom discourse. Building on a positively evaluated PD approach (e.g., Heppt et al., 2022), we investigated whether in-service teachers who were trained for language-supportive teaching in elementary school science classes (intervention group [IG], $n = 9$) provide their students with a more varied and more elaborate language input than teachers from the control group (CG; $n = 19$). Based on transcripts of video recordings of two science lessons per teacher (Topic 1: floating and sinking, Topic 2: evaporation and condensation), we found only a small number of features that point to more sophisticated and complex language use in the IG than in the CG. Moreover, group differences were less pronounced for Topic 2 than for Topic 1, suggesting that PD participation had only small effects on teachers' oral language use during science instruction.

Keywords: Teacher professional development, language support, elementary school, science education, linguistic complexity, computational linguistic analysis methods

1. Introduction

Language-supportive teaching that fosters both domain-specific learning and language development is increasingly seen as a general teaching principle across subjects (e.g., Becker-Mrotzek & Woerfel, 2020; Prediger & Hardy, 2023). Although considered beneficial for all students, such classroom instruction is particularly aimed at students who are at risk of falling behind. Among them are students with an immigrant background, often growing up as dual language learners, who have repeatedly been shown to lag behind their peers without immigrant background in academic achievement and (academic) language proficiency (Henschel et al., 2022; Ludewig et al., 2022). Similarly, students from families with low socioeconomic status (SES) consistently perform, on average, below their peers from high-SES families (e.g., Sachse et al.,

2022). Both student groups are likely to have only limited opportunities for acquiring the register of schooling within their families. Children's language development depends on the amount and quality of language input at home, which, in turn, varies with family SES. Children from high-SES families, thus, typically outperform their counterparts from low-SES families in their mastery of academic language, helping them to benefit from classroom discourse, accomplish written assignments, or understand specialized texts (e.g., Volodina et al., 2021). Students with an immigrant background face similar obstacles as they often grow up in families with low SES (e.g., Henschel et al., 2022). Moreover, many of them have limited access to the language of instruction within their homes, thus further hampering their mastery of the language of schooling (for an overview, see Heppt & Schröter, 2023). With the aim of adapting to these students' needs, language-supportive teaching is considered a potential remedy for tackling educational inequalities.

The need for integrating language support into regular classroom teaching is widely accepted, resulting in an expansion of course offerings on language support and second language acquisition in university teacher training in Germany (Paetsch & Heppt, 2021). While an increasing number of German elementary school teachers participates in such courses during university teacher training, many of them are still ill-prepared for providing subject-integrated language support in their daily classroom instruction (Henschel & Heppt, 2024). This highlights the importance of effective professional development (PD).

In investigating the effectiveness of PD for language-supportive teaching, research frequently focuses on teachers' classroom practice (e.g., Gabler et al., 2024; Heppt et al., 2022; van Dijk et al., 2019). Based on laborious coding of selected language-support strategies (e.g., use of language-supportive questions) or highly inferential ratings of the overall quality of language support, studies have pointed to the general effectiveness of PD for developing teachers' classroom practice (for a meta-analysis, see Kalinowski et al., 2020).

With the present study, we aim at exploring the potential of using computational linguistic analysis methods for examining PD effects. As respective methods allow for an automatic analysis of a wide range of linguistic features, they may facilitate an in-depth analysis of teachers' oral language use in classroom discourse (cf. Weiss et al., 2022). Building on a PD approach that has been positively evaluated regarding participants' knowledge on language support (Heppt et al., 2022) and their use of specific language-support strategies in science instruction (e.g., the use of language-supportive questions; Gabler et al., 2024), this study focuses on teachers' oral language input. Specifically, we investigate whether, upon completing the PD, German elementary school

teachers use more stimulating and sophisticated oral language in their science instruction than teachers who did not take part in the PD.

2. Theoretical and Empirical Background

2.1. Language-Supportive Classroom Instruction

Effective language-supportive classroom instruction helps students attain domain-specific learning goals, while developing the necessary academic language skills. In elementary school science classes, for instance, experiments on the floating and sinking of objects are clearly aimed at developing students' conceptual knowledge on phenomena like water displacement or water pressure. Yet, in order to construct this type of knowledge in a co-constructive process, students need to be able to formulate and justify assumptions and to describe and explain their observations (e.g., Vorholzer & Aufschnaiter, 2019). These language functions form a core part of the academic language register (e.g., Bailey et al., 2007; Prediger & Hardy, 2023). Performing language functions such as "hypothesizing" or "justifying" requires the precise understanding of their respective meaning. In addition, students need knowledge of adequate linguistic structures (e.g., the correct use of causal connectives such as "therefore" or "due to") and domain-specific academic vocabulary (e.g., "to displace", "wax", "Styrofoam"). As an important prerequisite for learning, language functions need to be systematically developed in classroom instruction, along with the underlying lexical and syntactical skills.

This can best be achieved by cognitively activating instruction that engages students in higher-order thinking and connects concepts with activities (e.g., by conducting experiments), as such learning environments typically offer multiple opportunities for using language in meaningful contexts (e.g., Bravo & Cervetti, 2014). In using language functions and expanding their vocabulary knowledge, students should be assisted through language-support strategies (e.g., open-ended questions, rich and elaborate language input, linguistic feedback that adequately expands the students' utterances; e.g., Gabler et al., 2020; Heppt et al., 2022; Mahan, 2020). The linguistic scaffolding approach (Gibbons, 2002), which builds on the theory of social learning (cf. Wood et al., 1976), considers these linguistic aids as a scaffold that helps students master linguistically demanding tasks. These aids are adapted to students' linguistic needs and gradually reduced as the students increasingly gain proficiency in academic language. The linguistic scaffolding approach has been proven effective in promoting students' domain-specific knowledge with no pronounced

differences across student groups (e.g., multilingual and monolingual learners; Prediger & Neugebauer, 2021; Prediger & Wessel, 2017). Research on its effectiveness for (academic) language development has mostly been conducted in the United States (US) with a focus on students with limited language proficiency (English Language Learners; ELLs). Overall, results indicate that integrating inquiry-based science instruction with the adaptive use of a range of language-support strategies benefits ELLs' academic language proficiency (e.g., Bravo & Cervetti, 2014; Llosa et al., 2016).

2.2. Language Input and its Associations with Content- and Language Learning

The quality and amount of oral language input are among the fundamental drivers of children's language development. Since the seminal study by Hart and Risely (1995) on the huge socioeconomic differences in the number of child-directed words, amounting to a "30-million-word gap" for children from high vs. low-SES families by the age of four, numerous studies have highlighted the relation between caregivers' language input and their child's language proficiency. In line with these findings, the amount and quality of a (preschool) teacher's oral language use are also considered an essential language-support strategy in instructional settings (e.g., Gabler et al., 2020; Kane et al., 2023).

With the aim of acting as language role models, this language-support strategy requires (preschool) teachers to provide frequent, rich, and elaborate language input in instruction. This can be achieved by using thinking-aloud techniques or by mapping one's own or students' actions with language (e.g., "I have a wooden dice in my left hand and a metal dice in my right hand. Now, I put the wooden dice into the water basin."). In doing so, teachers should try and use important general and domain-specific vocabulary, ideally in multiple contexts and by contrasting them with other words (e.g., "Did we hypothesize this or did we establish this? We established this in the end. So this is for sure."), contributing to students' enhanced and differentiated vocabulary knowledge.

Prior research on teacher's oral language use has been conducted in preschool settings and, to a lesser extent, in elementary school. Overall, this research showed that preschool teachers' use of elaborate language can increase students' domain-specific learning and language development (e.g., Kane et al., 2023; Studhalter et al., 2021). However, (preschool) teachers tend to use this high-quality language input rather infrequently (see however,

Weiss et al., 2022). Studhalter et al. (2021), for instance, found that preschool teachers' use of domain-specific vocabulary (e.g., "wax", "clay", "iron") during a 4-week learning unit on the topic "floating and sinking" significantly predicted children's conceptual learning gains. Focusing on the interplay between the quality of preschool teachers' talk and students' language development, Dickinson and Porche (2011) conducted a longitudinal study from preschool to fourth grade. Among the indicators used for evaluating the quality of preschool teachers' oral language were sophisticated vocabulary (i.e., low-frequency words) and utterances aimed at focusing children's attention, correcting or expanding their oral expressions. The authors found that each of these indicators of preschool teachers' language quality contributed to students' reading skills in Grade 4 (Dickinson & Porche, 2011).

Despite its pivotal role for student learning, children seem to receive relatively low amounts of high-quality language input during regular classroom teaching in elementary school. Based on videotaped classroom observations of five upper elementary classrooms in the US, Ernst-Slavit and Mason (2011) found that less than 12% of teachers' oral language input across classrooms and subjects can be classified as "academic language", as reflected in vocabulary (e.g., general and domain-specific vocabulary), grammar (e.g., syntactically long and complex sentences with clause connectives) and discourse (e.g., factual and information-dense style). In a similar vein, a Dutch study aimed at assessing elementary school teachers' use of academic language in whole-classroom discourse in mathematics classrooms in Grades 1 and 2 (Dokter et al., 2017). Drawing on transcripts of two eight-minute sequences of classroom instruction per teacher, the authors found that all teachers used math-specific language to some extent. However, the overall lexical and grammatical complexity of teacher talk was rather low. In sum, prior research suggests that teachers do not deliberately use their oral language input in classroom discourse as a means for modeling students' (academic) language development, pointing to the need for effective teacher PD.

2.3. Effectiveness of Teacher PD for Language-Supportive Teaching

Research on teacher PD in general (e.g., Darling-Hammond et al., 2009; Lipowsky & Rzejak, 2015) as well as the emerging literature on PD for language-supportive teaching (for an overview, see Kalinowski et al., 2020) have identified key characteristics of effective teacher PD. Importantly, teacher PD that helps teachers gain knowledge and skills for integrating language-support strategies into regular classroom teaching combines phases of input with

opportunities for actively using the newly acquired knowledge (e.g., in role plays or in classroom instruction). During implementation phases, teachers should receive feedback and have the chance to reflect upon their experiences, for instance by analyzing and discussing video-recordings of their own classroom teaching (Piwowar et al., 2013; van Dijk et al., 2019).

PD programs that consider these core principles have, indeed, been proven effective in developing teachers' knowledge and skills for subject-integrated language support (e.g., Babinski et al., 2018; van Dijk et al., 2019). A meta-analysis incorporating ten studies, showed small, albeit statistically non-significant gains in teachers' self-efficacy and self-assessed knowledge regarding language-supportive teaching, whereas larger PD effects occurred for teachers' language-supportive classroom behavior (Kalinowski et al., 2020). Most of the studies included in the meta-analysis were conducted in the US with a focus on ELLs and only few studies used (quasi-)experimental designs with intervention group (IG) and control group (CG).

A couple of more recent (quasi-)experimental studies from Germany and the Netherlands aimed at teachers in mainstream classrooms, helping them to provide language-supportive science instruction for *all* students (Gabler et al., 2020; Henrichs & Leseman, 2014; van Dijk et al., 2019). In one such PD program based on the linguistic scaffolding approach (Gabler et al., 2020), teachers were familiarized with and actively used core language-support strategies in group work and classroom teaching. The PD covered the following language-support strategies: (1) language modeling by providing elaborate and targeted language input, (2) asking language-stimulating questions, (3) giving language-supportive feedback (e.g., by elaborating on or rephrasing students' answers), and (4) shifting students' attention to important vocabulary and sentence structures (e.g., by using visual aids). This PD program, which also forms the basis of the present study, has been proven effective in advancing teachers' knowledge on subject-integrated language support (Heppt et al., 2022). Moreover, after dealing intensively with a science curriculum on "floating and sinking" and the implementation of possible language scaffolds, trained teachers showed a higher overall quality and amount of language-supportive behavior and provided better learning-related feedback in classroom teaching (Heppt et al., 2022). They also used language-stimulating questions more frequently than their counterparts who did not participate in the PD (Gabler et al., 2024). Regarding the overall quality of instructional support, including language-supportive teaching, no group differences were observed when participants were required to transfer their newly acquired knowledge and skills for language-supportive teaching to a different science curriculum (i.e., on "evaporation and condensation"). Interestingly, though, instructional

quality of both IG- and CG-teachers was higher when teaching this second topic, possibly reflecting that it was easier to implement with regard to both conceptual demands and classroom organization. At the same time, it should be noted that the two science topics require the use of different science vocabulary and focus on partly different language functions and structures. Hence, the topics may also differ in their potential for modeling sophisticated vocabulary, syntactic structures, and language functions. However, group differences in terms of the amount and quality of teachers' language input have not yet been investigated. Consequently, it remains an open question whether teachers' oral language use differs across topics.

2.4. The Potential of Using Computational Linguistic Analysis Methods for Investigating Teachers' Oral Language Input

In general, research on PD for language-supportive teaching has rarely considered the amount and linguistic sophistication (e.g., the use of low-frequency vocabulary) of teachers' oral language in classroom discourse as an outcome (for exceptions, see Henrichs & Leseman, 2014; van Dijk et al., 2019). At least in part, this is probably due to the enormous amount of time and effort required to assess the frequency of selected features in oral language. In their study on the effectiveness of a short intervention on academic language use in early science instruction, Henrichs and Leseman (2014), for instance, focused on two lexical features: lexical diversity, as measured in the number of word types (i.e., number of different words), and lexical sophistication, as measured in the number of general and domain-specific academic words. The number of word types was determined by counting the number of different words used across all (videotaped and transcribed) lessons. Assessing the number of general and domain-specific vocabulary required the authors to (1) create lists of all words used in the teacher-student conversations, (2) manually double-code all words that were classified as cross-disciplinary vocabulary (e.g., "experiment", "describe") or domain-specific vocabulary (e.g., "air pressure", "force"), and (3) determine interrater-reliability and discuss divergent ratings (Henrichs & Leseman, 2014).

Using computational linguistic analysis methods, by contrast, allows for assessing a wider range of linguistic features automatically in a timely manner. This enables a very fine-grained analysis of teacher talk, specifically incorporating a variety of features that are focused upon in the respective intervention. Moreover, computational linguistic analysis methods come along with maximum accuracy regarding the detection of specific linguistic features in

the texts (compared to human raters) and facilitate the implementation of computationally complex and sophisticated features. Whereas determining the number of word types as an indicator of lexical diversity certainly is utterly time-consuming when conducted manually, for example, computational linguistics enable an easy application of methodologically considerably more advanced measures of lexical diversity, resulting in less biased estimates and, thus, higher data quality (McCarthy & Jarvis, 2010; Weiss, 2023). Computational linguistic analysis methods, therefore, are a promising approach for investigating PD effects on teachers' oral language input in science teaching.

3. Research Questions and Hypotheses

Expanding prior findings on a PD approach for language-supportive teaching in elementary school science classes (Gabler et al., 2024; Heppt et al., 2022), the present study focuses on teacher talk in oral classroom interaction.

Specifically, we investigate the following research questions and hypotheses:

(1) Do IG-teachers differ from CG-teachers in the amount and quality (e.g., lexical diversity, lexical elaboration, morphological complexity) of oral language use in elementary school science classrooms after having participated in a PD program for language-supportive teaching in science classrooms? As IG-teachers were intensively trained for using selected language-support strategies in their science classes, including language modeling, we expected them to outperform the CG-teachers in terms of both quantity and quality of language input.

(2) Do these IG-CG-differences in oral language use persist across learning units on two different elementary school science topics (Topic 1: floating and sinking, Topic 2: evaporation and condensation)? Building on prior findings that point to within-group variability in instructional quality across topics (Heppt et al., 2022) and assuming that language demands differ across topics, we do not expect stable IG-CG-differences in teachers' oral language across topics.

4. Method

4.1. Study Design

The present study uses data from the project ProSach (“Professional development training on content-focused language support in elementary school science instruction”; German: “*Professionalisierungsmaßnahmen zur bedeutungsfokussierten Sprachförderung im Sachunterricht der Grundschule*”¹), in which a newly developed PD program for language-supportive science teaching was evaluated in a quasi-experimental field trial in two German federal states (Berlin and Hesse). The study was conducted over two school years and consisted of a PD phase in the first school year (2016–2017) and an implementation phase in the second school year (2017–2018). During the PD phase, teachers from IG and (waiting) CG received PD on the elementary school science topics “floating and sinking” and “evaporation and condensation”, focusing on the content knowledge and pedagogical content knowledge (PCK) needed for teaching these topics in Grade 3 and 4 and familiarizing participants with the lesson plans and teaching material of the learning units². Both PD courses were led by members of the project team who were experienced researchers (PostDocs) and had a background in elementary school science and didactics. The PD courses were conducted face-to-face and had a duration of 5 hours each. They included a range of didactical elements and methods, such as trainer input, hands-on experiments conducted collaboratively in small groups, and plenum discussions.

IG-teachers additionally participated in a comprehensive PD for subject-integrated language support in elementary school science. This PD program was developed along the lines of well-known characteristics of effective PD (e.g., Darling-Hammond et al., 2009; Kalinowski et al., 2020; see Section 2.3), thus combining phases of input with active trials in classroom and subsequent coaching sessions, which included feedback and joint reflection. The PD program was primarily developed and implemented by a member of the

-
- 1 The project was funded by the German Federal Ministry of Education and Research (BMBF) through Grant 01JI1602A, awarded to Humboldt-Universität zu Berlin, Germany (principal investigators: Dr. Sofie Henschel, Prof. Dr. Petra Stanat) and through Grant 01JI1602B awarded to Goethe University Frankfurt, Germany (principal investigator: Prof. Dr. Ilonca Hardy). The authors assume full responsibility of the content of the present publication.
 - 2 A third PD referred to the elementary school science topic “education for sustainable development”. Due to high dropout rates during the PD phase and, in particular, during the implementation phase, this topic could not be considered in the analyses and, thus, is not described in this chapter (cf. Heppt et al., 2022).

project team who was an experienced researcher (PostDoc) and had a strong background in linguistics and language support. The PD program comprised two modules (Module 1: Basics of language scaffolding, Module 2: Coaching and video-feedback) which took place over roughly half a school year.

Module 1 consisted of three 4-hour face-to-face workshops, familiarizing participants with the basic components of the language scaffolding approach. This approach consists of an extensive phase of language-supportive lesson planning (macro-scaffolding), followed by the actual language-supportive classroom teaching (micro-scaffolding). Building on the contents of the lesson unit on “floating and sinking”, IG-teachers, thus, identified the linguistic challenges of the topic and the learning materials (e.g., syntactical structures and vocabulary needed for formulating and justifying hypotheses, domain-specific vocabulary), discussed methods for assessing their students’ language skills, and learned how to define appropriate language-related learning goals (macro-scaffolding). Moreover, they were familiarized with a range of well-established language-support strategies for supporting their students’ language development in classroom teaching (micro-scaffolding). Amongst others, they were encouraged to model their students’ language skills by frequently using core language functions (i.e., hypothesizing, justifying, comparing) and by introducing and explaining important general (e.g., “to verify”, “to assume”) and domain-specific vocabulary (e.g., “water cycle”, “gaseous”), for example by using it in different contexts, contrasting it with words with a different meaning, or juxtaposing morphological derivations (e.g., “to displace” vs. “displacement”; see Section 2.3 for further information on the language-support strategies). Similarly to the PD courses on elementary school science topics, the workshops for subject-oriented language support incorporated a variety of didactical elements, were application-oriented, and fostered participants’ collaboration and reflection.

In the application phase in Module 2, each teacher delivered classroom trials of at least two lessons of the curriculum on “floating and sinking” in their regular elementary school science classrooms. We videotaped one lesson per teacher. The classroom trials were followed by (1) one-on-one feedback with the trainer and (2) a video-feedback session in a small group of participants and with the trainer. These video-feedback sessions had a duration of 2 to 3 hours and aimed at identifying successful examples of language support and at reflecting on possible improvements. To complete the PD phase, IG-teachers participated in a final 3-hour whole-group meeting, summarizing the learnings from the small group sessions (for detailed descriptions of the PD program, see Gabler et al., 2020; Heppt et al., 2022).

During the implementation phase in the second school year, teachers from both groups delivered the learning units on “floating and sinking” and “evaporation and condensation” in their regular Grade 3 or Grade 4 science classrooms. The learning unit on “floating and sinking” comprised six double lessons (90 minutes each), the one on “evaporation and condensation” five double lessons. The second double lesson of both topics was videotaped in each classroom, resulting in two videotaped double lessons per teacher. These videotaped lessons served as a basis for determining the quantity and quality of teachers’ classroom talk. After the implementation phase and, thus, at the beginning of the school year 2018–2019, participants from the (waiting) CG were offered a shortened and optimized version of the PD program for language-supportive teaching.

4.2. Participants

The present analyses are based on data from 27 elementary school teachers from 17 schools who participated in the PD phase and implementation phase of the project. Within each school, all participating teachers were assigned to the same quasi-experimental condition. The IG consisted of 9 teachers from 5 schools ($M_{\text{age}} = 40.33$ years, $SD = 4.04$, 8 female [88.89%], 1 male [11.11%]) and the CG comprised 19 teachers from 8 schools ($M_{\text{age}} = 42.17$ years, $SD = 7.82$, 14 female [73.68%], 5 male [26.32%]). There were no meaningful group differences in age ($t(19) = -0.39$, $p = .700$, $d = -0.25$) and gender distribution ($\chi^2(1) = 0.84$, $p = 0.36$, $\phi < .17$). The groups did not differ in the number of teachers who were trained for teaching in elementary schools ($\chi^2(1) = 2.22$, $p = 0.136$, $\phi < .29$), but more IG-teachers than CG-teachers had attended courses on German as a second language or language support during university teacher training ($\chi^2(1) = 5.03$, $p = 0.025$, $\phi < .45$). Moreover, whereas the IG consisted of teachers from Berlin only, the CG was predominantly located in Hesse ($\chi^2(1) = 20.50$, $p < .001$, $\phi < .86$). This uneven distribution of IG and CG across states resulted from difficulties in the recruiting process and teacher dropout during the PD phase in Berlin. This made it necessary to additionally recruit teachers from Hesse, who, due to time constraints during the ongoing project, could only be assigned to the less time-consuming PD in the CG (for further details on sample attrition and comparability of groups, see Heppt et al., 2022).

Upon PD completion and, thus, before teaching the elementary school science topics to their science classrooms during the implementation phase, IG teachers (as expected) outperformed CG teachers in their knowledge on lan-

guage-supportive teaching ($t(18) = 0.96, p = .174, d = 0.64$). Yet, teachers from both groups were on par in their PCK on floating and sinking ($t(20) = -0.06, p = .951, d = -0.03$) and evaporation and condensation ($t(17) = 0.43, p = .672, d = 0.21$).

The 27 teachers taught 489 elementary school students from Grades 3 ($n = 450$) and 4 ($n = 39$). Of these 489 students, 134 ($M_{\text{age}} = 8.71$ years, $SD = 0.87$, 61 girls [47.30%], 68 boys [50.70%]) were instructed by IG-teachers and 355 ($M_{\text{age}} = 8.40$ years, $SD = 0.56$, 179 girls [51.40%], 169 boys [48.60%]) were instructed by (waiting) CG-teachers. As cross-year teaching was only implemented in Berlin, all Grade 4 students formed part of the IG. Based on data from classroom-level aggregates, IG-students came from lower-SES families than CG-students ($t(10.66) = -1.66, p = .126, d = -0.82$), but there were no pronounced differences regarding the share of multilingual students ($t(24) = -0.90, p = .378, d = -0.38$). CG-students outperformed IG-students in prior knowledge on “floating and sinking” ($t(24) = -1.56, p = .133, d = -0.64$) but not on “evaporation and condensation” ($t(24) = -0.64, p = .53, d = -0.26$). Group differences on the language-related measures were negligible (science vocabulary: $t[24] = -0.57, p = .571, d = -0.24$; general academic vocabulary: $t[24] = -0.54, p = .598, d = -0.24$).

4.3. Assessment of the Amount and Quality of Teachers’ Language Input During Classroom Instruction

We used transcripts of a total of 50 videotaped lessons, 25 on “floating and sinking” ($n_{\text{IG}} = 8, n_{\text{CG}} = 17$) and 25 on “evaporation and condensation” ($n_{\text{IG}} = 8, n_{\text{CG}} = 17$), as a basis for assessing the amount and quality of teachers’ language input³. The transcripts cover 45 minutes of each double lesson, focusing on whole-classroom interaction. In particular, these were the introductory sequence (activation of prior knowledge), the instructional sequence (presentation and explanation of experiments for group work), and the reflection sequence at the end of the lesson (discussion and explanation of observations).

Teacher utterances were subsequently automatically coded regarding a variety of linguistic complexity features using the software *Common Text Analysis Platform* (CTAP; Chen & Meurers, 2016; Weiss & Meurers, 2022), a fully web-based, freely available and extensively researched platform for linguistic complexity analyses. Built on a Java Framework (cf. Chen & Meurers, 2016 for a detailed technical description of the architecture), CTAP allows for

3 Not all of the 27 teachers participated in the video-recordings of both double lessons. This means that the teacher samples for the two science topics are not exactly identical.

broad linguistic complexity modeling using a Natural Language Processing (NLP) pipeline that segments texts into sentences, tokens, characters, and syllables and calculates various theory-based measures of linguistic complexity. This automated extraction process allows to capture a large set of measures across different linguistic dimensions, including lexicon, syntax, morphology, discourse, language use and human processing – areas where manual analyses are often limited due to the high cost of implementation. Amongst others, CTAP has been used to characterize academic language development (Weiss & Meurers, 2019) and is currently the most extensive feature set for German complexity assessment. Of the 543 linguistic complexity features currently implemented in CTAP, we selected 18 features that matched the linguistic focus of our intervention. As shown in detail in Table 1, the majority of features considered in our analyses referred to teachers' vocabulary use (e.g., lexical diversity, lexical sophistication, morphological complexity), given the lesson focus on the introduction and use of new concepts and adequate terminology. Both videotaped lessons included active experimentation and the formulation and justification of hypotheses. With regard to syntactical features, we therefore considered teachers' use of causal connectives and, as these should lead to longer sentences, average sentence length.

4.4. Analytical Procedure

We used *t*-tests for comparing IC-CG-differences in teachers' oral language use. While normal distribution within samples is typically considered a precondition for conducting parametric tests, *t*-tests have been shown to be robust against violations of the normality assumption (e.g., Kubinger et al., 2009). In the present sample, almost all features were normally distributed within IG and CG for both topics. Within-group homoscedasticity was given for all features for Topic 1 and for all but two features for Topic 2. In these cases (i.e., number of tokens, lexical diversity as measured by the HDD), we report results for the *Welch*-test. Additional analyses based on the nonparametric Mann-Whitney-U-Test yielded very similar result patterns as those presented below. We used the effect size *d* for evaluating the practical relevance of our findings, interpreting 0.20 as a small, 0.50 as a medium, and 0.80 as a large effect (Cohen, 1988).

Teachers might adapt their oral language input to their students' prior knowledge and language proficiency. Therefore, we additionally investigated correlations (not displayed here) between the 18 features on the amount and complexity of teachers' language input in both topics and various characteris-

tics of classroom composition (i.e., prior knowledge on the topics “floating and sinking” and “evaporation and condensation”, mastery of science vocabulary and general academic vocabulary, share of multilingual students in class, classroom average of family SES). Yet, hardly any significant correlations emerged and there was no systematic pattern across topics. As there was, thus, no clear indication that teachers’ language input covaried with their students’ prior knowledge or the sociodemographic classroom composition, we refrained from controlling for any of these features in subsequent analyses. Moreover, students’ oral language use and related features (e.g., length of student-teacher interactions in whole-classroom interactions) were not considered in the present analyses. However, most of the selected features are not directly impacted by the length of the analyzed utterances (see Table 1).

Table 1: Linguistic Features and Their Alignment With the Intervention

#	Feature Set	Description and Selected Features	Alignment With the Intervention
1	Surface features	Overall amount of words and sentences spoken by the teachers, i.e., (1) number of tokens (total number of words) and (2) number of types (number of unique words), (3) number of sentences, (4) average sentence length	Increasing the amount of (high-quality) language input by using strategies such as thinking-aloud techniques or mapping actions with language, acting as a language role model by using adequate sentence structures for hypothesizing and justifying
2	Lexical diversity	Variability of teachers’ vocabulary as measured by indicators that most efficiently control for average text length, i.e., (5) <i>Hypergeometric Distribution Diversity</i> (HD-D; McCarthy & Jarvis, 2010) and (6) <i>Measure of Textual Lexical Diversity</i> (MTLD; McCarthy & Jarvis, 2010)	Definition, use, and consolidation of domain-specific vocabulary such as “to displace”, “water cycle”, or “gaseous”
3	Lexical density	Share of selected parts of speech, e.g., interrogative pronouns, nouns, and verbs, per overall tokens in the text (7–12)	As language role models, teachers should increasingly use academic language which is characterized by a high information density (e.g., Heppt & Schröter, 2023); asking language-supportive questions was one of the core language-support strategies targeted in the PD and should come along with a higher density of interrogative pronouns
4	Lexical elaboration	Lexical sophistication as reflected in age of active use (with more sophisticated vocabulary being actively used at an older age). Measures were extracted based on the corpus <i>Karlsruhe Children’s Texts</i> (13–14; Lavalley et al., 2015; Weiss, 2023). While such corpora can be used for determining the age at which specific words are actively used, we use them for assessing the overall lexical sophistication of a teacher’s vocabulary without identifying the exact words the teacher used	Definition, use, and consolidation of domain-specific vocabulary, which is likely to occur infrequently in everyday conversations and should, thus, be primarily acquired in instructional settings (Heppt & Schröter, 2023)
5	Cohesion (discourse)	Use of causal connectives, measured by (15) the overall number of connectives and (16) the number of connectives per token	Acting as a language role model in formulating justifications and explanations
6	Morphological complexity	Measures assessing (17) the share of tokens and (18) the share of types with at least three syllables	Definition, use, and consolidation of domain-specific vocabulary, which frequently consists of compounds or derivations, resulting in relatively long and morphologically complex words (Bailey et al., 2007; Köhne et al., 2015)

5. Results

We examined group differences in IG and CG teachers' amount and quality of language input during science teaching separately for Topic 1 ("floating and sinking") and Topic 2 ("evaporation and condensation"). The descriptive statistics, the statistics of the *t*-tests as well as the effect size *d* for evaluating the practical relevance of the effects are displayed in Tables 2 and 3.

For "floating and sinking", findings indicate that, overall, IG-teachers provided more language input than CG-teachers, resulting in large effect sizes (see Table 2). However, contrary to our expectations, CG teachers produced longer sentences than teachers from the IG. Regarding the selected features of lexical complexity, only sporadic group differences emerged, some of them in favor of the IG, others in favor of the CG. Specifically, teachers from IG and CG did not differ in the lexical diversity used in oral classroom talk. While we observed a higher density of verbs in the IG and a higher density of adverbs in the CG, no differences emerged for any of the other density features, including interrogative pronouns and content words. There is, thus, no clear indication that teachers' classroom talk was characterized by a higher information density in the IG than in the CG. Indicators used for assessing lexical elaboration do not point to systematic differences between IG and CG either.

In line with our assumptions, IG-teachers used more causal connectives. This large effect can mainly be explained by the overall larger amount of teacher talk in the IG, as the group differences do not persist when considering the number of tokens. Finally, compared to the CG, IG-teachers' classroom talk was characterized by a higher morphological complexity, as reflected in the higher share of tokens and types with at least three syllables.

In a next step, we compared these findings with the results for Topic 2 with the aim of exploring whether group differences in oral language input persist across topics. Investigating the results for Topic 2 revealed three following major findings (Table 3). First, when teaching the double lesson on "evaporation and condensation", IG and CG teachers did not differ in the overall amount of speech delivered. This finding is due to an increase of language input in the CG (difference in the number of tokens: $t(14) = -4.57$, $p < .001$, $d = 1.33$). In line with this finding, no group differences emerged in the number of causal connectives used in oral classroom talk. Second, the pattern of results remained largely stable for lexical complexity. Third, and in line with the findings for Topic 1, IG teachers used a higher share of morphologically more complex (i.e., longer) words.

Table 2: Descriptive and Inferential Statistics for Comparisons Between IG ($n = 8$) and CG ($n = 17$) in Language Input During Topic 1

Feature Set	#	Feature	IG		CG		t	df	p	d	Evaluation of IG-CG differences
Surface Features	1	Number of tokens	3299.38	567.34	2455.29	471.69	3.92	23	<.001***	1.68	IG > CG
	2	Number of types	685.25	101.48	566.65	82.50	3.12	23	.002**	1.34	IG > CG
	3	Number of sentences	602.25	152.83	431.18	139.73	2.77	23	.005**	1.19	IG > CG
	4	Average sentence length	6.17	0.26	6.35	0.44	-1.05	23	.153	-0.46	IG < CG
Lexical diversity	5	HD-D	0.85	0.01	0.85	0.01	0.28	23	.392	0	IG = CG
	6	MTLD	58.29	7.39	58.15	7.96	0.04	23	.484	0.02	IG = CG
Lexical density	7	Interrogative pronouns	0.03	0.01	0.03	0.01	1.17	23	.126	0	no systematic IG-CG differences
	8	Adjectives	0.05	0.01	0.05	0.01	0.27	23	.359	0	
	9	Adverbs	0.16	0.02	0.18	0.02	-2.21	23	.019*	-1.00	
	10	Content words	0.51	0.02	0.51	0.02	-0.69	23	.248	0	
	11	Nouns	0.15	0.01	0.15	0.02	0.35	23	.365	0	
	12	Verbs	0.22	0.01	0.21	0.01	2.50	23	.010*	1.00	
	13	Mean age at active use (content words, tokens)	10.33	0.07	10.35	0.09	-0.61	23	.275	-0.24	IG = CG
Lexical elaboration	14	Mean age at active use (content words, types)	10.38	0.06	10.38	0.08	-0.11	23	.454	0	IG = CG
	15	Number of causal connectives	56.63	17.02	37.82	14.93	2.81	23	.005**	1.21	IG > CG
Cohesion (dis-course)	16	Number of causal connectives per token	0.02	0.004	0.02	0.01	0.84	23	.205	0	IG = CG
	17	Share of tokens with at least 3 syllables	0.36	0.02	0.35	0.02	1.10	23	.141	0.50	IG > CG
	18	Share of types with at least three syllables	0.61	0.03	0.58	0.03	2.16	23	.021	1.00	IG > CG

Note. Group differences of at least medium effect size in favor of the IG are printed in **bold**. Group differences of at least medium effect size in favor of the CG are printed in *italics*. HD-D = Hypergeometric Distribution Diversity; MTLD = Measure of Textual Lexical Diversity.
 * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3: Descriptive and Inferential Statistics for Comparisons Between IG ($n = 8$) and CG ($n = 17$) in Language Input during Topic 2

Feature Set	#	Feature	IG		CG		t	df	p	d	Evaluation of IG-CG-differences
			M	SD	M	SD					
Surface features	1	Number of tokens	3124.75	390.44	3196.35	694.18	-0.33	22.01	.373	-0.12	
	2	Number of types	713.25	69.38	689.00	105.77	0.59	23	.281	0.25	IG = CG
	3	Number of sentences	463.25	107.36	464.18	120.65	-0.02	23	.493	-0.01	
	4	Average sentence length	6.49	0.46	6.60	0.38	-0.60	23	.279	-0.27	IG = CG
Lexical diversity	5	HD-D	0.86	0.01	0.86	0.01	0.09	8.91	.466	0	IG = CG
	6	MTLD	63.40	11.54	62.63	7.19	0.20	23	.421	0.09	
Lexical density	7	Interrogative pronouns	0.03	0.01	0.03	0.01	-0.01	23	.496	0	no systematic IG-CG-differences
	8	Adjectives	0.06	0.01	0.06	0.01	-0.23	23	.410	0	
	9	Adverbs	0.17	0.02	0.18	0.02	-0.78	23	.221	-0.50	
	10	Content words	0.52	0.02	0.51	0.02	0.84	23	.204	0.50	
	11	Nouns	0.14	0.01	0.14	0.02	0.81	23	.212	0	
	12	Verbs	0.22	0.01	0.21	0.01	2.37	23	.013*	1.00	
Lexical elaboration	13	Mean age at active use (content words, tokens)	10.35	0.07	10.35	0.07	0.01	23	.498	0	IG = CG
	14	Mean age at active use (content words, types)	10.44	0.11	10.41	0.09	0.63	23	.269	0.31	
Cohesion (dis-course)	15	Number of causal connectives	49.25	20.18	49.53	15.10	-0.04	23	.485	-0.02	IG = CG
	16	Number of causal connectives per token	0.02	0.01	0.02	0.01	-0.08	23	.468	0	
Morphological complexity	17	Share of tokens with at least 3 syllables	0.36	0.03	0.34	0.02	1.64	23	.057†	0.85	IG > CG
	18	Share of types with at least three syllables	0.61	0.03	0.59	0.04	1.16	23	.129	0.54	

Note. Group differences of at least medium effect size in favor of the IG are printed in **bold**. Group differences of at least medium effect size in favor of the CG are printed in *italics*.

† $p < .10$. * $p < .05$.

6. Discussion

Delivering classroom instruction that benefits all students and that fosters both conceptual understanding and language proficiency is highly demanding. There is, thus, a need for effective PD for helping teachers acquire the necessary skills. Drawing on a PD approach that has been shown to promote teachers' knowledge on subject-integrated language-support and benefits their language-supportive classroom teaching (Gabler et al., 2024; Heppt et al., 2022), the present study investigated teachers' oral language use in regular elementary school science classrooms. Specifically, using CTAP (Chen & Meurers, 2016; Weiss & Meurers, 2022), we investigated (1) whether teachers who had participated in a PD program on language-supportive teaching provided their students with quantitatively and qualitatively more stimulating language input than did teachers from the (waiting) CG and (2) whether these effects were sustained over time. For Topic 1 on "floating and sinking", we found that IG teachers talked more than CG teachers during classroom teaching and that they used more causal connectives. Moreover, their language input was characterized by higher morphological complexity, as indicated by the share of words with more than three syllables, and this effect persisted irrespective of the overall amount of language input. For the other features of linguistic complexity, however, there were either no group differences (i.e., lexical diversity) or the result pattern was very heterogeneous and did not allow for clear conclusions (i.e., lexical density and lexical elaboration). The result pattern for Topic 2 on "evaporation and condensation" was similar, except for the overall amount of language input. As CG teachers increased the amount of language input from Topic 1 to Topic 2, no group differences emerged in the amount of language input and the use of causal connectives. Overall, considering both topics, only a small number of features point to more sophisticated and complex language use in the IG.

The study adds to the small body of research focusing on teachers' amount and sophistication of language input when investigating the effectiveness of PD for language-supportive teaching. Prior research reported strong positive effects of a 3-hour PD on teachers' lexical diversity, measured in the number of word types, and the use of general and domain-specific academic vocabulary (Henrichs & Leseman, 2014). This study used very specific definitions of what was rated as general and specific vocabulary based on detailed word lists. Another intervention study that was conducted over the course of several months and included video feedback sessions, increased teachers' use of syntactically complex sentences and their lexical sophistication, as indicated in the overall word frequency (van Dijk et al., 2019). In the present study, IG

and CG classroom talk differed in just a few linguistic features. Yet, whereas previous studies typically considered only very few language features, we included six feature sets with a total of 18 features.

The finding that IG teachers provided more language input during Topic 1, along with a more frequent use of causal connectives and an enhanced use of morphologically complex words, is basically in line with our assumptions. IG teachers were trained for the use of the language-support strategy language modeling, incorporating an increase of language input by thinking-aloud techniques and mapping their own or their students' actions with language. Acting as language role models, IG teachers deliberately aimed at performing important language functions of scientific inquiry, such as justifying their hypotheses and conclusions, which requires the use of causal connectives (e.g., Gabler et al., 2020). Their more frequent use of long and morphologically complex words, which persisted across both topics, matches the nature of domain-specific and general academic vocabulary. Such terms have repeatedly been identified to be relatively long (Bailey et al., 2007), in particular in the German language, where the use of compounds and verbs with prefixes is quite common (Köhne et al., 2015).

In explaining the overall small amount of systematic IG-CG-differences, which were still more pronounced in Topic 1 than in Topic 2, several aspects need to be considered. First, the PD on subject-integrated language support drew on the lesson plans and teaching materials of Topic 1. It required teachers to teach at least two lessons on “floating and sinking” in their Grade 3 classrooms, while actively implementing core language-support strategies such as language modeling. The implementation of language-supportive behavior was further supported by detailed suggestions for language support that were included in the lesson plans for Topic 1 in the IG. Feedback and video-coaching sessions were also based on teachers' experiences while teaching the topic “floating and sinking”. IG teachers, thus, had ample opportunity to familiarize with Topic 1 and to plan and use language-support strategies while teaching this topic. This probably helped them to integrate strategies such as thinking-aloud techniques or the use of core domain-specific vocabulary in different contexts and contrastive ways, resulting in an enhanced language input during Topic 1. Topic 2 required IG teachers to transfer their knowledge and skills without prior classroom trials and without receiving suggestions for language support in the lesson plans, possibly resulting in reduced IG-CG-differences (Heppt et al., 2022).

Second, with the aim of ensuring a high level of comparability across intervention groups, teachers from both groups taught exactly the same lessons, using detailed lesson plans and accompanying teaching materials. It

can therefore be assumed that teachers' language use was, at least to a certain degree, predetermined by the lesson content. In particular, academic vocabulary needed for co-constructing knowledge on core concepts and for verbalizing assumptions and observations, occurred in lesson plans and work sheets and had to be used by teachers in IG and CG alike. In a similar line of reasoning, it should be noted that language use across topics surely does not only depend on PD participation but also on differences that are inherent to the topics. Thus, vocabulary like "metal", or "plastic", that plays an important role for "floating and sinking", is more frequent and, thus, less elaborate than vocabulary such as "to rise", or "teapot warmer", that occurs in the lesson unit on "evaporation and condensation".

Third, teachers from both groups participated in PD courses on the two elementary school science topics, suggesting that both IG- and (waiting) CG-teachers were well-prepared for providing high-quality science instruction and were motivated to engage in teacher PD. As previous research demonstrated a positive link between certain teacher characteristics, such as occupational motivation, constructivist beliefs, and content knowledge, and the uptake of teacher PD (Richter et al., 2021), it is possible that teachers of both groups had better prerequisites for delivering cognitively activating and language-supportive instruction than the average teacher has. Possibly, group differences in language use would have been more pronounced if CG teachers had delivered the elementary school science classes based on the curricula and teaching materials only, without participating in respective PD courses.

6.1. Limitations and Conclusion

A major limitation of our study is that it is based on only a small convenience sample of teachers from two German federal states. This resulted in limited test power and prevents generalization of results to the larger population of elementary school teachers. Moreover, data on teachers' instructional practice and, hence, their oral language use in classroom instruction have only been collected after PD completion. Consequently, it is not possible for us to investigate changes in the quantity and quality of teachers' language input after participating in our PD program. Yet, previous analyses with the same sample showed that IG-teachers' knowledge on subject-integrated language support increased after PD completion and that they outperformed their counterparts from the CG (Heppt et al., 2022). It therefore seems reasonable to assume that differences in teachers' classroom talk are, at least partly, attributable to different treatment conditions.

The linguistic complexity features used for comparing the quasi-experimental groups were chosen with great care to yield maximum alignment with the intervention. Although word length is a widely established indicator of morphologic complexity (e.g., Bailey et al., 2007; Köhne et al., 2015), one could still argue that it is a rather simplistic feature. A more direct assessment of different forms of derivations, such as nominalizations or compounds, might prove even more insightful. Along similar lines, for gaining a better understanding of teachers' language-supportive instructional practice, it might be helpful to not only learn to what degree teachers use long words or causal connectives but rather which words they use and in which instructional contexts and phases of instruction.

Despite these shortcomings, we believe that using computational linguistic analysis methods for investigating PD effects offers valuable insights. In particular, the results of the present study complement and validate prior findings on the same PD approach, pinpointing that differences between the intervention groups are most pronounced with Topic 1 and get attenuated with Topic 2 – a topic which was less familiar to IG teachers in terms of material and respective linguistic scaffolding. Using computational linguistic analysis methods also expands on prior findings. Whereas IG teachers' science instruction on "floating and sinking" has previously been shown to be of higher language-supportive quality, relying on a highly inferential global rating (Heppt et al., 2022), the fine-grained linguistic analyses yielded through computational linguistic analyses gives hints on the language features that might have contributed to these differences. In order to gain an even more comprehensive understanding of teachers' amount and complexity of oral language, it would be highly valuable for future research to also consider students' language use, ideally also including the number and length of teacher-student as well as teacher-whole classroom interactions. In our study, there were no measurable covariations between students' prior knowledge, indicated by performance on achievement tests, and teachers' language use. Considering student-teacher interactions would allow for further insights as to whether teachers adapt their language input to their students' knowledge and language proficiency, which is an important prerequisite for fostering learning.

References

- Babinski, L. M., Amendum, S. J., Knotek, S. E., Sánchez, M., & Malone, P. (2018). Improving young English learners' language and literacy skills through teacher professional development: A randomized controlled trial. *American Educational Research Journal*, 55(1), 117–143. <https://doi.org/10.3102/0002831217732335>

- Bailey, A. L., Butler, F. A., Stevens, R., & Lord, C. (2007). Further specifying the language demands of school. In A. L. Bailey (Ed.), *The language demands of school. Putting academic English to the test* (pp. 103–156). Yale University Press.
- Becker-Mrotzek, M., & Woerfel, T. (2020). Sprachsensibler Unterricht und Deutsch als Zweit-sprache als Gegenstand der Lehrerbildung [Language-supportive teaching and German as a second language as a topic for teacher education]. In C. Cramer, J. König, M. Rothland, & S. Blömeke (Eds.), *Handbuch Lehrerinnen- und Lehrerbildung* (pp. 98–104). Klinkhardt. <https://doi.org/10.35468/hblb2020-011>
- Bravo, M. A., & Cervetti, G. N. (2014). Attending to the language and literacy needs of English learners in science. *Equity & Excellence in Education*, 47(2), 230–245. <https://doi.org/10.1080/10665684.2014.900418>
- Chen, X., & Meurers, D. (2016). CTAP: A web-based tool supporting automatic complexity analysis. *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity*, 113–119. COLING.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession*. National Staff Development Council. <https://edpolicy.stanford.edu/sites/default/files/publications/professional-learning-learning-profession-status-report-teacher-development-us-and-abroad.pdf>
- Dickinson, D. K., & Porche, M. V. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development*, 82(3), 870–886. <https://doi.org/10.1111/j.1467-8624.2011.01576.x>
- Dokter, N., Aarts, R., Kurvers, J., Ros, A., & Kroon, S. (2017). Academic language in elementary school mathematics: Academicness of teacher input during whole class instruction. *Dutch Journal of Applied Linguistics*, 6(2), 213–230. <https://doi.org/10.1075/dujal.17007.dok>
- Ernst-Slavit, G., & Mason, M. R. (2011). “Words that hold us up:” Teacher talk and academic language in five upper elementary classrooms. *Linguistics and Education*, 22(4), 430–440. <https://doi.org/10.1016/j.linged.2011.04.004>
- Gabler, K., Henschel, S., Hardy, I., Sontag, C., & Heppt, B. (2024). Sprachförderliches Unterrichten im Sachunterricht der Grundschule [Language-supportive teaching in elementary school science]. *Zeitschrift für Grundschulforschung*, 17, 87–110. <https://doi.org/10.1007/s42278-024-00191-8>
- Gabler, K., Mannel, S., Hardy, I., Henschel, S., Heppt, B., Hettmannsperger-Lippolt, R., Sontag, C., & Stanat, P. (2020). Fachintegrierte Sprachförderung im Sachunterricht der Grundschule: Entwicklung, Erprobung und Evaluation eines Fortbildungskonzepts auf der Grundlage des Scaffolding-Ansatzes [Subject-integrated language support in elementary school subject classes: Development, testing, and evaluation of a training concept based on the scaffolding approach]. In C. Titz, S. Weber, H. Wagner, A. Ropeter, C. Geyer, & M. Hasselhorn (Eds.), *Sprach- und Schriftsprachförderung wirksam gestalten: Innovative Konzepte und Forschungsimpulse* (pp. 59–83). Kohlhammer.
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning. Teaching second language learners in the mainstream classroom*. Heinemann.
- Hart, B., & Risely, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes.
- Henrichs, L. F., & Leseman, P. P. M. (2014). Early science instruction and academic language development can go hand in hand. The promising effects of a low-intensity teacher-focused intervention. *International Journal of Science Education*, 36(17), 2978–2995. <https://doi.org/10.1080/09500693.2014.948944>

- Henschel, S., & Heppt, B. (2024). *Kompetenzeinschätzungen von Grundschullehrkräften zum sprachbildenden Unterrichten im Fach Mathematik: Zusammenhänge mit Merkmalen der Aus- und Fortbildung, Unterrichtsaktivitäten und mathematischen Kompetenzen der Schüler:innen* [Primary school teachers' self-perceived competencies in language-supportive mathematics instruction: relationships with characteristics of teacher education and professional development, teaching activities, and students' mathematical competencies] [Manuscript submitted for publication]. Institut zur Qualitätsentwicklung im Bildungswesen (IQB).
- Henschel, S., Heppt, B., Rjosk, C., & Weirich, S. (2022). Zuwanderungsbezogene Disparitäten [Immigration-related disparities]. In P. Stanat, S. Schipolowski, R. Schneider, K. Sachse, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (pp. 181–219). Waxmann.
- Heppt, B., Henschel, S., Hardy, I., Hettmannsperger-Lippolt, R., Gabler, K., Sontag, C., Mannel, S., & Stanat, P. (2022). Professional development for language support in science classrooms: Evaluating effects for elementary school teachers. *Teaching and Teacher Education*, 109, Article 103518. <https://doi.org/10.1016/j.tate.2021.103518>
- Heppt, B., & Schröter, P. (2023). Bildungssprache als übergeordnetes Ziel sprachlicher Bildung [Academic language as superordinate goal of language education]. In M. Becker-Mrotzek, I. Gogolin, H.-J. Roth, & P. Stanat (Eds.), *Grundlagen der sprachlichen Bildung* (pp. 139–153). Waxmann.
- Kalinowski, E., Egert, F., Gronostaj, A., & Vock, M. (2020). Professional development on fostering students' academic language proficiency across the curriculum – A meta-analysis of its impact on teachers' cognition and teaching practices. *Teaching and Teacher Education*, 88, Article 102971. <https://doi.org/10.1016/j.tate.2019.102971>
- Kane, C., Sandilos, L., Hammer, C. S., Komaroff, E., Bitetti, D., & López, L. (2023). Teacher language quality in preschool classrooms: Examining associations with DLLs' oral language skills. *Early Childhood Research Quarterly*, 63, 352–361. <https://doi.org/10.1016/j.ecresq.2023.01.006>
- Köhne, J., Kronenwerth, S., Redder, A., Schuth, E., & Weinert, S. (2015). Bildungssprachlicher Wortschatz – linguistische und psychologische Fundierung und Itementwicklung [Academic vocabulary – linguistic and psychological foundation and item development]. In A. Redder, J. Naumann, & R. Tracy (Eds.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS) – Ergebnisse* (pp. 67–92). Waxmann.
- Kubinger, K. D., Rasch, D., & Moder, K. (2009). Zur Legende der Voraussetzungen des t-Tests für unabhängige Stichproben [On the legend of the preconditions of the t-test for independent samples]. *Psychologische Rundschau*, 60(1), 26–27. <https://doi.org/10.1026/0033-3042.60.1.26>
- Lavalley, R., Berkling, K., & Stüker, S. (2015). Preparing children's writing database for automated processing. *Proceedings of the L1 Teaching, Learning and Technology* (pp. 9–15). <https://isl.iar.kit.edu/pdf/Lavalley2015.pdf>
- Lipowsky, F., & Rzejak, D. (2015). Key features of effective professional development programmes for teachers. *RICERCAZIONE*, 7(2), 27–51.
- Llosa, L., Lee, O., Jiang, F., Haas, A., O'Connor, C., Van Booven, C. D., & Kieffer, M. J. (2016). Impact of a large-scale science intervention focused on English language learners. *American Educational Research Journal*, 53(2), 395–424. <https://doi.org/10.3102/0002831216637348>
- Ludewig, U., Schlitter, T., Lorenz, R., Kleinkorres, R., Schaufelberger, R., Frey, A., & McElvany, N. (2022). *Die COVID-19 Pandemie und Lesekompetenz von Viertklässler*innen. Ergebnisse der IFS-Schulpanelstudie 2016–2021* [The COVID-19 pandemic and reading achievement in Grade 4. Results of the IFS school panel study 2016–2021]. https://ifs.ep.tu-dortmund.de/storages/ifs-ep/r/Downloads_allgemein/IFS_Schulpanelstudie.pdf [access on Januar 3, 2023]

- Mahan, K. R. (2020). The comprehending teacher: Scaffolding in content and language integrated learning (CLIL). *The Language Learning Journal*, 50(1), 74–88. <https://doi.org/10.1080/09571736.2019.1705879>
- McCarthy, P. M., & Jarvis, S. (2010). MTLT, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392. <https://doi.org/10.3758/BRM.42.2.381>
- Paetsch, J., & Heppt, B. (2021). Teaching German as a second language: Implications for teacher education in Germany. In J. Chin-Kin Lee & T. Ehmke (Eds.), *Quality in teacher education and professional development: Issues and prospects in Chinese and German systems* (pp. 253–267). Routledge.
- Piwowar, V., Thiel, F., & Ophardt, D. (2013). Training inservice teachers' competencies in classroom management. A quasi-experimental study with teachers of secondary schools. *Teaching and Teacher Education*, 30, 1–12. <https://doi.org/10.1016/j.tate.2012.09.007>
- Prediger, S., & Hardy, I. (2023). Fachliches und sprachliches Lernen im mathematisch-naturwissenschaftlichen Unterricht [Content- and language-related learning in mathematics and science instruction]. In M. Becker-Mrotzek, I. Gogolin, H.-J. Roth, & P. Stanat (Eds.), *Grundlagen der sprachlichen Bildung* (pp. 171–184). Waxmann.
- Prediger, S., & Neugebauer, P. (2021). 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., & Wessel, L. (2017). Brauchen mehrsprachige Jugendliche eine andere fach- und sprachintegrierte Förderung als einsprachige? [Do multilingual students require other kinds of content- and language integrated interventions than monolingual students?]. *Zeitschrift für Erziehungswissenschaft*, 21(2), 361–382. <https://doi.org/10.1007/s11618-017-0785-8>
- Richter, E., Kunter, M., Marx, A. & Richter, D. (2021). Who participates in content-focused teacher professional development? Evidence from a large scale study. *Frontiers in Education*, 6, Article 722169. <https://doi.org/10.3389/educ.2021.722169>
- Sachse, K. A., Jindra, C., Schumann, K., & Schipolowski, S. (2022). Soziale Disparitäten [Social disparities]. In P. Stanat, S. Schipolowski, R. Schneider, K. Sachse, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (pp. 151–180). Waxmann.
- Studhalter, U. T., Leuchter, M., Tettenborn, A., Elmer, A., Edelsbrunner, P. A., & Saalbach, H. (2021). Early science learning: The effects of teacher talk. *Learning and Instruction*, 71, Article 101371. <https://doi.org/10.1016/j.learninstruc.2020.101371>
- van Dijk, M., Menninga, A., Steenbeek, H., & van Geert, P. (2019). Improving language use in early elementary science lessons by using a video feedback intervention for teachers. *Educational Research and Evaluation*, 25(5–6), 299–322. <https://doi.org/10.1080/13803611.2020.1734472>
- Volodina, A., Heppt, B., & Weinert, S. (2021). Effects of socioeconomic status and language use on academic language proficiency in children with a migration background: An evaluation using quantile regressions. *Contemporary Educational Psychology*, 65, Article 101973. <https://doi.org/10.1016/j.cedpsych.2021.101973>
- Vorholzer, A., & Aufschnaiter, C. (2019). Guidance in inquiry-based instruction – An attempt to disentangle a manifold construct. *International Journal of Science Education*, 41(11), 1562–1577. <https://doi.org/10.1080/09500693.2019.1616124>
- Weiss, Z. (2023). *An integrative approach to linguistic complexity analysis for German* [Doctoral dissertation, University of Tübingen]. University Library Publication System Tübingen. <https://doi.org/10.15496/publikation-93806>

- Weiss, Z., Lange-Schubert, K., Geist, B., & Meurers, D. (2022). Sprachliche Komplexität im Unterricht [Linguistic complexity in the classroom]. *Zeitschrift für Germanistische Linguistik*, 50(1), 159–201. <https://doi.org/10.1515/zgl-2022-2052>
- Weiss, Z., & Meurers, D. (2019). Analyzing linguistic complexity and accuracy in academic language development of German across elementary and secondary school. *Proceedings of the 14th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 380–393). <https://doi.org/10.18653/v1/W19-4440>
- Weiss, Z., & Meurers, D. (2022). Assessing sentence readability for German language learners with broad linguistic modeling or readability formulas: When do linguistic insights make a difference? *Proceedings of the 17th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2022)*, 141–153. Association for Computational Linguistics.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>

Enthusiastische Lehrkräfte, motivierte Lernende? Eine differenzielle Betrachtung der Verbindung von Unterrichts- enthusiasmus und intrinsischer Motivation in heterogenen Lerngruppen

Anne Kosubek, Hanna Gaspard, Hannes Schröter &
Fani Lauermann

Zusammenfassung

Der Unterrichtsenthusiasmus von Lehrkräften zeigt häufig positive Zusammenhänge mit der intrinsischen Motivation von Lernenden, wobei eine Generalisierung für sprachlich-heterogene Lerngruppen noch aussteht. Diese Studie untersucht daher den Zusammenhang zwischen dem affektiven Erleben und dem emotionalen Ausdruck des Unterrichtsenthusiasmus von Lehrkräften sowie deren Zusammenhang mit der intrinsischen Motivation von Lernenden in heterogenen Lerngruppen auf Klassenebene. Der ausgedrückte Unterrichtsenthusiasmus wurde anhand standardisierter Beobachtungen von Videosequenzen erfasst, während Lehrkräfte ihren erlebten Unterrichtsenthusiasmus und Lernende ihre intrinsische Motivation selbst einschätzten. Insgesamt wurden Daten von $n = 33$ Lehrkräften und ihren $n = 314$ Lernenden in Vorbereitungsklassen (Regelschulen) sowie von $n = 35$ Lehrkräften und ihren $n = 377$ Lernenden in Integrationskursen (Erwachsenenbildung) analysiert. Hierarchisch-lineare Regressionsmodelle zeigten lediglich in den Integrationskursen Zusammenhänge zwischen dem affektiven Erleben des Unterrichtsenthusiasmus der Lehrkräfte und der intrinsischen Motivation der Lernenden. Mögliche Gründe für diese differenziellen Zusammenhänge beispielsweise durch die unterschiedliche Altersstruktur der Lerngruppen werden diskutiert.

Schlagworte: Unterrichtsenthusiasmus von Lehrkräften, intrinsische Motivation von Lernenden, DaZ, sprachlich-heterogene Lerngruppen, hierarchisch-lineare Regression

Abstract

Teachers' enthusiasm for teaching often seems to be positively associated with desirable learning outcomes, such as students' intrinsic motivation. However, these findings are yet to be generalized for heterogeneous learning groups. Accordingly, the present study investigated the associations between teachers' affective experiences and the observable emotional expressivity of teachers' enthusiasm for teaching, aggregated at the classroom level, as well as the associations of both components of teachers' enthusiasm with students' class-level intrinsic motivation in heterogeneous learning groups. The expressivity of enthusiasm for teaching was assessed using standardized observations of video sequences, teachers reported their experienced enthusiasm for teaching, and students reported their intrinsic motivation. In total, data from $n = 33$ teachers and their $n = 314$ students in German-as-a-second-language classrooms in secondary schools and $n = 35$ teachers and their $n = 377$ students in adult education settings were analyzed. Hierarchical linear regression models showed significant associations between teachers' affective experience of teaching enthusiasm and students' intrinsic motivation only in adult education. Possible reasons for these differential correlations, for example due to different students' age composition, are discussed.

Keywords: Teacher enthusiasm for teaching, student intrinsic motivation, German-as-a-second-language, linguistically heterogeneous students, hierarchical-linear regression

1. Einleitung

Der Enthusiasmus von Lehrkräften für das Unterrichten steht in positivem Zusammenhang mit wünschenswerten Ergebnissen von Lernenden im Lernprozess (siehe Keller et al., 2016; Rosenshine, 1970a). Lernende von enthusiastischeren Lehrkräften sind demnach stärker motiviert (z. B. Patrick et al., 2000) und berichten meist von mehr Lernfreude (z. B. Frenzel et al., 2009), stärkerem Interesse (z. B. Fauth et al., 2019; Keller et al., 2014a) sowie mehr Motivation, die mit der Beziehung zur Lehrperson zusammenhängt (affiliative motivation; Kim & Schallert, 2014).

Somit nimmt der Unterrichtsenthusiasmus von Lehrkräften eine Schlüsselfunktion für effektive Lehr-Lernprozesse (Schoofs, 2019) und die Motivation von Lernenden ein (Sekundarstufe I: Kunter & Holzberger, 2014; College: Patrick et al., 2000). Insbesondere vor dem Hintergrund der Förderung der intrinsischen Motivation von Lernenden und lebenslangem Lernen als wesentlichem Ziel in Lern- und Bildungsprozessen erscheint der Unterrichtsenthusiasmus von Lehrkräften somit als wichtiger Einflussfaktor.

Allerdings steht eine Generalisierung dieser Erkenntnisse für sprachlich-heterogene Lerngruppen noch aus (Hachfeld et al., 2012; Keller et al., 2016). Angesichts der heutigen Sprachvielfalt von Lerngruppen (Becker-Mrotzek, 2022) gewinnt ein enthusiastisches Lehrkraftverhalten jedoch gerade in sprachlich-heterogenen Lerngruppen im Zweitspracherwerb durch die animierte und inhaltsunterstützende Präsentation von Unterrichtsinhalten (Rosenshine, 1970a; Suleiman & Hashem, 1995), wie beispielsweise durch Gestik und Mimik von Lehrkräften, an Bedeutung (Sahlender & ten Hagen, 2024; Stam & Tellier, 2022).

2. Forschungsgegenstand und Forschungsdesiderat

2.1. Die Verbindung von Unterrichtsenthusiasmus und intrinsischer Motivation

Enthusiasmus von Lehrkräften

Der Enthusiasmus von Lehrkräften stellt eine wichtige Determinante für die Qualität des Handelns der Lehrkraft im Unterricht dar (Kunter et al., 2011). Neben dem Erleben von charakteristischen Emotionen wie Freude, Aufregung und Vergnügen der jeweiligen Lehrkraft (Kunter et al., 2008) kann der Enthusiasmus von Lehrkräften als ein positiver Ausdruck charakterisiert und damit als Verhalten im Unterricht (Keller et al., 2016) und als veränderbare Dimension der professionellen Kompetenz von Lehrkräften je nach Zeitpunkt (z. B. Unterrichtsstunde) und Kontext (z. B. Lerngruppe) angesehen werden (Gaspard & Lauermann, 2021). Eine holistische Begriffsdefinition von Enthusiasmus bezieht daher mit dem affektiven Erleben als Disposition der Lehrkraft und dem emotionalen Ausdruck als Verhalten im Unterricht zwei gleichberechtigte und sich ergänzende Komponenten ein (Keller et al., 2014a; 2016).

Das affektive Erleben des Enthusiasmus im Unterricht zeichnet sich durch positive Emotionen der Lehrkraft wie Begeisterung, positive Aufregung oder

Vergnügen aus (Kunter et al., 2008) und ist damit als internaler Prozess zu betrachten (Frenzel et al., 2009). Als positiven Gefühlszustand charakterisiert, ähnelt es der intrinsischen Motivation sehr stark. Allerdings ist die intrinsische Motivation als aktive Handlung anzusehen, während das Erleben des Enthusiasmus eher einen passiven Vorgang darstellt, der nicht willentlich gesteuert wird (Bösel, 2008).

Mit dem emotionalen Ausdruck von Enthusiasmus werden enthusiastische Verhaltensweisen von Lehrkräften im Unterricht definiert. Der Ausdruck des Enthusiasmus umfasst sowohl verbale als auch nonverbale Komponenten (Keller et al., 2016), wie den Gesichtsausdruck, Gestik, Mimik und Stimmvariationen, Bewegungen der Lehrkraft durch den Raum sowie eine verstärkte Intonation, die Wortwahl oder den Blickkontakt mit den Lernenden (Collins, 1978; Rosenshine, 1970b). Differenzierungen im Einsatz und in der Wertigkeit der einzelnen Indikatoren sind möglich (Murphy & Walls, 1994). Dabei deutet ein hohes Ausmaß an Energie, eine lebhafteste, dynamische Ausdrucksweise sowie ein hohes Tempo in den Bewegungen auf eine hohe Ausprägung des Enthusiasmus hin (Collins, 1978; Rosenshine, 1970a). Dagegen erscheint eine wenig enthusiastische Lehrkraft eher gelangweilt, lethargisch und monoton in ihrer Ausdrucksweise (Rosenhine, 1970a).

Entsprechend kann der emotionale Ausdruck des Enthusiasmus im Unterricht als persönlicher, individueller Stil der Lehrkraft in Bezug auf die Informationsvermittlung an die Lernenden definiert werden (Valentín et al., 2022), durch den die innere Einstellung der Lehrkraft zu den Lernenden und zum Unterricht nach außen sichtbar wird (Hooda, 2017). Daher sind enthusiastische Verhaltensweisen als externe Indikatoren zu definieren und im Gegensatz zu dem internalen Prozess des erlebten Enthusiasmus leicht von Außenstehenden zu identifizieren, weshalb sie auch für Lernende im Lernprozess leicht wahrnehmbar sind (Keller et al., 2016). So können Lehrkräfte durch enthusiastische Verhaltensweisen wichtige Inhalte im Unterricht hervorheben (Wood, 1999) oder die Lernenden bei der vertiefenden Verarbeitung des Lernmaterials unterstützen (Huangfu et al., 2022).

Werden beide Komponenten des Enthusiasmus gemeinsam betrachtet, ist von einem Zusammenhang zwischen dem affektiven Erleben und dem emotionalen Ausdruck des Enthusiasmus auszugehen (z. B. Bleck, 2017; Frenzel et al., 2009; Keller et al., 2014a; 2016; Kunter et al., 2008). Das emotionale Erleben von Enthusiasmus im Unterricht kann sich beispielsweise in Unterschieden hinsichtlich der Sprechweise, der verbalen Äußerungen (Kunter et al., 2011) oder auch des Unterrichtsstils äußern (Frenzel et al., 2009). Damit wäre das affektive Erleben von Enthusiasmus als Ursprung zu sehen, der den emotionalen Ausdruck im Verhalten begünstigt sowie fördert (Bettencourt

et al., 1983; Keller et al., 2016). Lehrkräfte sollten demnach sowohl Enthusiasmus für den Unterricht affektiv erleben, als auch im Unterricht ausdrücken, um einen positiven Zusammenhang mit der Motivation ihrer Lernenden herzustellen (Keller et al., 2016; Kunter et al., 2013). Allerdings transformiert sich das affektive Erleben scheinbar nicht automatisch in einen emotionalen Ausdruck des Enthusiasmus im Unterricht (Keller et al., 2018; Taxer & Frenzel, 2018) und überträgt sich damit möglicherweise auch nicht automatisch auf die Motivation der Lernenden. Grundsätzlich scheinen jedoch zwei Varianten der Übertragung möglich: 1) eine direkte Übertragung des Unterrichtsenthusiasmus der Lehrkräfte auf die Motivation der Lernenden sowie 2) eine indirekte Übertragung über weitere Konstrukte, die als Mediatoren fungieren.

Motivationale Konsequenzen des Unterrichtsenthusiasmus für Lernende

Für den emotionalen Ausdruck des Unterrichtsenthusiasmus zeigten Studien vielfach eine direkte Übertragung von positiven Lehrkraft-Emotionen (emotionale Ansteckung, Hatfield et al., 1993) auf Lernende, indem die Lernenden den Gesichtsausdruck, stimmliche Ausdrücke sowie die Körperhaltung oder die Emotionen ihrer Lehrkräfte direkt übernahmen (Hatfield et al., 1993). Ebenso scheint eine indirekte Übertragung über eine erhöhte Konzentration und Aufmerksamkeit der Lernenden durch die Wahrnehmung des ausgedrückten Enthusiasmus der Lehrkräfte möglich (Kunter & Holzberger, 2014). So konnten Frenzel und Kolleg*innen (2009) auf Klassenebene eine Übertragung der von Lehrkräften berichteten Freude auf die Freude von jugendlichen Lernenden vermittelt über den von Lernenden wahrgenommenen Enthusiasmus der Lehrkraft feststellen, wobei hier von einem wechselseitigen Prozess ausgegangen werden kann (Frenzel et al., 2018). Ein direkter Zusammenhang zwischen der berichteten Freude der Lehrkräfte und der Lernenden zeigte sich hier jedoch nicht (Frenzel et al., 2009).

Zudem deuten Studien auf positive Zusammenhänge zwischen dem erlebten und ausgedrückten Unterrichtsenthusiasmus von Lehrkräften und dem Interesse von jugendlichen Lernenden vermittelt über den von Lernenden wahrgenommenen ausgedrückten Enthusiasmus der Lehrkräfte hin (Keller et al., 2014a). Eine Vermittlung scheint auch zwischen der von jungen Erwachsenen im College wahrgenommenen enthusiastischen Vortragsweise der Lehrkräfte in Vorlesungen und der intrinsischen Motivation der Lernenden hinsichtlich der Bearbeitung des Vorlesungsmaterials (Patrick et al., 2000) sowie auf Individualebene zwischen dem von jugendlichen Lernenden wahrgenommenen Enthusiasmus der Lehrkraft und dem intrinsischen Wert

dieser Lernenden für dieses Unterrichtsfach zu funktionieren (Lazarides et al., 2018). Dabei konnte gezeigt werden, dass nicht nur einzelne Lernende, sondern die gesamte Lerngruppe durch den Enthusiasmus der Lehrkraft angesprochen werden kann (Palmer, 2020).

Darüber hinaus schaffen Lehrkräfte mit einem hohen Maß an Unterrichtsenthusiasmus bessere Lernbedingungen im Unterricht (Hooda, 2017) sowie ein günstigeres Unterrichtsklima (Fauth et al., 2019) und wählen häufiger relevanzbezogene Unterrichtsinhalte (Parrisius et al., 2020), welche die Motivation ihrer Lernenden positiv beeinflussen können.

2.2. Forschungsinteresse, Forschungsfragen und Hypothesen

Diese Erkenntnisse beziehen sich jedoch vielfach auf Untersuchungen in homogenen Lerngruppen, in denen Lernende hinsichtlich ihres Alters, ihrer sprachlichen Voraussetzungen oder ihres kulturellen Hintergrunds ähnliche Merkmale aufweisen. Dagegen sind Erkenntnisse über den Zusammenhang des Unterrichtsenthusiasmus von Lehrkräften und der intrinsischen Motivation von Lernenden in heterogenen Lerngruppen rar (Alter: Dewaele & Li, 2021; Frommelt et al., 2021; Keller et al., 2016; Valentín et al., 2022; Sprachhintergrund: siehe Keller et al., 2016; Migrationshintergrund: Hachfeld et al., 2012). Dabei könnte der Unterrichtsenthusiasmus von Lehrkräften insbesondere für sprachlich- und kulturell-heterogene Lerngruppen von Vorteil sein, indem Unterrichtsinhalte animiert und damit unterstützend zu sprachlichen Äußerungen präsentiert werden (Rosenshine, 1970a; Suleiman & Hashem, 1995; Gestik: Sahlender & ten Hagen, 2024; Stam & Tellier, 2022). Damit lässt sich möglicherweise Verständnisschwierigkeiten besser begegnen und durch die Stimme und Gestik einen positiven Einfluss auf die Lernendenmotivation nehmen (Rosenshine, 1970a).

Zudem beziehen sich die genannten Erkenntnisse vorrangig auf den mathematischen bzw. naturwissenschaftlichen Bereich (z. B. Frommelt et al., 2021; Keller et al., 2014a; Kunter et al., 2013; Review siehe Keller et al., 2016). Eine Übertragung auf andere Schulformen, Klassenstufen und Domänen steht dagegen noch aus und könnte zu differenziellen Erkenntnissen führen (Keller et al., 2014a).

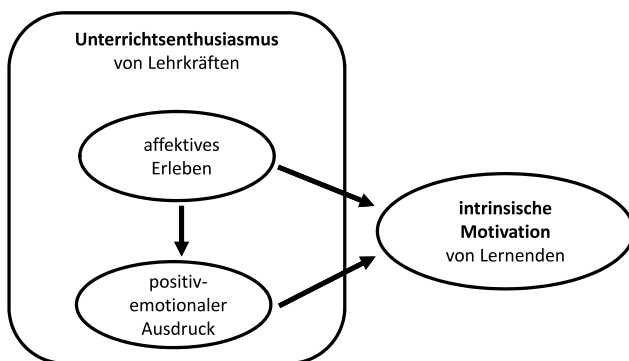
Darüber hinaus existiert ein breites Spektrum an Forschungserkenntnissen, die meist auf einem eindimensionalen Verständnis von Unterrichtsenthusiasmus basieren (siehe Keller et al., 2016; Kunter & Holzberger, 2014), weshalb Fragen hinsichtlich einer holistischen Betrachtungsweise und den Zusammenhängen mit anderen Konstrukten von Lernenden offenbleiben

(Keller et al., 2016). Dementsprechend besteht Forschungsbedarf nach dem Zusammenhang beider Komponenten des Unterrichtsenthusiasmus sowie einem möglicherweise differenziellen oder additiven Zusammenhang mit der Motivation von Lernenden (Keller et al., 2014a, 2016, 2018; Kunter & Holzberger, 2014).

Zudem erfolgte die Erfassung des affektiven Erlebens von Unterrichtsenthusiasmus bisher vielfach über Selbsteinschätzungen von Lehrkräften, während der emotionale Ausdruck des Unterrichtsenthusiasmus meist durch die Wahrnehmung von Lernenden erhoben wurde (siehe Keller et al., 2016). Beobachtungen des tatsächlichen Verhaltens von Lehrkräften im Unterricht sind dagegen selten, könnten allerdings wertvolle Einblicke in den emotionalen Ausdruck von Lehrkräften im Unterricht geben (Kunter & Holzberger, 2014).

Daher soll mit dieser Untersuchung eine Prüfung der Generalisierung bisheriger Forschungserkenntnisse auf den Sprachenunterricht – speziell auf den Unterricht von Deutsch als Zweitsprache (DaZ) – vorgenommen und dabei auf alters-, sprachlich- und kulturell-heterogene Lerngruppen fokussiert werden. In diesem Beitrag wird von einer holistischen Betrachtungsweise des Unterrichtsenthusiasmus von Lehrkräften ausgegangen, die beide Komponenten in den Blick nimmt. Die Datenerfassung soll auf Selbstberichtsdaten von Lehrkräften hinsichtlich ihres affektiven Erlebens und auf externen Beobachtungen des emotionalen Ausdrucks des Unterrichtsenthusiasmus der Lehrkräfte basieren. Anhand bisheriger Erkenntnisse könnten daher folgende Zusammenhänge auch für heterogene Lerngruppen vermutet werden:

Abbildung 1: Zusammenhänge der Komponenten von Unterrichtsenthusiasmus und der intrinsischen Motivation



Anmerkung: Weiterentwicklung in Anlehnung an Keller et al., 2016.

Zusätzlich werden die Leistung der Lernenden sowie das Alter und die Erfahrung der Lehrkraft im Unterrichten Beachtung finden, da Studien zeigten, dass jüngere Lehrkräfte und Lehrkräfte mit mehr Erfahrung im Mittel enthusiastischer als ältere Kolleg*innen mit weniger Erfahrung sind (Alter: Kunter et al., 2008; Unterrichtserfahrung: Kunter et al., 2011).

Zur Überprüfung dieses Zusammenhangs wird in dieser Untersuchung den folgenden Forschungsfragen (FF) nachgegangen:

FF1: Inwieweit besteht auf Klassenebene ein linearer Zusammenhang zwischen dem affektiven Erleben und dem emotionalen Ausdruck des Unterrichtsenthusiasmus von Lehrkräften in heterogenen Lerngruppen?

FF2: Inwieweit besteht auf Klassenebene ein linearer Zusammenhang zwischen dem affektiven Erleben, dem emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte und der intrinsischen Motivation der Lernenden in heterogenen Lerngruppen unter Einfluss der Kontrollvariablen Alter und Erfahrung der Lehrkraft im Unterrichten von DaZ sowie der Leistung der Lernenden?

Anhand der dargestellten bisherigen Erkenntnisse werden folgende Hypothesen (H) angenommen:

Es besteht ein signifikant moderat-positiver, linearer Zusammenhang zwischen dem affektiven Erleben und dem emotionalen Ausdruck des Unterrichtsenthusiasmus von Lehrkräften (H1).

Es besteht ebenso ein signifikant positiver, linearer Zusammenhang zwischen der intrinsischen Motivation von Lernenden und dem affektiven Erleben des Unterrichtsenthusiasmus von Lehrkräften (H2.1) und dem emotionalen Ausdruck des Unterrichtsenthusiasmus von Lehrkräften (H2.2).

Weiterhin wird angenommen, dass ein signifikant positiver, linearer Zusammenhang zwischen beiden Komponenten des Unterrichtsenthusiasmus der Lehrkräfte in gemeinsamer Betrachtung mit der intrinsischen Motivation der Lernenden (H2.3) besteht, der ebenfalls unter Einfluss der Kontrollvariablen Alter und Erfahrung der Lehrkraft im Unterrichten von DaZ und der Leistung der Lernenden bestehen bleibt (H2.4).

3. Methode

3.1. Kontext der Studie

Die dieser Untersuchung zugrundeliegenden Daten wurden im Rahmen des Projekts Verbundprojekts „Competencies of school teachers and adult educators in teaching German as a second language in linguistically diverse

classrooms“ (COLD) im Februar 2020 sowie von September 2020 bis März 2022 in Nordrhein-Westfalen (33 Vorbereitungsklassen, 18 Integrationskurse), Baden-Württemberg (6 Integrationskurse), Hessen (3 Integrationskurse), Berlin (4 Integrationskurse), Niedersachsen (3 Integrationskurse) und Rheinland-Pfalz (1 Integrationskurs), Deutschland, erhoben. Hierbei wurden Lehrkräfte und ihre Lernenden in Vorbereitungsklassen (VK) von Regelschulen (Sekundarstufe I) und Integrationskursen (IK) der Erwachsenenbildung mittels Fragebögen zur Selbsteinschätzung von Kompetenzen, Einstellungen und der Leistung der Lerngruppe befragt. Zudem wurden Videosequenzen von je einer 45-minütigen Unterrichtsstunde dieser befragten Lehrkräfte und Lernenden aufgezeichnet. Dabei erfolgte die Rekrutierung der Teilnehmenden über eine Ausschreibung auf der Internetseite des Deutschen Instituts für Erwachsenenbildung, über die sich interessierte Lehrkräfte bzw. Schulen freiwillig für die Studie anmelden konnten.

3.2. Stichprobe

Im Rahmen der Studie wurden Daten von $n = 33$ Lehrkräften (78.8% weiblich) aus 33 VK an 26 unterschiedlichen Regelschulen (Sekundarstufe I) erhoben. Die Lehrkräfte waren zum Zeitpunkt der Erhebung durchschnittlich 45 Jahre alt und hatten im Mittel eine allgemeine Lehrerfahrung von etwa 12 Jahren im Unterrichten sowie etwa 8 Jahre Erfahrung im Unterrichten von DaZ. Im Durchschnitt unterrichteten diese Lehrkräfte 10 Lernende pro Lerngruppe (siehe Tabelle 1).

Tabelle 1: Stichprobe

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<i>Lehrkräfte</i>					
Alter ^a	33/35	45.0/48.1	10.1/11.5	28.0/29.0	62.0/72.0
Allg. Lehrerfahrung ^a	33/35	12.2/15.3	8.4/10.0	2.1/0.0	31.0/36.3
DaZ Lehrerfahrung ^a	33/35	8.1/11.4	7.7/9.5	0.2/0.0	31.0/36.3
Anzahl LN pro LK	33/35	9.6/10.8	2.9/3.2	4.0/5.0	15.0/19.0
<i>Lernende</i>					
Alter ^a	314/377	13.5/35.0	2.1/9.9	10.0/16.0	24.0/74.0
Anzahl Monate in D	314/377	18.6/31.0	15.2/33.6	0.0/0.0	90.0/336.0

Anmerkung. Vorbereitungsklassen/Integrationskurse; DaZ = Deutsch als Zweitsprache; D = Deutschland; LK = Lehrkräfte; LN = Lernende.

^aAngaben in Jahren.

Zudem wurden $n = 314$ Lernende (43.0% weiblich) der o. g. Lehrkräfte befragt. Diese Lernenden waren zum Zeitpunkt der Erhebung im Durchschnitt 13.5 Jahre alt. Sie stammten aus 57 unterschiedlichen Herkunftsländern und spra-

chen 45 unterschiedliche Familiensprachen. Im Durchschnitt lebten die Lernenden seit 18.6 Monaten in Deutschland.

Darüber hinaus wurden $n = 35$ Lehrkräfte (74.3% weiblich) in 35 IK befragt. Diese Lehrkräfte waren zum Zeitpunkt der Erhebung im Durchschnitt etwa 48 Jahre alt, hatten eine Lehrererfahrung von durchschnittlich 15.3 Jahren und im Mittel 11.4 Jahre Lehrererfahrung im Unterrichten von DaZ. Diese Lehrkräfte unterrichteten etwa 11 Lernende pro Lerngruppe.

In den IK wurden Daten von $n = 377$ Lernenden erhoben (64.5% weiblich), die zum Zeitpunkt der Erhebung durchschnittlich 35 Jahre alt waren. Die Lernenden stammten aus 74 unterschiedlichen Herkunftsländern und sprachen 70 unterschiedliche Familiensprachen. Im Durchschnitt lebten diese Lernenden seit etwa 31 Monaten in Deutschland.

3.3. Erhebungsinstrumente

Affektives Erleben des Unterrichtsenthusiasmus der Lehrkräfte

Die Lehrkräfte schätzten ihr affektives Erleben des Unterrichtsenthusiasmus für die jeweilige Lerngruppe anhand von zwei Items, wie zum Beispiel „Ich unterrichte in dieser Klasse mit Begeisterung“ auf einer Skala von 1 = *Stimme gar nicht zu* bis 7 = *Stimme ganz zu* mittels eines Fragebogens selbst ein (Bauert et al., 2009). In der vorliegenden Untersuchung zeigte sich eine exzellente bis gute interne Konsistenz (VK: Cronbachs $\alpha = .960$ /IK: Cronbachs $\alpha = .877$).

Emotionaler Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte

Im Rahmen der Datenerhebung wurde mit Hilfe von zwei Videokameras eine Unterrichtsstunde von etwa 45 Minuten pro Lehrkraft und dazugehöriger Lerngruppe aufgezeichnet. Für die Erfassung des emotionalen Ausdrucks des Unterrichtsenthusiasmus der Lehrkräfte in der jeweiligen Lerngruppe wurden je zwei fünfminütige Sequenzen aus dieser Unterrichtsstunde ausgewählt.

Um die Vergleichbarkeit der Sequenzen zu gewährleisten, wurden Videosequenzen ausgewählt, die Unterrichtssituationen mit der gesamten Lerngruppe im Plenum enthielten, welche sowohl einen Redeanteil der Lehrkraft als auch Redeanteile von Lernenden umfassten. Um eine mögliche Verzerrung der Daten in den ersten Minuten einer offenen Beobachtung zu umgehen (Döring & Bortz, 2016), wurde der Beginn des ersten Messzeitpunkts frühestens 7 Minuten nach Beginn der Videoaufnahme angesetzt. Die zweite Videosequenz enthielt einen Zeitabstand von mindestens 10 Minuten ab dem

Ende der ersten Sequenz, um eventuelle Veränderung im emotionalen Ausdruck der Lehrkraft (z. B. stärkerer Ausdruck des Enthusiasmus zu Beginn der Unterrichtsstunde im Vergleich zum Ende) in die Datenerfassung einfließen lassen zu können. Danach wurde entsprechend der o. g. Bedingungen eine geeignete Sequenz ausgewählt.

Es folgte eine standardisierte Beobachtung durch zwei Beobachtungspersonen anhand von acht Indikatoren auf einer Skala von 1 = *Nicht enthusiastisch* bis 4 = *Sehr enthusiastisch* (Weiterentwicklung in Anlehnung an Collins, 1976 und Keller et al., 2014b). Die Indikatoren waren neben der (1) Stimme der Lehrkraft, (2) Augen und Blickkontakt zu den Lernenden, (3) Gestik und (4) Körperbewegungen der Lehrkraft im Raum, der (5) Gesichtsausdruck, (6) die Wortwahl der Lehrkraft (7) die Akzeptanz von Ideen und Gefühlen der Lernenden und (8) das allgemeine Energielevel der Lehrkraft. Aufgrund der Gesichtsbedeckung vieler Lehrkräfte mit einem Mund-Nasen-Schutz konnte der Indikator Gesichtsausdruck bei einem Großteil nicht beobachtet werden und wurde daher von den weiteren Analysen ausgeschlossen. Als Gesamtwert für den emotionalen Ausdruck der Lehrkräfte wurde ein Gesamtmittelwert der übrigen sieben Indikatoren anhand beider Messzeitpunkte und beider Beobachtungspersonen gebildet (VK: Cronbachs $\alpha = .937$ /IK: Cronbachs $\alpha = .917$).

Die Übereinstimmung zwischen den beiden Beobachtungspersonen wurde mit einem etwa 30-minütigen Training sichergestellt, in dem eine kurze Übersicht zum theoretischen Hintergrund des Themas und des Forschungsvorhabens gegeben wurde. Zudem ist der Beobachtungsbogen sowie das zu verwendende Codebook zur Einschätzung des Lehrkraft-Enthusiasmus vorgestellt worden. Anschließend erfolgte eine angeleitete Probeeinschätzung anhand zweier Videosequenzen aus der Pilotstudie, die nicht in die finale Untersuchung einfließen. Nach diesem Training schätzen die Beobachtungspersonen jeweils fünf Lehrkräfte anhand von jeweils zwei Videosequenzen ein, worauf die Übereinstimmung mittels des Intraklassen-Korrelationskoeffizienten (ICC_{just}) ermittelt wurde. Hier wurde eine gute bis sehr gute Übereinstimmung festgestellt, sodass die Beobachtungspersonen das Rating ohne weiteres Training fortsetzten. Die Übereinstimmung für alle in dieser Studie verwendeten Videodaten über beide Beobachtungspersonen hinweg war schlussendlich ebenfalls gut bis sehr gut (VK: $ICC_{just} = .708$ /IK: $ICC_{just} = .868$).

Intrinsische Motivation der Lernenden für den Deutschunterricht

Die Lernenden gaben ihre intrinsische Motivation für den Deutschunterricht anhand von vier Items (adaptiert an den Deutschunterricht nach Frey et al., 2009) mittels eines Fragebogens selbst an. Ein Beispielitem war „Der Deutschunterricht macht mir Spaß“, welches die Lernenden auf einer Skala von 1 = *Stimme gar nicht zu* bis 5 = *Stimme ganz zu* einschätzten. Insgesamt hatte die Skala in dieser Untersuchung eine sehr gute interne Konsistenz (VK: Cronbachs $\alpha = .859$ /IK: Cronbachs $\alpha = .902$).

Kontrollvariablen

Die Lehrkräfte machten Angaben u. a. hinsichtlich ihres Alters sowie ihrer Erfahrung im Unterrichten von DaZ. Für die Analysen wurden diese Angaben in Monate umgerechnet. Die Deutsch-Leistung der Lernenden wurde mittels eines Leistungstests (C-Test) erfasst und die erreichte Gesamtpunktzahl als Anhaltspunkt für die Analysen einbezogen.

3.4. Analysen

Zunächst wurden die Daten mit der Software SPSS (Version 29) aufbereitet, z-standardisiert sowie mit Hilfe von MPlus (Version 8.10) analysiert. Zur Untersuchung der Zusammenhänge des erlebten und ausgedrückten Unterrichtsenthusiasmus der Lehrkräfte mit der intrinsischen Motivation der Lernenden wurden einfache lineare Regressionsanalysen und hierarchisch-lineare Regressionsmodelle (HLM; Raudenbush & Bryk, 2001) berechnet, die die genestete Datenstruktur (Lernende in Lerngruppen) berücksichtigten. Aufgrund des Forschungsinteresses auf Klassenebene wurden die Angaben der Lernenden aggregiert.

Der Zusammenhang zwischen dem erlebten Unterrichtsenthusiasmus und dem emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte (FF1) wurde anhand einer einfachen linearen Regressionsanalyse untersucht (Modell 1, siehe Tabelle 3). Für die statistische Berechnung galt das affektive Erleben des Unterrichtsenthusiasmus als unabhängige Variable, während der emotionale Ausdruck als abhängige Variable festgelegt wurde.

Zur Überprüfung des Zusammenhangs zwischen den Komponenten des Unterrichtsenthusiasmus der Lehrkräfte und der intrinsischen Motivation der Lernenden auf Klassenebene (FF2) wurden HLM berechnet. Hierfür wurde zunächst die intrinsische Motivation der Lernenden als abhängige Variable und das affektive Erleben sowie der emotionale Ausdruck des Unterrichtsent-

husiasmus als unabhängige Variablen festgelegt. Für die HLM wurde zunächst ein Random Intercept-Only Modell für die intrinsische Motivation auf Basis der Maximum Likelihood berechnet. Anschließend wurden mehrere hierarchische Regressionsmodelle berechnet, in denen Zusammenhänge der intrinsischen Motivation mit dem affektiven Erleben des Unterrichtsenthusiasmus (Modell 2), dem emotionalen Ausdruck (Modell 3) sowie gemeinsam mit beiden Komponenten des Unterrichtsenthusiasmus als unabhängige Variablen (Modell 4) berechnet wurden. Anschließend wurde der Zusammenhang aus Modell 4 unter Kontrolle des Alters und der Erfahrung der Lehrkraft im Unterrichten von DaZ sowie der Leistung der Lernenden auf Klassenebene (Modell 5) berechnet.

4. Ergebnisse

4.1. Deskriptive Statistiken

Die befragten Lehrkräfte schätzten das eigene affektive Erleben des Unterrichtsenthusiasmus im Durchschnitt sehr hoch ein (VK: $M = 6.61$ bzw. IK: $M = 6.70$). Ebenso wurde der emotionale Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte durch die externen Beobachtungspersonen als eher enthusiastisch eingestuft (VK: $M = 2.67$ /IK: $M = 2.54$; siehe Tabelle 2).

Die intrinsische Motivation der Lernenden für den Deutschunterricht war im Durchschnitt ebenfalls recht hoch (VK: $M = 4.07$ /IK: $M = 4.31$). Hier zeigte sich eine geringe Standardabweichung (VK: $SD = 0.88$ /IK: $SD = 0.84$), was auf eine geringe Varianz zwischen den untersuchten Lerngruppen hindeuten könnte. Die durchschnittliche Klassenleistung der Lernenden im Deutschunterricht anhand des Deutschtests lag eher im unteren Drittel (VK: $M = 22.52$ /IK: $M = 25.45$ Punkte) mit einer recht großen Streuung (VK: $SD = 15.52$ /IK: $SD = 14.07$), was ein Hinweis auf eine hohe Heterogenität der Lernenden innerhalb der Stichprobe sowie auch innerhalb der Lerngruppen sein könnte.

Tabelle 2: Deskriptive Statistiken

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Cronbachs α</i>	<i>ICC</i>
<i>Lehrkräfte</i>					
Erlebter Unterrichtsenthusiasmus	33/35	6.61/6.70	0.62/0.55	.960/.877	
Ausgedrückter Unterrichtsenthusiasmus ^a	33/35	2.67/2.54	0.44/0.45	.937/.917	.708/.868
<i>Lernende</i>					
Intrinsische Motivation	299/364	4.07/4.31	0.88/0.84	.859/.902	
Deutschleistung im Test	312/372	22.52/25.45	15.52/14.07		

Anmerkung: Vorbereitungsklassen/Integrationskurse.

^a Interrater-Reliabilität: two-way mixed, consistency

4.2. Zusammenhang der Komponenten des Unterrichtsenthusiasmus

Zur Überprüfung des linearen Zusammenhangs zwischen dem affektiven Erleben als unabhängige Variable und dem emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte als abhängige Variable (H1) wurde eine einfache lineare Regressionsanalyse durchgeführt. Diese zeigte weder für die VK noch die IK einen signifikanten Zusammenhang (VK: $\beta = .008$ /IK: $\beta = .057$; siehe Tabelle 3).

4.3. Zusammenhang des Unterrichtsenthusiasmus und der intrinsischen Motivation

Für den Zusammenhang zwischen dem affektiven Erleben, dem emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte als unabhängige Variablen und der intrinsischen Motivation der Lernenden für den Deutschunterricht als abhängige Variable wurde sowohl für die separate (H2.1 & H2.2) als auch die gemeinsame Betrachtung (H2.3) der Komponenten des Unterrichtsenthusiasmus ein signifikant positiver, linearer Zusammenhang mit der intrinsischen Motivation der Lernenden auf Klassenebene angenommen.

Zur Überprüfung dieser Hypothesen wurde zunächst ein Random Intercept-Only Modell für die intrinsische Motivation der Lernenden für den Deutschunterricht berechnet. Sowohl in den VK als auch den IK zeigten sich Unterschiede der intrinsischen Motivation innerhalb der Klassen zwischen den Lernenden (VK: $\sigma^2_w = .982$, $p < .001$; IK: $\sigma^2_w = .893$, $p < .001$), aber keine signifikanten Unterschiede zwischen den beobachteten Klassen (VK: $\sigma^2_b = .052$, $p = .192$; IK: $\sigma^2_b = .043$, $p = .116$).

Mittels hierarchisch-linearer Regressionsmodelle ließ sich in den VK für die intrinsische Motivation der Lernenden weder mit dem selbstberichteten affektiven Erleben (Modell 2; $\beta = -.020$), noch mit dem emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte (Modell 3; $\beta = .046$) ein linearer Zusammenhang auf Klassenebene feststellen (siehe Tabelle 3). Daher werden die Hypothesen 2.1 und 2.2 für die VK abgelehnt.

Diese Ergebnisse zeigten sich für die VK ebenso in dem hierarchisch-linearen Regressionsmodell, in dem der Zusammenhang beider Komponenten des Unterrichtsenthusiasmus der Lehrkraft und der intrinsischen Motivation der Lernenden für den Deutschunterricht geprüft wurde (Modell 4). In diesem Modell war sowohl für das affektive Erleben ($\beta = -.021$), ebenso wie für den emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkraft ($\beta = .046$) kein signifikanter linearer Zusammenhang mit der intrinsischen Motivation der Lernenden auf Klassenebene zu erkennen (siehe Tabelle 3), weshalb Hypothese 2.3 für die VK ebenfalls abgelehnt wird.

Für die IK zeigte sich auf Klassenebene ebenso wie für die VK kein signifikant linearer Zusammenhang der intrinsischen Motivation der Lernenden mit dem emotionalen Ausdruck (Modell 3; IK: $\beta = .074$); wohl aber mit dem selbstberichteten affektiven Erleben des Unterrichtsenthusiasmus der Lehrkräfte (Modell 2; IK: $\beta = .123$). Daher wird Hypothese 2.1 für die IK angenommen; während Hypothese 2.2 auch für die IK abgelehnt werden muss.

Diese Ergebnisse wiederholten sich für die IK in dem hierarchisch-linearen Regressionsmodell (Modell 4), mit dem der Zusammenhang beider Komponenten des Unterrichtsenthusiasmus der Lehrkraft und der intrinsischen Motivation der Lernenden für den Deutschunterricht geprüft wurde. In diesem gemeinsamen Modell war für das affektive Erleben ($\beta = .120$), aber nicht für den emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkraft ($\beta = .069$) ein signifikanter linearer Zusammenhang mit der intrinsischen Motivation der Lernenden auf Klassenebene zu erkennen (siehe Tabelle 3), weshalb Hypothese 2.3 für die IK lediglich teilweise angenommen werden kann.

Tabelle 3: Ergebnisse der hierarchisch-linearen Regressionsanalysen

		Vorbereitungs- klassen			Integrations- kurse		
Unabhängige Variable(n)	Abhängige Variable	β	SE	p	β	SE	p
1 Erlebter	Ausgedrückter						
Unterrichtsenthusiasmus	Unterrichtsenthusiasmus	.008	.157	.961	.057	.137	.681
2 Erlebter	Intrinsische Motivation						
Unterrichtsenthusiasmus		-.020	.062	.750	.123	-.057	.032
3 Ausgedrückter	Intrinsische Motivation						
Unterrichtsenthusiasmus		.046	.085	.589	.074	-.052	.153
4 Erlebter	Intrinsische Motivation						
Unterrichtsenthusiasmus,		-.021	.059	.716	.120	-.054	.028
ausgedrückter							
Unterrichtsenthusiasmus		.046	.085	.584	.069	.051	.176
5 Erlebter	Intrinsische Motivation						
Unterrichtsenthusiasmus,		.028	.076	.713	.157	.051	.002
ausgedrückter							
Unterrichtsenthusiasmus,		.053	.082	.515	.021	-.056	.714
Alter LK,		.002	.008	.776	-.005	.006	.368
Erfahrung DaZ LK,		-.008	.006	.194	-.004	.006	.500
Leistung LN		-.001	.001	.260	.000	.001	.831

Anmerkung. DaZ = Deutsch als Zweitsprache; LK = Lehrkraft; LN = Lernende.

4.4. Zusammenhang des Unterrichtsenthusiasmus und der intrinsischen Motivation inklusive der Kontrollvariablen

Auch unter Einbezug der Kontrollvariablen zeigten sich in den VK auf Klassenebene (Modell 5) weder für das affektive Erleben ($\beta = .028$), noch für den emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkraft ($\beta = .053$) ein signifikanter Zusammenhang mit der intrinsischen Motivation der Lernenden (siehe Tabelle 3). Entsprechend muss auch Hypothese 2.4 für die VK abgelehnt werden.

In den IK dagegen zeigte sich auch unter Einbezug der Kontrollvariablen auf Klassenebene (Modell 5) für das affektive Erleben ($\beta = .157$) ein signifikanter Zusammenhang mit der intrinsischen Motivation der Lernenden für den Deutschunterricht, aber wiederholt nicht für den emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkraft ($\beta = .021$; siehe Tabelle 3). Entsprechend kann die Hypothese 2.4 für die IK lediglich teilweise angenommen werden.

5. Diskussion

5.1. Interpretation und Einordnung der Ergebnisse

Zusammenfassend lässt sich in der vorliegenden Untersuchung sowohl in den VK der Regelschulen als auch in den IK der Erwachsenenbildung auf Klassenebene kein Zusammenhang zwischen dem affektiven Erleben und dem emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte finden. Ebenso zeigten sich in den VK weder in einer separaten Analyse, noch in der gemeinsamen Betrachtung beider Komponenten des Unterrichtsenthusiasmus der Lehrkräfte Zusammenhänge mit der intrinsischen Motivation der Lernenden. In den IK dagegen war auf Klassenebene ein Zusammenhang des erlebten Unterrichtsenthusiasmus der Lehrkräfte mit der intrinsischen Motivation zu beobachten. Jedoch bestand auch hier kein Zusammenhang des emotionalen Ausdrucks des Unterrichtsenthusiasmus der Lehrkräfte mit der intrinsischen Motivation der Lernenden.

Zusammenhang der Komponenten des Unterrichtsenthusiasmus

Da in der vorliegenden Untersuchung sowohl in den VK als auch den IK auf Klassenebene kein linearer Zusammenhang zwischen dem affektiven Erleben und dem emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte gefunden werden konnte, scheint sich bei Lehrkräften, die von einem hohen Maß an erlebtem Unterrichtsenthusiasmus berichten, dieser Unterrichtsenthusiasmus nicht zwangsläufig in ihrem Verhalten widerzuspiegeln. Dies steht im Gegensatz zu bisherigen Forschungserkenntnissen, die eine Verknüpfung zwischen dem affektiven Erleben und dem emotionalen Ausdruck des Unterrichtsenthusiasmus von Lehrkräften zeigten (z. B. Frenzel et al., 2009; Keller et al., 2014a; Kunter et al., 2008; Taxer & Frenzel, 2018). Die Ergebnisse der vorliegenden Untersuchung sprechen vielmehr für ein differenzielles Auftreten beider Komponenten. Dies stützt die These von Keller und Kolleg*innen (2018), nach der die Lehrkräfte durchaus Enthusiasmus im Unterricht erleben, diesen jedoch nicht zwangsläufig in ihrem Verhalten äußern.

Ein weiterer möglicher Grund für diesen fehlenden Zusammenhang könnte in der Operationalisierung dieser Untersuchung liegen, in der das affektive Erleben des Unterrichtsenthusiasmus der Lehrkräfte als allgemeine Komponente erhoben, während der emotionale Ausdruck lediglich anhand einer Unterrichtsstunde erfasst wurde. Zudem wäre ein nicht-authentischer emotionaler Ausdruck der Lehrkräfte in dieser Untersuchung denkbar (Keller et al., 2018), sodass sich keine Zusammenhänge mit dem affektiven Erleben

des Unterrichtsenthusiasmus und der intrinsischen Motivation der Lernenden finden.

Zusammenhang des erlebten Unterrichtsenthusiasmus und der intrinsischen Motivation

In der Betrachtung des Zusammenhangs des affektiven Erlebens des Unterrichtsenthusiasmus der Lehrkraft und der intrinsischen Motivation der Lernenden zeigte sich in den VK, dass Lerngruppen von Lehrkräften mit einem hohen Grad an erlebtem Unterrichtsenthusiasmus nicht intrinsisch motivierter sind als Lerngruppen von Lehrkräften, die von einem geringen Erleben des Enthusiasmus für den Unterricht berichteten. In den IK ließ sich jedoch ein positiver linearer Zusammenhang feststellen. Diese Erkenntnisse bestätigen das heterogene Bild bisheriger Untersuchungen, die zum einen positive Zusammenhänge der Motivation von Lernenden und dem erlebten Unterrichtsenthusiasmus von Lehrkräften zeigten (z. B. Kunter et al., 2008, 2011; Liew et al., 2017; Patrick et al., 2000); aber zum anderen auch keinen direkten Zusammenhang feststellen konnten (z. B. Frenzel et al., 2009; Frommelt et al., 2021).

In der vorliegenden Untersuchung könnte die unterschiedliche Altersstruktur zwischen VK und IK ein möglicher Grund für die verschiedenen Ergebnismuster in der Betrachtung des Zusammenhangs zwischen dem erlebten Unterrichtsenthusiasmus der Lehrkräfte und der intrinsischen Motivation der Lernenden für den Deutschunterricht sein. Dementsprechend wäre mit steigendem Alter – also möglicherweise in Lerngruppen im Erwachsenenalter – eher ein Zusammenhang mit dem erlebten Unterrichtsenthusiasmus der Lehrkräfte denkbar, während sich dieser Zusammenhang bei jüngeren Lernenden im Schulalter nicht findet. Innerhalb der Lerngruppen zeigten sich keine Zusammenhänge zwischen dem Alter der Lernenden und der intrinsischen Motivation sowie dem erlebten Unterrichtsenthusiasmus, sodass dieses Argument tatsächlich nur für die älteren Lernenden in der Erwachsenenbildung denkbar ist.

Im Vergleich der Stichproben der Lehrkräfte sowie der Lernenden ließen sich einige Unterschiede und damit auch mögliche Erklärungsansätze der unterschiedlichen Ergebnismuster für VK und IK finden. Die Lehrkräfte in den IK unterrichteten im Durchschnitt doppelt so viele Unterrichtsstunden pro Woche in DaZ-Kursen sowie durchschnittlich etwa 4 Stunden mehr in dem jeweils eingeschätzten Kurs im Vergleich zu den Lehrkräften in den VK. Zudem schätzten die Lehrkräfte in den IK ihre eigenen Fähigkeiten und Kenntnisse für das Unterrichten von DaZ höher ein und besitzen häu-

figer zusätzliche Erfahrungen im Unterrichten von DaZ. Allerdings finden sich für die IK keine Zusammenhänge dieser Konstrukte mit dem erlebten Unterrichtsenthusiasmus der Lehrkräfte. Für die VK zeigten sich lediglich positive lineare Zusammenhänge des erlebten Unterrichtsenthusiasmus mit den Unterrichtsstunden in diesem Kurs sowie mit zusätzlichen DaZ-Erfahrungen, was jedoch nicht die differenziellen Befunde zwischen VK und IK auf Klassenebene erklären dürfte.

Ein weiterer möglicher Erklärungsansatz mit Blick auf die Stichprobe der Lernenden könnte in der geringen Varianz der intrinsischen Motivation der Lernenden auf Klassenebene liegen, die keine signifikanten Unterschiede zwischen den beobachteten Klassen zeigte (VK: $\sigma^2_b = .052$, $p = .192$; IK: $\sigma^2_b = .043$, $p = .116$). Zu diskutieren wäre in diesem Kontext darüber hinaus die Anzahl der beobachteten Klassen, die möglicherweise mit $n = 33$ für die VK und $n = 35$ für die IK nicht ausreichte, um diese Zusammenhänge auf Klassenebene systematisch zu untersuchen. Die Betrachtung der Konstrukte auf Klassenebene lässt sich unter Berücksichtigung der intrinsischen Motivation der Lernenden als individuelles Merkmal, aber auch in Bezug auf den Unterrichtsenthusiasmus der Lehrkräfte für das Unterrichten einzelner Lernender kritisch betrachten. Daher lässt sich für eine Untersuchung auf Individualebene in zukünftiger Forschung plädieren.

Darüber hinaus wären indirekte Zusammenhänge weiterer Konstrukte mit der intrinsischen Motivation und dem erlebten Unterrichtsenthusiasmus der Lehrkräfte möglich. Sowohl in den IK als auch den VK zeigte sich auf Individualebene ein positiver Zusammenhang des erlebten Unterrichtsenthusiasmus mit der wahrgenommenen Leistung der Lerngruppe durch die Lehrkräfte, welche wiederum in positivem Zusammenhang mit der tatsächlichen Testleistung der Lernenden im Deutschtest in den VK sowie auch den IK stand. Allerdings war lediglich in den IK ein positiver Zusammenhang ihrer Leistung im Deutschtest und der intrinsischen Motivation der Lerngruppe für den Deutschunterricht festzustellen. Zurückzuführen ist insbesondere der letztgenannte differenzielle Zusammenhang möglicherweise darauf, dass Lernende in den IK im Gegensatz zu Lernenden in den VK weniger durch extrinsische Faktoren, wie etwa die Noten motiviert lernen, sowie im Sinne der Erwartungs-Wert-Theorie (Wigfield & Eccles, 2000) insbesondere dem Deutschunterricht einen stärkeren intrinsischen Wert zuschreiben. Allerdings zeigten sich diese Zusammenhänge nicht auf Klassenebene und bieten somit keine plausible Erklärung für die differenziellen Zusammenhänge von VK und IK auf Klassenebene in dieser Untersuchung.

Ein weiterer möglicher Erklärungsansatz könnte in der Wahrnehmung von zwischenmenschlichen Hinweisen liegen, die in Zusammenhang mit

der Motivation von Lernenden stehen (Wild et al., 1997). Dementsprechend würden Lernende in den IK anders als in den VK die zwischenmenschliche Kommunikation ihrer Lehrkraft wahrnehmen und dadurch für den Deutschunterricht motiviert werden. Allerdings lässt sich diese These anhand der vorliegenden Studie nicht überprüfen.

Zusammenhang des emotionalen Ausdrucks und der intrinsischen Motivation

Für den emotionalen Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte zeigte sich weder in den VK noch in den IK ein linearer Zusammenhang mit der intrinsischen Motivation der Lernenden auf Klassenebene. Demnach sind Lerngruppen von Lehrkräften, die ihrem Enthusiasmus im Unterricht einen starken Ausdruck verleihen, nicht stärker intrinsisch motiviert als Lerngruppen, die von Lehrkräften mit einem niedrigen emotionalen Ausdruck des Unterrichtsenthusiasmus unterrichtet wurden.

Ein möglicher Grund hierfür könnte der im Vergleich zum affektiven Erleben vergleichsweise niedrigere Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte sein. Eventuell reicht insbesondere in den VK das affektive Erleben von Enthusiasmus im Unterricht allein nicht aus, um damit die intrinsische Motivation der Lernenden anzusprechen (Keller et al., 2018). Somit könnten die Ergebnisse dieser Untersuchung ein Hinweis darauf sein, dass in den VK anders als in den IK kein direkter Zusammenhang zwischen dem affektiven Erleben von Unterrichtsenthusiasmus der Lehrkraft und der intrinsischen Motivation der Lernenden existiert. Vielmehr scheinen hier weitere oder andere Faktoren als Mediatoren fungieren zu müssen, um die Motivation der Lernenden anzusprechen. Denkbar wäre der Einbezug von schüler*innenorientierten Unterrichtspraktiken in die Untersuchung, da beispielsweise ein positiver Zusammenhang des Unterrichtsenthusiasmus von Lehrkräften mit einer stärkeren Förderung des Selbstbestimmungserlebens oder einem qualitativ höherwertigen Unterrichtshandeln mit Blick auf die drei Basisdimensionen von Unterrichtsqualität möglich erscheint (Kunter & Holzberger, 2014), die wiederum mit einer höheren intrinsischen Motivation von Lernenden einhergehen könnten. Diese müssten in weiteren Studien näher untersucht werden.

Eine möglicherweise fehlende Verknüpfung des emotional ausgedrückten Unterrichtsenthusiasmus mit dem Inhalt der jeweiligen Unterrichtsstunde könnte ebenso ein Grund für die fehlenden Zusammenhänge in den VK sein (Wood, 1999). Eventuell steht der emotionale Ausdruck des Unterrichtsenthusiasmus in diesen Lerngruppen erst dann mit der intrinsischen Motivation

der Lernenden in Verknüpfung, wenn sie zur Unterstreichung der Unterrichtsinhalte genutzt wird.

Weiterhin könnte auch die gezielte Nutzung von Gestik und Mimik im DaZ-Unterricht, um beispielsweise einen geringen Wortschatz der Lernenden zu kompensieren und so nonverbal Inhalte zu vermitteln, nicht im Zusammenhang mit dem affektiven Erleben des Unterrichtsenthusiasmus und damit einem *emotionalen* Ausdruck der Lehrkräfte stehen. Daher wäre eine emotionale Ansteckung (siehe Hatfield et al., 1993) der Lernenden nicht möglich, was sich in dem fehlenden Zusammenhang mit der intrinsischen Motivation der Lernenden widerspiegelt. So könnte in den VK das Ausbleiben des Ausdrucks des Unterrichtsenthusiasmus negativ von Lernenden wahrgenommen worden sein und dementsprechend nicht zu einer Steigerung der intrinsischen Motivation geführt haben (Burić & Mornar, 2023; Keller et al., 2018). Warum sich dennoch ein Zusammenhang des erlebten Unterrichtsenthusiasmus und der intrinsischen Motivation der Lernenden in den IK zeigt, obwohl auch hier keine Zusammenhänge mit dem emotionalen Ausdruck bestehen, müsste in zukünftigen Studien genauer untersucht werden.

Ebenso ist anzunehmen, dass die Emotionen von Lehrkräften von Unterrichtsstunde zu Unterrichtsstunde variieren (Becker et al., 2015) und sich sogar innerhalb von Unterrichtsstunden ändern können (Keller et al., 2018). Damit ist nicht von einem permanenten, gleichbleibenden Erleben (Becker et al., 2015) und ebenso nicht von einem permanenten emotionalen Ausdruck des Enthusiasmus der Lehrkräfte auszugehen. Der fehlende Zusammenhang insbesondere des emotionalen Ausdrucks des Unterrichtsenthusiasmus der Lehrkraft mit dem affektiven Erleben sowie mit der intrinsischen Motivation der Lernenden könnte somit auf die Erfassung des affektiven Erlebens als generelle Einschätzung und des emotionalen Ausdrucks des Unterrichtsenthusiasmus anhand einer Unterrichtsstunde in dieser Untersuchung zurückzuführen sein. Möglicherweise unterscheidet sich der emotionale Ausdruck hinsichtlich spezieller Inhalte oder Arbeitsphasen, sodass die Beobachtungen in dieser Untersuchung einen situativen emotionalen Ausdruck der Lehrkraft in der beobachteten Unterrichtsstunde bzw. -sequenz darstellen (Keller et al., 2018) und diese nicht zu einer dauerhaft anhaltenden Steigerung der intrinsischen Motivation der Lernenden für den Unterricht führen (Patrick et al., 2000).

Denkbar wäre zudem, dass ein mittleres Ausmaß des emotionalen Ausdrucks des Unterrichtsenthusiasmus eher in Zusammenhang mit der intrinsischen Motivation von Lernenden steht als ein hohes Niveau. Eine extreme Form des emotionalen Ausdrucks des Unterrichtsenthusiasmus könnte eher zu einem verstärkten Fokus auf die Animation der Lehrkraft und damit zu

einer verminderten Konzentration auf den Inhalt und den Lernprozess führen (Rosenshine, 1970a). Ein solcher kurvilinearere Zusammenhang wäre jedoch mit der Untersuchung eines linearen Zusammenhangs des Unterrichtsenthusiasmus der Lehrkräfte mit der intrinsischen Motivation der Lernenden in dieser Studie nicht abzubilden.

Übertragbarkeit der Ergebnisse auf heterogene Lerngruppen den DaZ-Unterricht

Neben den bislang diskutierten Faktoren für die fehlenden Zusammenhänge in dieser Untersuchung, wären Auswirkungen auf die Ergebnisse aufgrund Besonderheiten der Stichprobe denkbar. Heterogene Lerngruppen stehen – anders als homogene Lerngruppen – vor besonderen Herausforderungen, was sowohl Lernende als auch Lehrkräfte beeinflusst. Die heterogenen Lerngruppen in dieser Untersuchung sind häufig von Wechseln der Lerngruppenzusammensetzung oder Neuzugängen von Lernenden geprägt, sodass Zusammenhänge von Lehrkraft- und Lernendenkompetenzen aufgrund einer kurzfristigen Zusammenarbeit noch nicht beobachtbar sind und erst nach einer längeren Zusammenarbeit sicht- und wahrnehmbar werden (z. B. ähnlich dem Zusammenhang von wahrgenommenem Unterrichtsmanagement und Unterrichtsenthusiasmus von Lehrkräften; Lazarides et al., 2021).

Darüber hinaus könnte für die Lernenden ein Erleben der Lehrkraft in unterschiedlichen Unterrichtssituationen essenziell sein, um den Enthusiasmus der Lehrkraft im Unterricht zu übernehmen und damit stärker intrinsisch motiviert für den Unterricht zu sein (Wood, 1999). Lernende, die erst kürzlich von einer Lehrkraft unterrichtet werden, hatten so eventuell noch nicht die Gelegenheit, den Enthusiasmus der Lehrkraft im Unterricht in unterschiedlichen Situationen zu erleben, als solchen zu erkennen und dadurch intrinsisch stärker motiviert zu werden. Allerdings lässt sich diese These in der vorliegenden Untersuchung nicht bestätigen, da sich weder für die VK noch für die IK ein Zusammenhang zwischen dem Enthusiasmus der Lehrkräfte, der Motivation der Lernenden für den Deutschunterricht und der Dauer des Unterrichts bei dieser Lehrkraft finden lässt.

Darüber hinaus könnte die kulturelle Heterogenität der Lernenden in den Lerngruppen dazu geführt haben, dass auf Klassenebene kein linearer Zusammenhang zwischen den Komponenten des Unterrichtsenthusiasmus und der intrinsischen Motivation der Lernenden gefunden werden konnte. So sind differenzielle Zusammenhänge aufgrund beispielsweise der Zugehörigkeit zur gleichen kulturellen Gruppe der Lehrkraft oder aufgrund der Fähigkeit

zur Interpretation von bestimmter Gestik und Mimik durch die Lernenden möglich (Orosz et al., 2015).

Ebenso lässt sich anhand dieser Arbeit keine Übertragbarkeit bisheriger Erkenntnisse für den DaZ-Unterricht erkennen. Beim Erlernen einer neuen Sprache ist der besondere emotionale Kontext der Lernenden zu berücksichtigen (Dewaele & Li, 2020), der sich durchaus von den Emotionen in anderen Unterrichtsfächern unterscheiden kann. Mögliche Zusammenhänge dieser Emotionen und der intrinsischen Motivation der Lernenden sind in Zukunft zu untersuchen.

Darüber hinaus stehen Lehrkräfte im DaZ-Unterricht vor der Herausforderung, die Lernenden in einer für sie fremden Sprache zu unterrichten, sodass nicht nur die Inhalte des Unterrichts, sondern auch die Sprache selbst zum Lerngegenstand wird. Dementsprechend entfällt eine wichtige Methode der Vermittlung von Unterrichtsinhalten durch die Sprache und muss durch andere Methoden ersetzt bzw. ergänzt werden. Dies kann beispielsweise Gestik oder Mimik sein, die gezielt eingesetzt werden, um sprachliche Inhalte zu untermauern und verständlich zu machen (Sahlender & ten Hagen, 2024).

5.2. Limitationen und weitere Forschung

So wäre es denkbar, dass der in dieser Untersuchung erfasste Ausdruck der Lehrkraft nicht auf dem affektiven Erleben des Unterrichtsenthusiasmus und damit einem *emotionalen* Ausdruck basiert, was eine der Limitationen dieser Untersuchung aufzeigt. Die Erhebung des Ausdrucks des Unterrichtsenthusiasmus im Fremdsprachenunterricht ist daher möglicherweise mit Einschränkungen aufgrund des gewählten Beobachtungsinstruments zu betrachten, da der Einsatz von Gestik, Mimik oder Körperbewegungen nicht unbedingt ein Indikator für den Ausdruck des Unterrichtsenthusiasmus der Lehrkraft sein muss. Eventuell bildet der beobachtete emotionale Ausdruck neben dem Enthusiasmus der Lehrkraft andere Faktoren ab, wie einen methodisch-didaktischen Einsatz. Darüber hinaus ist für die Einschätzung des emotionalen Ausdrucks des Unterrichtsenthusiasmus als zeitlich und situativ instabile Größe zwei Sequenzen einer Unterrichtsstunde möglicherweise nicht ausreichend. Hier wäre eine Einschätzung über mehrere Unterrichtsstunden hinweg in zukünftigen Studien wünschenswert; ebenso wie die Hinzunahme des Indikators des *Gesichtsausdrucks*, der in dieser Untersuchung aufgrund des häufigen Tragens eines Mund-Nasen-Schutzes der Lehrkräfte von den Analysen ausgeschlossen werden musste.

Zudem ist neben der Einschätzung des emotionalen Ausdrucks des Unterrichtsenthusiasmus durch externe Beobachtungspersonen eine Einschätzung durch die Lernenden in zukünftigen Untersuchungen anzustreben. Da der von Lernenden wahrgenommene Unterrichtsenthusiasmus der Lehrkraft einen wesentlichen Aspekt in der Betrachtung dieses Zusammenhangs darstellt (Frenzel et al., 2009; Keller et al., 2014b), könnte der fehlende Zusammenhang zwischen den Komponenten des Unterrichtsenthusiasmus und der intrinsischen Motivation der Lernenden durchaus in der fehlenden Wahrnehmung des Ausdrucks des Unterrichtsenthusiasmus durch die Lernenden liegen. In zukünftigen Untersuchungen ist dementsprechend eine zusätzliche Erfassung des wahrgenommenen Unterrichtsenthusiasmus durch die Lernenden sinnvoll, um mögliche differenzierte Wahrnehmungen des Unterrichtsenthusiasmus durch die Lernenden im Gegensatz zu externen Beobachtungspersonen abbilden zu können.

In diesem Zusammenhang ließe sich ein weiteres Mal für eine zukünftige Längsschnittuntersuchung plädieren, da in der vorliegenden Untersuchung keine Aussagen über kausale oder auch reziproke Effekte getroffen werden können. Möglich wäre demnach, dass die intrinsische Motivation der Lernenden als ausschlaggebender Faktor fungiert, der das Erleben und auch den Ausdruck des Unterrichtsenthusiasmus der Lehrkräfte beeinflusst, während ein umgekehrter Effekt ebenso möglich erscheint.

Darüber hinaus erscheint eine weiterführende Erforschung der Zusammenhänge auf Individualebene sinnvoll, um die differenzielle Wahrnehmung des Unterrichtsenthusiasmus innerhalb der einzelnen Lerngruppen und dementsprechend individuelle Unterschiede zwischen den Lernenden herausstellen zu können. Ebenso ist die Erforschung des Zusammenhangs des Unterrichtsenthusiasmus der Lehrkräfte und der intrinsischen Motivation in heterogenen Lerngruppen neben dem DaZ-Unterricht für weitere Unterrichtsfächer wünschenswert.

5.3. Implikationen für die Praxis

Mit Blick auf die schulische Praxis sprechen die Ergebnisse dieser Untersuchung insbesondere für ein hohes Maß des erlebten Unterrichtsenthusiasmus durch die Lehrkräfte und dessen Zusammenhang mit der intrinsischen Motivation der Lernenden in den IK. Entsprechend lässt sich speziell in der Erwachsenenbildung für eine Steigerung des affektiven Erlebens des Unterrichtsenthusiasmus der Lehrkräfte plädieren, um möglicherweise damit die intrinsische Motivation der Lernenden anzusprechen. Eine Möglichkeit zur

Steigerung des affektiven Erlebens des Unterrichtsenthusiasmus wäre beispielsweise die Förderung der Bedürfnisbefriedigung unter Berücksichtigung der Arbeitsbedingungen sowie der Zufriedenheit der Lehrkräfte (Moè & Katz, 2022), um ein entsprechendes Arbeitsumfeld und -klima zu schaffen, in dem Lehrkräfte Unterrichtsenthusiasmus sowohl erleben als auch ausdrücken.

Zudem legt der in dieser Untersuchung festgestellte, vergleichsweise niedrige, emotionale Ausdruck des Unterrichtsenthusiasmus und dessen fehlender Zusammenhang mit der intrinsischen Motivation der Lernenden die Vermutung nahe, dass ein erhöhter Ausdruck eventuell eher in Zusammenhang stehen würde. Durch Trainings und Schulungen könnten Lehrkräfte zu einem stärkeren authentischen emotionalen Ausdruck ihres Unterrichtsenthusiasmus animiert werden (z. B. Bettencourt et al., 1983; Collins, 1976; Rolider et al., 1984; Streeter, 1986). Durch regelmäßige Fortbildungen und Schulungen sowie Praktiken in der Lehrkraftausbildung wäre eine theoretische, aber vor allem praktische Schulung hinsichtlich einer passenden Intensivität des emotionalen Ausdrucks von Enthusiasmus zu überlegen. Damit könnte Lehrkräften dazu verholfen werden, bewusst Enthusiasmus für ihren Unterricht zu entwickeln, zu erleben und diesen in Form eines authentischen Ausdrucks im Unterricht strategisch einzusetzen, um gezielt die Lernprozesse ihrer Lernenden anzuregen (Wood, 1999).

Literatur

- Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2009). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV). Dokumentation der Erhebungsinstrumente* [Teachers' professional knowledge, cognitively activating mathematics teaching and the development of mathematical competence (COACTIV). Documentation of the survey instruments]. Max-Planck-Institut für Bildungsforschung.
- Becker, E. S., Keller, M. M., Goetz, T., Frenzel, A. C., & Taxer, J. L. (2015). Antecedents of teachers' emotions in the classroom: an intraindividual approach. *Frontiers in Psychology*, 6, Article 635. <https://doi.org/10.3389/fpsyg.2015.00635>
- Becker-Mrotzek, M. (2022). Sprachliche Bildung in der heterogenen Gesellschaft [Language education in a heterogeneous society]. *Leseräume-Zeitschrift für Literalität in Schule und Forschung*, 9(8), 1–11. <https://leseraeume.de/wp-content/uploads/2022/12/lr-2022-1-Becker-Mrotzek.pdf>
- Bettencourt, E. M., Gillett, M. H., Damien Gall, M., & Hull, R. E. (1983). Effects of teacher enthusiasm training on student on-task behavior and achievement. *American Educational Research Journal*, 20(3), 435–450. <https://doi.org/10.2307/1162610>
- Bleck, V. (2017). *Lehrerenthusiasmus: Entwicklung, Determinanten, Wirkungen* [Teacher enthusiasm: development, determinants, effects]. Springer. <https://doi.org/10.1007/978-3-658-23102-6>

- Bösel, B. (2008). *Philosophie und Enthusiasmus: Studien zu einem umstrittenen Verhältnis* [Philosophy and enthusiasm: studies on a controversial relationship]. Passagen-Verlag.
- Burić, I., & Mornar, M. (2023). Teacher dispositional affectivity, emotional labor, and self-efficacy: A longitudinal analysis. *Current Psychology*, 42, 18382–18395. <https://doi.org/10.1007/s12144-022-03029-7>
- Collins, M. L. (1976). *Effects of enthusiasm training on preservice elementary teachers* (ED129773). ERIC. <https://files.eric.ed.gov/fulltext/ED129773.pdf>
- Collins, M. L. (1978). Effects of enthusiasm training on preservice elementary teachers. *Journal of Teacher Education*, 29(1), 53–57. <https://doi.org/10.1177/002248717802900120>
- Dewaele, J. M., & Li, C. (2021). Teacher enthusiasm and students' social-behavioral learning engagement: The mediating role of student enjoyment and boredom in Chinese EFL classes. *Language Teaching Research*, 25(6), 922–945. <https://doi.org/10.1177/13621688211014538>
- Döring, N., Bortz, J., Pöschl, S., Werner, C. S., Schermelleh-Engel, K., Gerhard, C., & Gäde, J. C. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* [Research methods and evaluation in the social and human sciences] (5th ed.). Springer. <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1624548>
- Fauth, B., Decristan, J., Decker, A. T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education*, 86, 102882. <https://doi.org/10.1016/j.tate.2019.102882>
- Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, 101(3), 705–716. <https://doi.org/10.1037/a0014695>
- Frenzel, A. C., Becker-Kurz, B., Pekrun, R., Goetz, T., & Lüdtke, O. (2018). Emotion transmission in the classroom revisited: A reciprocal effects model of teacher and student enjoyment. *Journal of Educational Psychology*, 110(5), 628–639. <https://doi.org/10.1037/edu0000228>
- Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., & Pekrun, R. (2009). *PISA 2006 Skalenhandbuch. Dokumentation der Erhebungsinstrumente* [PISA 2006 Scale manual. Documentation of the survey instruments]. Waxmann.
- Frommelt, M. C., Schiefele, U., & Lazarides, R. (2021). Teacher Enthusiasm, Supportive instructional practices, and student motivation in mathematics classrooms. *Interdisciplinary Education and Psychology*, 2(3), Article 5. <https://doi.org/10.31532/InterdiscipEducPsychol.2.3.005>
- Gaspard, H., & Lauermaun, F. (2021). Emotionally and motivationally supportive classrooms: A state-trait analysis of lesson- and classroom-specific variation in teacher- and student-reported teacher enthusiasm and student engagement. *Learning and Instruction*, 75, Article 101494. <https://doi.org/10.1016/j.learninstruc.2021.101494>
- Hachfeld, A., Schroeder, S., Anders, Y., Hahn, A., & Kunter, M. (2012). Multikulturelle Überzeugungen [Multicultural beliefs]. *Zeitschrift für Pädagogische Psychologie*, 26(2), 101–120. <https://doi.org/10.1024/1010-0652/a000064>
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, 2(3), 96–100. <https://doi.org/10.1111/1467-8721.ep10770953>
- Hooda, M. (2017). Determinants of teacher enthusiasm. *International Journal of Research in Engineering, IT and Social Sciences*, 7(8), 1–3.
- Huangfu, Q., Li, H., Tang, S., Wang, J., Liu, Q., & Chen, G. (2022). How teacher enthusiasm affects students' learning of chemistry declarative knowledge in video lectures. *Chemistry Education Research and Practice*, 23(4), 898–912. <https://doi.org/10.1039/d2rp00095d>
- Keller, M. M., Becker, E. S., Frenzel, A. C., & Taxer, J. L. (2018). When teacher enthusiasm is authentic or inauthentic: Lesson profiles of teacher enthusiasm and relations to students' emotions. *AERA Open*, 4(2), 1–16. <https://doi.org/10.1177/2332858418782967>

- Keller, M. M., Goetz, T., Becker, E. S., & Morger, V. (2014a). Feeling and showing: A new conceptualization of dispositional teacher enthusiasm and its relation to students' interest. *Learning and Instruction*, 33, 29–38. <https://doi.org/10.1016/j.learninstruc.2014.03.001>
- Keller, M. M., Hoy, A. W., Goetz, T., & Frenzel, A. C. (2016). Teacher enthusiasm: Reviewing and redefining a complex construct. *Educational Psychology Review*, 28(4), 743–769. <https://doi.org/10.1007/s10648-015-9354-y>
- Keller, M. M., Neumann, K., & Fischer, H. E. (2013). Teacher enthusiasm and student learning. In J. Hattie & E. M. Anderman (Eds.), *Educational Psychology Handbook Ser. International Guide to Student Achievement* (pp. 247–250). Routledge.
- Keller, M. M., Neumann, K., & Fischer, H. E. (2014b). Enthusiastic teaching and its impact on students' interest and self-concept: An investigation of German physics classrooms. In H. E. Fischer (Ed.), *Quality of instruction in physics: Comparing Finland, Germany and Switzerland* (pp. 129–143). Waxmann.
- Kim, T., & Schallert, D. L. (2014). Mediating effects of teacher enthusiasm and peer enthusiasm on students' interest in the college classroom. *Contemporary Educational Psychology*, 39(2), 134–144. <https://doi.org/10.1016/j.cedpsych.2014.03.002>
- Kunter, M., Frenzel, A., Nagy, G., Baumert, J., & Pekrun, R. (2011). Teacher enthusiasm: Dimensionality and context specificity. *Contemporary Educational Psychology*, 36(4), 289–301. <https://doi.org/10.1016/j.cedpsych.2011.07.001>
- Kunter, M., & Holzberger, D. (2014). Loving Teaching. In P. W. Richardson, S. A. Karabenick, & H. M. G. Watt (Eds.), *Teacher Motivation* (pp. 83–99). Routledge. <https://doi.org/10.4324/9780203119273-6>
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820. <https://doi.org/10.1037/a0032583>
- Kunter, M., Tsai, Y. M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, 18(5), 468–482. <https://doi.org/10.1016/j.learninstruc.2008.06.008>
- Lazarides, R., Buchholz, J., & Rubach, C. (2018). Teacher enthusiasm and self-efficacy, student-perceived mastery goal orientation, and student motivation in mathematics classrooms. *Teaching and Teacher Education*, 69, 1–10. <https://doi.org/10.1016/j.tate.2017.08.017>
- Lazarides, R., & Schiefele, U. (2021). Von der Lehrermotivation zur Schülermotivation: Ein integratives Modell zur motivationalen Entwicklung im Unterricht [From teacher motivation to student motivation: An integrative model for motivational development in the classroom]. In R. Lazarides, D. Raufelder, I. Gogolin, A. Scheunflug, B. Hannover & E. Souvignier (Eds.), *Edition ZfE: Bd. 10. Motivation in unterrichtlichen Lehr-Lernkontexten: Perspektiven aus Pädagogik, Psychologie und Fachdidaktiken* (Vol. 10, pp. 3–28). Springer VS. https://doi.org/10.1007/978-3-658-31064-6_1
- Liew, T. W., Mat Zin, N. A., & Sahari, N. (2017). Exploring the affective, motivational and cognitive effects of pedagogical agent enthusiasm in a multimedia learning environment. *Human-centric Computing and Information Sciences*, 7(1), 1–21. <https://doi.org/10.1186/s13673-017-0089-2>
- Moè, A., & Katz, I. (2022). Need satisfied teachers adopt a motivating style: The mediation of teacher enthusiasm. *Learning and Individual Differences*, 99, 1–6. <https://doi.org/10.1016/j.lindif.2022.102203>
- Murphy, C. A., & Walls, R. T. (1994). *Concurrent and sequential occurrences of teacher enthusiasm behaviors* (ED375128). ERIC. <https://files.eric.ed.gov/fulltext/ED375128.pdf>
- Orosz, G., Tóth-Király, I., Bőthe, B., Kusztor, A., Kovács, Z. Ü., & Jánvári, M. (2015). Teacher enthusiasm: a potential cure of academic cheating. *Frontiers in Psychology*, 6, Article 318. <https://doi.org/10.3389/fpsyg.2015.00318>
- Palmer, D. (2020). Teacher enthusiasm and student motivation for learning. *Global Journal of Educational Studies*, 6(1), 40–49. <https://doi.org/10.5296/gjes.v6i1.16405>

- Parrisius, C., Gaspard, H., Trautwein, U., & Nagengast, B. (2020). The transmission of values from math teachers to their ninth-grade students: Different mechanisms for different value dimensions? *Contemporary Educational Psychology*, 62, Article 101891. <https://doi.org/10.1016/j.cedpsych.2020.101891>
- Patrick, B. C., Hisley, J., & Kempler, T. (2000). What's everybody so excited about? The effects of teacher enthusiasm on student intrinsic motivation and vitality. *The Journal of Experimental Education*, 68(3), 217–236. <https://doi.org/10.1080/00220970009600093>
- Raudenbush, S. W., & Bryk, A. S. (2001). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). *Advanced quantitative techniques in the social sciences* (Vol. 1). Sage Publications.
- Rolider, A., Siedentop, D., & van Houten, R. (1984). Effects of enthusiasm training on subsequent teacher enthusiastic behavior. *Journal of Teaching in Physical Education*, 3(2), 47–59. <https://doi.org/10.1123/jtpe.3.2.47>
- Rosenshine, B. (1970a). Enthusiastic teaching: A research review. *The School Review*, 78(4), 499–514. <https://www.jstor.org/stable/1084086>
- Rosenshine, B. (1970b). *Interpretive study of teaching behaviors related to student achievement. Final Report* (ED051116). ERIC. <https://files.eric.ed.gov/fulltext/ED051116.pdf>
- Sahlender, M., & ten Hagen, I. (2024). Do teachers adapt their gestures in linguistically heterogeneous second language teaching to learners' language proficiencies? *Gesture*, 22(2), 189–226. <https://doi.org/10.1075/gest.22023.sah>
- Schoofs, B. (2019). Enthusiasm in teaching. *Scholarly Papers and Articles*, Article 60. https://scholarworks.gvsu.edu/cgi/viewcontent.cgi?article=1060&context=library_sp
- Stam, G., & Tellier, M. (2022). Gesture helps second and foreign language learning and teaching. In A. Morgenstern & S. Goldin-Meadow (Eds.), *Gesture in language: Development across the lifespan*, pp. 335–363. De Gruyter Mouton; American Psychological Association.
- Streeter, B. B. (1986). The effects of training experienced teachers in enthusiasm on students' attitudes toward reading. *Reading Psychology*, 7(4), 249–259. <https://doi.org/10.1080/0270271860070403>
- Suleiman, M. F., & Hashem, M. E. (1995). *Cultural Factors Influencing the Communicative Process in the Diverse Classroom* (ED404381). ERIC. <https://files.eric.ed.gov/fulltext/ED404381.pdf>
- Taxer, J. L., & Frenzel, A. C. (2018). Inauthentic expressions of enthusiasm: Exploring the cost of emotional dissonance in teachers. *Learning and Instruction*, 53, 74–88. <https://doi.org/10.1016/j.learninstruc.2017.07.008>
- Valentín, A., Mateos, P. M., González-Tablas, M. M., & López, E. (2022). Effects of teacher enthusiasm and type of text on the motivation and achievement of schoolchildren. *Frontiers in Psychology*, 13, 1–11. <https://doi.org/10.3389/fpsyg.2022.842521>
- Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wild, T. C., Enzle, M. E., Nix, G., & Deci, E. L. (1997). Perceiving others as intrinsically or extrinsically motivated: Effects on expectancy formation and task engagement. *Personality and Social Psychology Bulletin*, 23(8), 837–848. <https://doi.org/10.1177/0146167297238005>
- Wood, A. M. (1999). *The effects of teacher enthusiasm on student motivation, selective attention, and text memory*. Dissertation Abstracts International Section A: Humanities and Social Sciences, 59(9–A), 3355.

In welchem Maße ist die Gymnasialempfehlung von Lehrkräften heute noch vom familiären Hintergrund der Schüler*innen abhängig?

Ruben Kleinkorres, Ulrich Ludewig & Ramona Lorenz

Zusammenfassung

Auch wenn die Empfehlung für die weiterführende Schule grundsätzlich auf leistungsbezogenen Kriterien beruhen sollte, wurden in der Vergangenheit wiederholt sozioökonomische und zuwanderungsbezogene Disparitäten hinsichtlich der Wahrscheinlichkeit festgestellt, eine Gymnasialempfehlung zu erhalten. Vor dem Hintergrund ständiger gesellschaftlicher und bildungspolitischer Entwicklungen und einem zunehmenden Bewusstsein für diese Problematik, war das Ziel des vorliegenden Beitrags, zu untersuchen, ob diesbezüglich – unter umfassender Berücksichtigung relevanter Kriterien für den Übergang – weiterhin Unterschiede bestehen. Auf Basis von Daten der Internationalen Grundschul-Lese-Untersuchung (IGLU) 2021 wurden $N = 4188$ Schüler*innen der vierten Klasse (49.0% weiblich) untersucht, die in solchen Bundesländern Deutschlands unterrichtet wurden, in denen am Ende der vierten Klasse eine Übergangsempfehlung für die weiterführende Schule ausgesprochen wird. Das durchschnittliche Alter lag bei 10.40 Jahren ($SD = 0.53$). Zur Beantwortung der Fragestellungen wurden logistische Regressionsmodelle spezifiziert, in die zur Vorhersage der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, sukzessive leistungsbezogene Merkmale und familiäre Hintergrundvariablen als Prädiktoren aufgenommen wurden. Die Ergebnisse wiesen auf hohe Zusammenhänge der Halbjahres-Zeugnisnoten in den Fächern Deutsch und Mathematik mit der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, hin. Darüber hinaus waren auch das durch die Lehrkraft benotete Arbeitsverhalten sowie durch objektive Tests gemessene Fähigkeiten und Kompetenzen (kognitive Fähigkeiten und Lesekompetenz) signifikant mit der Gymnasialempfehlung assoziiert. Unter Berücksichtigung der leis-

tungsbezogenen Variablen sowie der elterlichen Bildungsaspiration, die ihrerseits mit der Gymnasialempfehlung zusammenhing, wiesen Kinder mit Zuwanderungshintergrund keine abweichende Wahrscheinlichkeit dafür auf, eine Gymnasialempfehlung zu erhalten. Der sozioökonomische Status hingegen war auch unter Kontrolle aller anderen Variablen signifikant mit der Gymnasialempfehlung assoziiert. Limitationen der Studie sowie Implikationen für Forschung und Praxis werden diskutiert.

Schlagworte: Sozioökonomischer Hintergrund, Gymnasialempfehlung, leistungsbezogene Merkmale, Übergang auf die weiterführende Schule, Zuwanderungshintergrund

Abstract

Even though the recommendation for secondary school should generally be based on performance-related criteria, socio-economic and immigration-related disparities in the probability of receiving an academic track recommendation have been repeatedly identified in the past. Against the background of constant societal developments and changes in educational policy as well as an increasing awareness of this problem, the aim of this article was to investigate whether differences continue to exist in this respect, taking into account several relevant criteria for the transition. Based on data from the German subsample of the Progress in International Reading Literacy Study (PIRLS) 2021, $N = 4188$ fourth-grade students (49.0% female) were examined who were taught in those federal states in Germany where a transition recommendation for secondary school is given at the end of fourth grade. The average age was 10.40 years ($SD = 0.53$). To answer the research questions, logistic regression models were specified in which performance-related characteristics and family background variables were successively included as predictors of the probability of an academic track recommendation. The results showed high associations between half-year report card grades in German and mathematics and the probability of receiving an academic track recommendation. The work behavior graded by the teacher and achievement measured by objective tests (cognitive skills and reading literacy) were also significantly related to the academic track recommendation. Taking into account the performance-related variables and parental educational aspirations, which were also related to the academic track recommendation, children with an immigration background did not show any deviating probability of receiving an academic track recommendation. Socioeconomic status, on the other hand, was significantly associated with the academic track recommendation, even

when all other variables were controlled for. Limitations of the study and implications for research and practice are discussed.

Keywords: Socioeconomic background, academic track recommendation, performance-related characteristics, transition to secondary school, immigration background

1. Einleitung

Das deutsche Bildungssystem sieht in der Regel eine vierjährige – in wenigen Bundesländern eine sechsjährige – Grundschulzeit gefolgt von einer mehrjährigen Sekundarschulzeit vor. Am Ende der Primarstufe sprechen die Grundschullehrkräfte eine Empfehlung für die weiterführende Schulform aus, die in einigen Bundesländern verbindlich ist. Dabei entscheiden in den meisten Bundesländern in letzter Instanz zwar die Eltern, welche Schulform ihr Kind im folgenden Schuljahr besuchen wird (Bach, 2023; Autor:innengruppe Bildungsberichterstattung, 2024; Kuhn, 2023), die Empfehlung der Lehrkraft dürfte Elternentscheidungen jedoch bedeutend beeinflussen (Maaz & Nagy, 2009).

In seiner Grundidee zielt das mehrgliedrige deutsche Schulsystem auf Leistungsgruppierung ab, in der das Gymnasium die Schulform mit der höchsten Leistungsorientierung darstellt (Rost & Köller, 2004). Das bedeutet, dass Lehrkräfte Schüler*innen für das Gymnasium empfehlen sollten, wenn sie gute Leistungen in der Grundschule feststellen konnten und/oder ein hohes Leistungspotenzial der Schüler*innen für die weiterführende Schule sehen. Dem übergeordneten Ziel bildungspolitischer Anstrengungen folgend, Chancengleichheit und Bildungsgerechtigkeit zu erhöhen (KMK, 2019; Wößmann et al., 2023), sollten Kinder, die die gleichen Leistungen zeigen, mit der gleichen Wahrscheinlichkeit eine bestimmte Übergangsempfehlung erhalten, unabhängig von ihrem familiären Hintergrund. Übereinstimmend mit theoretischen Ansätzen zu Herkunftseffekten (Boudon, 1974) zeigten empirische Befunde in der Vergangenheit jedoch, dass die Einschätzung der Lehrkräfte bezüglich der Eignung von Schüler*innen für das Gymnasium vielmals auch von sozioökonomischen Faktoren abhängt, die unabhängig von der Leistung sind (Dumont et al., 2014; Maaz & Nagy, 2009).

Zwar ist die Entscheidung über die weiterführende Schule im Verlauf der Sekundarschulzeit noch korrigierbar und der Weg zum Abitur wurde durch den Ausbau integrierter Gesamtschulen erleichtert, jedoch erfolgen Schulformwechsel insgesamt eher selten und es finden häufiger Wechsel auf eine weniger leistungsorientierte Schulform statt als umgekehrt (Pfost et al.,

2018; Stubbe, Schaufelberger, et al., 2023). Darüber hinaus liegen Unterschiede nach familiärem Hintergrund vor, Gebrauch von der Möglichkeit zu machen, frühere Bildungsverlaufsentscheidungen zu revidieren (Biewen & Tapalaga, 2017). Daher bietet der Besuch eines Gymnasiums nach wie vor eine besonders gute Voraussetzung für eine erfolgreiche akademische Laufbahn (Stubbe, Schaufelberger, et al., 2023; Wößmann et al., 2023). Entsprechend bedeutsam ist es, dass Lehrkräfte zu einer realistischen und fairen Prognose hinsichtlich der zukünftigen Schullaufbahn gelangen.

Der Übergang von der Grundschule auf die Sekundarstufe wird seit den 1960er Jahren beforscht (Gresch et al., 2010). Mit Aufkommen der Large-Scale Untersuchungen ab den späten 1990er Jahren intensivierte sich die Forschung zu diesem Thema (Dumont et al., 2014) und die oben beschriebene Benachteiligung wurde wiederholt bestätigt. Die regelmäßige Untersuchung ist dabei sinnvoll, weil sich das Bildungssystem und die Umstände in Deutschland verändert haben und stetig verändern. So ist beispielsweise der Anteil an Bundesländern mit hohem Verbindlichkeitsgrad der Übergangsempfehlung der Lehrkräfte zurückgegangen (vgl. Bach, 2023; Tabelle 1) und es gibt vielfältige Möglichkeiten, das Abitur auch ohne eine Gymnasialempfehlung zu erlangen (Stubbe, Schaufelberger, et al., 2023). Weiterhin hat sich die Schülerschaft hinsichtlich ihrer familiären Hintergründe verändert: Der Anteil an Kindern mit Zuwanderungshintergrund ist gestiegen (Henschel et al., 2022) und es sind in den letzten Jahren viele Familien zugewandert, die aus ihrem Heimatland fliehen mussten (Feld et al., 2017). Auf der anderen Seite hat sich der Zugang zu Bildung in den letzten Jahrzehnten gewandelt, was zur Folge hat, dass heute mehr Eltern als früher einen tertiären Bildungsabschluss besitzen (Statistisches Bundesamt, 2022). Zuletzt ist anzumerken, dass die Problematik der sozioökonomischen und zuwanderungsbezogenen Disparitäten bei der Gymnasialempfehlung durch die zunehmende Beforschung über die Wissenschaft hinaus vielen Lehrkräften zugänglich gemacht wurde. So wäre es möglich, dass heute ein stärkeres Bewusstsein für diese Problematik besteht als früher, weil sie beispielsweise in der Lehrkräfteausbildung thematisiert wurde. Eine aktuelle Untersuchung der Abhängigkeit der Gymnasialempfehlung von familiären Merkmalen erscheint vor dem Hintergrund dieser dynamischen Entwicklungen angemessen.

Der vorliegende Beitrag untersuchte daher, welche Rolle der familiäre Hintergrund und die Bildungsaspiration der Eltern aktuell dabei spielen, eine Gymnasialempfehlung zu erhalten. Dabei wurde im Vorhinein sukzessive der Erklärungswert bewerteter schulischer Leistungen (Halbjahres-Zeugnisnoten), bewerteter Verhaltensweisen im schulischen Kontext (Arbeits- und Sozialverhalten) und objektiver Leistungen (kognitive Fähigkeiten, Lesekom-

petenz) untersucht, um ein breites Spektrum geeigneter Faktoren abzudecken, die in die Bewertung von Lehrkräften über die Eignung eines Kindes für den Besuch des Gymnasiums einfließen könnten.

2. Kriterien für eine Übergangsempfehlung

Um Chancengleichheit zu wahren und dem Leistungsprinzip gerecht zu werden, sollte die Empfehlung der Lehrkräfte, welche Schulform ein Kind ihrer Klasse nach der Grundschule besuchen soll, auf Faktoren basieren, die den Schulerfolg besonders zuverlässig vorhersagen können, damit jedes Kind den Bildungsweg verfolgen kann, der dessen Bildungsfähigkeit entspricht (KMK, 2015). Kriterien, die in den meisten Ländern der Bundesrepublik Deutschland als Basis für die Empfehlung für eine weiterführende Schule festgelegt sind, sind die Leistungen, die die Schüler*innen während der Grundschulzeit erbracht haben, ihr Lern- und Arbeitsverhalten (teilweise auch ihr Sozialverhalten; z. B. in Niedersachsen) sowie die Lern- und Leistungsentwicklung und das daraus abgeleitete Entwicklungspotenzial (KMK, 2015). In einigen Bundesländern wird zudem ein gewisser Notendurchschnitt für eine Gymnasialempfehlung vorausgesetzt (KMK, 2015; Kuhn, 2023).

Folgerichtig wurden in vergangenen Untersuchungen zur Vorhersage der Gymnasialempfehlung zuvorderst die Schulnoten in den Hauptfächern Deutsch und Mathematik herangezogen, um die Empfehlung der Lehrkräfte für eine bestimmte weiterführende Schule vorherzusagen (Brändle & Weirich, 2023; Maaz & Nagy, 2009; Stubbe, Schaufelberger, et al., 2023; für einen Überblick siehe Glock et al., 2013). Da die Schulnoten jedoch nicht ausschließlich durch die objektive Leistung von Schüler*innen bedingt, sondern beispielsweise an das Leistungsniveau der Klasse angepasst sein können (Glock et al., 2013), wurde in bisherigen Studien zusätzlich häufig die Leistung von Kindern in standardisierten Tests herangezogen (Brändle & Weirich, 2023; Maaz & Nagy, 2009; Stubbe, Schaufelberger, et al., 2023). Den in den einzelnen Bundesländern festgelegten Kriterien für eine Übergangsempfehlung folgend, sollten weiterhin lernrelevante Eigenschaften und Verhaltensweisen der Schüler*innen (Arbeits- und Sozialverhalten) und deren Zusammenhänge mit der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, untersucht werden (Klapproth et al., 2013; Lintorf et al., 2021).

2.1. Bewertete schulische Leistungen und Verhaltensweisen

Schulnoten

Die schulischen Leistungen in der Grundschule sollten ein zentrales Kriterium für die Übergangsempfehlung der Lehrkräfte sein (Ditton, 2010; KMK, 2015). Fähigkeiten und Verhaltensweisen, die Schüler*innen in der Grundschule gezeigt haben, um gute Noten zu erhalten, sollten auch für einen erfolgreichen Besuch des Gymnasiums ausschlaggebend sein. Insbesondere die Zeugnisnoten in den Hauptfächern Deutsch und Mathematik können den Lehrkräften dabei als Richtschnur für ihre Empfehlung dienen, da sie dokumentiert sind und auf früheren schriftlichen Beiträgen der Schüler*innen sowie deren Beteiligung im Unterricht basieren (Stubbe et al., 2017). In den meisten Bundesländern ist schriftlich festgehalten, welche Empfehlung für die weiterführende Schule ausgesprochen werden sollte, je nachdem welche Noten in diesen beiden Schulfächern erreicht wurden (KMK, 2015).

Verschiedene empirische Studien legen nahe, dass die Übergangsempfehlung (Anders et al., 2010; Brändle & Weirich, 2023; Maaz & Nagy, 2009) beziehungsweise Gymnasialpräferenz (d. h. die Prognose der Lehrkräfte, ob ein bestimmtes Kind nach der Grundschule auf das Gymnasium wechseln wird; Stubbe et al., 2012; Stubbe, Schaufelberger, et al., 2023) von Lehrkräften in hohem Maße von den Schulnoten der Schüler*innen abhängt. Glock et al. (2013) berichteten basierend auf verschiedenen Studien, dass der Anteil aufgeklärter Varianz durch die Schulnoten in Deutsch und Mathematik zwischen 66% und 77% liegt.

Arbeits- und Sozialverhalten

Während gut belegt ist, dass die Schulnoten die Gymnasialempfehlung durch die Lehrkräfte zu großen Teilen erklären können, wurden andere Merkmale der Schüler*innen, die durch die Lehrkräfte bewertet werden und den zukünftigen Schulerfolg erklären können, in diesem Zusammenhang seltener untersucht. Die Hinzunahme weiterer potenziell relevanter Variablen kann aber sinnvoll sein, um die Varianz in der Gymnasialempfehlung der Lehrkräfte systematisch aufzuklären und besser verstehen zu können, auf welcher Grundlage die Lehrkräfte ihr Urteil fällen. Das Arbeits- und Sozialverhalten der Schüler*innen könnten hierbei insofern relevante Variablen darstellen, als dass sie in einigen Bundesländern als Kriterien für die Übergangsempfehlung gelten (KMK, 2015). Das Arbeitsverhalten der Schüler*innen beschreibt zum einen deren Motivation, sich gestellten Aufgaben zuzuwenden und zum anderen die Ausdauer, Konzentration und Sorgfalt, diese im Anschluss zu

bearbeiten (Pröscholdt et al., 2011). Aus dieser Definition wird ersichtlich, dass das Arbeitsverhalten eine wichtige Rolle für die Prognose der zukünftigen akademischen Laufbahn spielen kann. Es ist jedoch zu prüfen, ob Lehrkräfte diese Eigenschaften unabhängig von den fächerspezifischen Schulnoten gewichten und in ihre Übergangsempfehlung einfließen lassen, da naheliegt, dass ein sorgfältiges und konzentriertes Arbeitsverhalten sich bereits in den Schulnoten niedergeschlagen hat.

Neben dem Arbeitsverhalten eignet sich das Sozialverhalten der Schüler*innen als möglicher weiterer Indikator für die Qualifikation für die zukünftige Schulform. Das Sozialverhalten zeichnet sich beispielsweise durch das Verhalten der Schüler*innen gegenüber Mitschüler*innen, ihre Hilfsbereitschaft und ihre Kooperationsfähigkeit aus (Böhmer et al., 2015). Die Fähigkeit mit anderen zu kooperieren, zu diskutieren, ihre Perspektiven einzunehmen und Konfliktlösungen zu finden, kann damit einhergehen, sich auch mit höherer Motivation intellektuellen Problemen zu widmen (Wentzel et al., 2019). Darüber hinaus kann die Fähigkeit zur Kooperation sich zusätzlich positiv auf das Lernen auswirken (Slavin, 2011; Wentzel et al., 2021). Soziale Kompetenzen von Schüler*innen könnten somit von Lehrkräften herangezogen werden, um deren Leistungspotenzial und das Bestehen auf einer höheren Schulform abzuschätzen. Gleichzeitig könnte sich das Sozialverhalten auf das Verhältnis zu den Lehrkräften auswirken und diese in ihrer Tendenz beeinflussen, eine Gymnasialempfehlung auszusprechen (Timmermans et al., 2019). Auf diese Weise könnte das Sozialverhalten auch einen Teil der familiären Herkunft erklären, da ein bestimmter erlernter Habitus in Familien mit hohem sozioökonomischen Status Einfluss darauf haben könnte, wie die Lehrkraft das Sozialverhalten einer Schülerin oder eines Schülers wahrnimmt (El-Mafaalani, 2014; Kramer, 2017). Ähnlich wie das Arbeitsverhalten könnte sich aber auch das Sozialverhalten bereits auf die fächerspezifischen Schulnoten auswirken, sodass zu prüfen ist, ob es über die Schulnoten hinaus mit der Gymnasialempfehlung zusammenhängt.

Vergangene Untersuchungen konnten zeigen, dass Lehrkräfte sowohl das Arbeits- als auch das Sozialverhalten als wichtige Indikatoren für die Eignung für die zukünftige Schulform erachten (Anders et al., 2010; Böhmer et al., 2015; Gräsel et al. 2010). Quantitative Studien wiesen dabei darauf hin, dass, auch unter Kontrolle von Leistungsindikatoren, positive Zusammenhänge zwischen dem Arbeitsverhalten und der Gymnasialempfehlung bestehen (Anders et al., 2010; Lintorf et al., 2021; Stubbe et al., 2017; Stubbe et al., 2020). So berichteten Anders et al. (2010) über einen inkrementellen Erklärungswert von Lerntugenden wie der Anstrengungsbereitschaft, der Gewissenhaftigkeit und des Fleißes bei gleichzeitiger Betrachtung von Leistungsmerkmalen der

Schüler*innen. Weiterhin konnten Klapproth et al. (2013) basierend auf einer luxemburgischen Stichprobe bei gleichzeitiger Berücksichtigung von verschiedenen Leistungsindikatoren und familiären Hintergrundmerkmalen einen Zusammenhang zwischen der Zuverlässigkeit und Sorgfalt von Schüler*innen und der Wahrscheinlichkeit, eine Empfehlung für die höchste Schulform zu erhalten, nachweisen. Im Hinblick auf das Sozialverhalten berichteten Anders et al. (2010) zwar erwartungsgemäß eine positive Korrelation mit der Gymnasialempfehlung, jedoch ein negatives Regressionsgewicht, wenn zugleich Leistungs- und Motivationsmerkmale berücksichtigt werden. Klapproth et al. (2013) berichteten ebenfalls eine positive bivariate Korrelation zwischen dem Sozialverhalten und der Empfehlung für die höchste Schulform, wobei dieser Zusammenhang im Regressionsmodell bei gleichzeitiger Betrachtung der Schulnoten, der Leistung in standardisierten Tests und des Arbeitsverhaltens verschwand.

Insgesamt liegen einige Belege dafür vor, dass das Arbeitsverhalten auch unter Kontrolle von Leistungsmerkmalen mit der Gymnasialempfehlung zusammenhängt. Das Sozialverhalten wurde hingegen weniger stark beforscht und die bisherigen Befunde deuten darauf hin, dass es über die Leistung hinaus keinen Zusammenhang mit der Gymnasialempfehlung aufweist.

2.2. Kognitive Leistungsfähigkeit und Lesekompetenz als grundlegende Fähigkeiten für Leistungsentwicklung

Fächerspezifische Schulnoten und die Noten für im Schulkontext gezeigte Verhaltensweisen bieten eine dokumentierte Orientierung für den Leistungsstand der Schüler*innen und geben den Lehrkräften die Möglichkeit, ihre Übergangsempfehlung vor Eltern und Schüler*innen zu rechtfertigen. Jedoch können die Schulnoten durch Referenzgruppeneffekte, wie beispielsweise die durchschnittliche Leistung in einer Klasse, beeinflusst sein (Glock et al., 2013), was dazu führen kann, dass bei gleicher individueller Leistungsstärke Schüler*innen in leistungsschwächeren Klassen bessere Noten erhalten als Schüler*innen in leistungsstärkeren Klassen (Klapproth et al., 2013). Standardisierte Fähigkeits- und Kompetenztestungen können in diesem Zusammenhang Aufschluss über den objektiven Leistungsstand von Schüler*innen geben.

Insbesondere die kognitiven Fähigkeiten und die Lesekompetenz spielen eine zentrale Rolle für die Beurteilung des zukünftigen schulischen Erfolgs: Kognitive Fähigkeiten begünstigen sowohl mathematische als auch lesebezogene Fertigkeiten (Liu et al., 2021) und gelten folglich als starker Prädiktor für

zukünftigen akademischen wie auch beruflichen Erfolg (Rhode & Thompson, 2007). Weiterhin nimmt die Lesekompetenz eine besondere Rolle ein, da sie es den Schüler*innen ermöglicht, dem Unterricht auf der weiterführenden Schule angemessen folgen zu können (Glock et al., 2013) und es ihnen erleichtert, Zugang zu neuem Wissen zu erhalten (Cromley, 2009; McElvany et al., 2023; Savolainen et al., 2008). Die Lesekompetenz ist entsprechend wichtig für das erfolgreiche Lernen im weiteren Verlauf des schulischen Werdegangs. Im Unterschied zu Schulnoten liegen den Lehrkräften meist keine konkreten Leistungskennwerte für die allgemeinen kognitiven Fähigkeiten (z. B. Ergebnisse eines Intelligenztests) vor. Ähnliches gilt für die Lesekompetenz, da standardisierte Verfahren in der Praxis eher selten eingesetzt werden (Schilcher et al., 2023) und die Ergebnisse von Large-Scale Assessment Studien wie der *Internationalen Grundschul-Lese-Untersuchung* den Lehrkräften zum Zeitpunkt, zu dem die Übergangsempfehlung ausgesprochen wird, nicht bekannt sind und ihnen zudem nicht personenspezifisch mitgeteilt werden. Wie bereits dargelegt wurde, ist es als wahrscheinlich anzunehmen, dass sich kognitive Fähigkeiten und Lesekompetenz auf aktuelle und zukünftige schulische Leistungen auswirken und somit zu einem gewissen Teil bereits in die Schulnoten, die in der Grundschule erzielt wurden, eingeflossen sind. Eine Vielzahl an empirischen Studien legt jedoch einen eigenständigen Erklärungswert von objektiven Leistungen über die Schulnoten hinaus nahe (Anders et al., 2010; Brändle & Weirich, 2023; Klapproth et al., 2013; Neugebauer, 2011; Stubbe, Schaufelberger, et al., 2023).

3. Rolle des familiären Hintergrunds beim Übergang auf die weiterführende Schule

Neben den Faktoren, die direkt mit der Leistung zusammenhängen, konnten in verschiedenen Studien auch Zusammenhänge mit familiären Merkmalen, insbesondere dem sozioökonomischen Hintergrund der Kinder, festgestellt werden (Maaz & Nagy, 2009; Stubbe et al., 2012; Stubbe, Schaufelberger, et al., 2023; Wößmann et al., 2023). Als Gründe für die Benachteiligung nach familiärem Hintergrund werden häufig primäre und sekundäre Herkunftseffekte angeführt (Dumont et al., 2014; Gresch et al., 2010; Dumont et al., 2019).

3.1. Primäre und sekundäre Herkunftseffekte

Nach Boudon (1974) bezeichnen *primäre Herkunftseffekte* Disparitäten, die auf Unterschieden im sozialen, ökonomischen und kulturellen Kapital basieren und sich in tatsächlichen akademischen Leistungsunterschieden niederschlagen. Als dahinterliegender Mechanismus wird vermutet, dass Kinder aus Familien, die in Bezug auf diese Faktoren privilegiert sind, ein anregenderes familiäres Umfeld sowie mehr materielle Ressourcen im Elternhaus vorfinden und günstigere Förderung erhalten als Kinder aus in dieser Hinsicht benachteiligten Familien (Klapproth et al., 2013; Maaz & Nagy, 2009). Diese Herkunftseffekte können zu einem Kompetenzvorsprung führen und die weitere Kompetenzentwicklung begünstigen, indem verfügbare schulische Ressourcen besser genutzt werden. *Sekundäre Herkunftseffekte* gehen über den Einfluss des Elternhauses auf Kompetenzunterschiede und Differenzen in der Kompetenzentwicklung hinaus und beschreiben den zusätzlichen Einfluss des familiären Hintergrunds durch unterschiedliche Erwartungen der Eltern an die akademische Laufbahn ihrer Kinder (Boudon, 1974). So haben Eltern mit einem höheren sozioökonomischen Status häufig ausgeprägtere Bildungsaspirationen, also höhere Ambitionen hinsichtlich des Bildungsgrades, den ihr Kind erreichen soll (Paulus & Blossfeld, 2007). Ein in diesem Zusammenhang häufig angeführtes Motiv ist das des Statuserhalts (Dumont et al., 2014; Stocké, 2007). Demnach streben sozioökonomisch wohlhabende Eltern danach, dass ihre Kinder einen Bildungsgrad und in der Folge auch beruflichen Status erreichen, der ihrem eigenen entspricht. Weiterhin lassen sich unterschiedliche Bildungsaspirationen auch mit Wert-Erwartungstheorien erklären (z. B. Breen & Goldthorpe, 1997): Eltern aus sozioökonomisch privilegierten Verhältnissen schreiben der Bildung einen hohen Wert und Nutzen zu und sind daher motivierter, ihrem Kind eine Schullaufbahn auf dem Gymnasium zu ermöglichen, während sozioökonomisch benachteiligte Eltern die Kosten dieses längeren Bildungswegs stärker gewichten und daher seltener eine höhere Schulform präferieren (Gresch et al., 2010).

Maaz und Nagy (2009) schlugen eine Ausdifferenzierung dieser Theorie vor, um primäre und sekundäre Herkunftseffekte spezifischer quantifizieren zu können. In ihrem Modell unterscheiden sie zwischen primären und sekundären Herkunftseffekten auf die Leistungsbeurteilung, die Laufbahnbeurteilung und das Übergangsverhalten. Basierend auf ihrem Modell lässt sich der Einfluss primärer Herkunftseffekte auf die Gymnasialempfehlung zum einen dadurch erklären, dass objektive Leistungen (siehe Kapitel 2.2), die bereits durch die Herkunft bedingt sind, in die Schulnoten eingehen. Zum anderen wird den objektiven Leistungen ein direkter Einfluss auf die Übergangsempfehlung zugeschrieben. Während sich der sekundäre Herkunftseffekt nach

der klassischen Definition insbesondere im Übergangsverhalten der Eltern beziehungsweise in der Abweichung der Übergangsentscheidung der Eltern von der Übergangsempfehlung der Lehrkräfte äußert, spiegeln sich sekundäre Herkunftseffekte nach dem Modell von Maaz und Nagy (2009) bereits in der Notenvergabe und der Übergangsempfehlung wider. Als Grund dafür werden unterschiedliche Bewertungsmaßstäbe der Lehrkräfte je nach sozioökonomischem Hintergrund der Kinder vermutet (Maaz & Nagy, 2009). Weiterhin berichteten Gräsel et al. (2010) in diesem Zusammenhang auf Grundlage von Lehrkraftbefragungen, dass diese teilweise die Unterstützung aus dem Elternhaus in ihre Übergangsempfehlung mit einbeziehen. Darüber hinaus wäre auch denkbar, dass Lehrkräfte durch den direkten Austausch mit den Eltern eine Vorstellung von deren Motivation und Bildungsaspiration erhalten und dadurch in ihrer Tendenz beeinflusst werden, eine bestimmte Übergangsempfehlung auszusprechen (Dumont et al., 2014; Dumont et al., 2019). Basierend auf den Ergebnissen einer qualitativen Erhebung stellten Dumont et al. (2019) in diesem Zusammenhang die Vermutung an, dass Eltern mit einem hohen sozioökonomischen Status zu einem früheren Zeitpunkt und regelmäßiger Kontakt zu Lehrkräften suchen und diesen subtil vermitteln, welche zukünftige Schulform sie sich für ihr Kind wünschen. Die Studienergebnisse wiesen zudem darauf hin, dass auch ambitionierte Eltern mit einem niedrigen sozioökonomischen Status den Lehrkräften ihre Bildungsaspirationen mitteilen, dies jedoch meist zu einem späteren Zeitpunkt tun (Dumont et al., 2019).

Sowohl für primäre als auch für sekundäre Herkunftseffekte konnten in der Vergangenheit empirische Belege vorgebracht werden. So wurde regelmäßig über Leistungsunterschiede zwischen Kindern aus sozioökonomisch benachteiligten und privilegierten Elternhäusern berichtet (Ehmke & Jude, 2010; McElvany et al., 2023). Maaz und Nagy (2009) konnten die Effekte des sozioökonomischen Status weiter zerlegen und nachweisen, dass der sozioökonomische Hintergrund die Schulnoten – sowohl direkt als auch indirekt über die objektive Leistung – beeinflusst und darüber hinaus sowohl Einfluss auf die Übergangsempfehlung als auch auf das Übergangsverhalten nimmt. Weiterhin berichteten in der Vergangenheit verschiedene Autor*innengruppen über Unterschiede zwischen Kindern unterschiedlichen familiären Hintergrunds in der Wahrscheinlichkeit, eine Gymnasialempfehlung beziehungsweise -prognose zu erhalten, auch wenn deren Leistung und Schulnoten vergleichbar waren (Müller-Benedict, 2007; Stubbe et al., 2012; Stubbe, Schaufelberger, et al., 2023; van Leest et al., 2021). Brändle und Weirich (2023) kamen hingegen jüngst bei der Untersuchung von Hamburger Schulen zu dem Ergebnis, dass der sozioökonomische Status unter Kontrolle von objektiv gemessenen Leistungen und Schulnoten nicht mit der Gymnasialempfehlung

der Lehrkräfte zusammenhing. Hinsichtlich der Bildungsaspiration der Eltern fanden Dumont et al. (2019), dass diese auch unter Kontrolle von objektiver Leistung und Schulnoten mit der Gymnasialempfehlung zusammenhing.

3.2. Zuwanderungsbezogene Disparitäten

Auch in Bezug auf den Zuwanderungshintergrund liegen Disparitäten beim Übergang auf die weiterführende Schule vor (Wößmann et al., 2023). Diese lassen sich jedoch im Wesentlichen auf zwei Gründe zurückführen, die bereits Einfluss auf die Übergangsempfehlung durch die Lehrkräfte haben können: Zum einen liegen bei Kindern mit Zuwanderungshintergrund häufig geringere Kenntnisse in der Unterrichtssprache vor als bei Kindern ohne Zuwanderungshintergrund (Esser, 2006; Relikowski et al., 2010; Stubbe, Kleinkorres, et al., 2023) und zum anderen weisen Familien mit Zuwanderungshintergrund meist zusätzlich einen niedrigeren sozioökonomischen Status auf (Dumont et al., 2014; Relikowski et al., 2010). Beide Gründe können dazu führen, dass geringere schulische Leistungen erreicht werden, was dann zu einer niedrigeren Wahrscheinlichkeit führt, eine Gymnasialempfehlung zu erhalten. Entsprechend lassen sich zuwanderungsbedingte Unterschiede beim Übergang auf die weiterführende Schule weitgehend auf primäre Herkunftseffekte zurückführen (Dumont et al., 2014; Relikowski et al., 2010).

Darüber hinaus kommen auch bei Kindern mit Zuwanderungshintergrund sekundäre Herkunftseffekte zum Tragen, die jedoch anders als bei Kindern aus sozioökonomisch benachteiligten Familien nicht negativ, sondern positiv ausfallen. Diese Bevorteilung von Kindern mit Zuwanderungshintergrund zeigt sich insbesondere bei der Übergangsentscheidung der Eltern (Dumont et al., 2013; Relikowski et al., 2010), es liegt jedoch auch Evidenz dafür vor, dass Kinder mit Zuwanderungshintergrund bei gleichen Leistungen bevorzugt Gymnasialempfehlungen von den Lehrkräften erhalten (Arnold et al., 2007; Gresch, 2012; Nishen et al., 2022). Diese Befunde wurden zum einen mit unterschiedlichen Standards begründet, die die Lehrkräfte gegenüber Kindern mit und ohne Zuwanderungshintergrund anlegen (Nishen et al., 2022); zum anderen wurden sie häufig damit begründet, dass zugewanderte Eltern höhere Bildungsaspirationen aufweisen als solche ohne Zuwanderungshintergrund (Becker & Gresch, 2016; Dumont et al., 2014; Astleithner et al., 2021). Für diese unterschiedlichen Bildungsaspirationen liegen verschiedene Theorien vor (vgl. Astleithner et al., 2021). Beispielsweise erklärt die These des Zuwanderungsoptimismus unterschiedliche Bildungsaspirationen von zugewanderten Eltern damit, dass es sich bei den zugewanderten Eltern

um eine selektierte Gruppe handelt, die das Risiko einzugehen bereit ist, in einem anderen Land Fuß zu fassen, um ihre gesellschaftliche Position zu verbessern (Kao & Tienda, 1995; Hunkler & Tjaden, 2018). Weitere Theorien begründen die höhere Aspiration damit, dass zugewanderte Eltern die Leistungen ihrer Kinder auf Grund fehlenden Wissens über das Bildungssystem des Aufnahmelandes überschätzen (Informationsdefizitthese; Gresch, 2012; Kao & Tienda, 1998) oder weil sie sich für ihre Kinder erhoffen, in einem höherqualifizierten Bereich weniger Diskriminierung zu erfahren (These der antizipierten Diskriminierung; Heath & Brinbaum, 2007). Da die Bildungsaspiration der Eltern in diesen Theorien als Motor dafür angesehen wird, dass Unterschiede hinsichtlich der erhaltenen Übergangsempfehlung und des Übergangsverhaltens je nach Zuwanderungshintergrund vorliegen, sollten hier, hinsichtlich des Einflusses auf die Gymnasialempfehlung der Lehrkräfte, ähnliche Mechanismen greifen, wie sie in Kapitel 3.1 bereits für die sekundären Herkunftseffekte nach sozioökonomischem Status beschrieben wurden. Es muss jedoch beachtet werden, dass es sich bei der Gruppe von Kindern mit Zuwanderungshintergrund, um eine sehr heterogene Gruppe handelt. Hinsichtlich der These des Zuwanderungsoptimismus ist beispielsweise zu berücksichtigen, dass nicht alle Familien ihr Heimatland verlassen, weil sie nach einer Verbesserung ihrer gesellschaftlichen Stellung streben, sondern weil sie aus anderen Gründen dazu gezwungen sind (z. B. Flucht auf Grund von Unruhen). Im Hinblick auf die Gymnasialempfehlung durch die Lehrkräfte macht sich die ethnische Heterogenität auch in den Chancen bemerkbar, eine Empfehlung für das Gymnasium zu erhalten (Gresch, 2012). Darüber hinaus fanden einige Studien keinen Einfluss des Zuwanderungshintergrundes auf die Schullaufbahneempfehlung der Lehrkräfte, wenn die soziale Herkunft und die Schulleistungen der Schüler*innen berücksichtigt wurden (Dumont et al., 2013; Relikowski et al., 2010).

4. Fragestellungen

Basierend auf dem vorgestellten theoretischen Hintergrund nahmen wir an, dass insbesondere die fächerspezifischen Schulnoten (Anders et al., 2010; Brändle & Weirich, 2023; Glock et al., 2013; Maaz & Nagy, 2009; Stubbe, Schaufelberger, et al., 2023) mit der Übergangsempfehlung der Lehrkräfte zusammenhängen sollten. Darüber hinaus wurden das Arbeits- und das Sozialverhalten der Schüler*innen als weitere Prädiktoren untersucht, da diese Verhaltensweisen in einigen Ländern der Bundesrepublik Deutschland als Kriterien für eine Übergangsempfehlung festgelegt sind (KMK, 2015).

Dabei erwarteten wir basierend auf bisherigen Studien zumindest für das durch die Lehrkräfte bewertete Arbeitsverhalten einen Zusammenhang mit der Gymnasialempfehlung (z. B. Klapproth et al., 2013; Lintorf et al., 2021). Wir erwarteten zudem einen positiven Zusammenhang zwischen der objektiv gemessenen Leistung (Lesekompetenz und kognitiven Fähigkeiten) der Schüler*innen und der Gymnasialempfehlung (Anders et al., 2010; Brändle & Weirich, 2023; Klapproth et al., 2013; Neugebauer, 2011; Stubbe, Schaufelberger, et al., 2023). Auf Grundlage verschiedener gesellschaftlicher und bildungspolitischer Entwicklungen, die sich in den letzten Jahren vollzogen haben, war zudem von Interesse zu untersuchen, wie der familiäre Hintergrund mit der Gymnasialempfehlung zusammenhängt. Während in der Vergangenheit wiederholt über sozioökonomische und zugewanderungsbezogene Disparitäten berichtet wurde (Maaz & Nagy, 2009; Stubbe et al., 2012; Stubbe, Schaufelberger, et al., 2023; Wößmann et al., 2023), weisen jüngste Ergebnisse auf Grundlage von Daten einzelner Länder der Bundesrepublik Deutschlands darauf hin, dass der sozioökonomische Status nicht mit der Wahrscheinlichkeit zusammenhängt, eine Gymnasialempfehlung zu erhalten, wenn die Leistung der Schüler*innen berücksichtigt wird (Brändle & Weirich, 2023; Dumont et al., 2019). Eine deutschlandweite, aktuelle Untersuchung des Zusammenhangs zwischen der Gymnasialempfehlung und dem familiären Hintergrund scheint vor dem Hintergrund von dynamischen Entwicklungen wie der Veränderung des Verbindlichkeitsgrads der Übergangsempfehlung der Lehrkräfte (vgl. Bach, 2023), vielfältigeren Möglichkeiten, das Abitur auch ohne eine Gymnasialempfehlung zu erlangen (Stubbe, Schaufelberger, et al., 2023), eines gestiegenen Anteils an Kindern mit Zuwanderungshintergrund (Henschel et al., 2022) sowie Eltern mit einem tertiären Bildungsabschluss (Statistisches Bundesamt, 2022) und dem zunehmenden Bewusstsein der Problematik der sozioökonomischen und zugewanderungsbezogenen Disparitäten bei der Gymnasialempfehlung angemessen. Entsprechend wurden die folgenden Forschungsfragen formuliert:

1. Welchen Erklärungswert besitzen die Halbjahres-Zeugnisnoten in Deutsch und Mathematik sowie das Arbeits- und das Sozialverhalten der Schüler*innen für die Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten?
2. Besitzen objektiv gemessene Leistungen (kognitive Leistungsfähigkeit und Lesekompetenz) über die Halbjahres-Zeugnisnoten in Deutsch und Mathematik sowie das Arbeits- und das Sozialverhalten hinaus einen Erklärungswert für die Gymnasialempfehlung der Lehrkräfte?

3. Hat der familiäre Hintergrund (Zuwanderungshintergrund und sozioökonomischer Hintergrund) über Leistungsbewertungen der Lehrkraft und objektiv gemessene Leistungen hinaus einen Erklärungswert für die Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten?
4. Kann die Bildungsaspirationen der Eltern etwaige zugewanderungsbezogene und sozioökonomische Unterschiede hinsichtlich der Chancen, eine Gymnasialempfehlung zu erhalten, erklären?

5. Methode

5.1. Stichprobe

Grundlage für die interessierenden Fragestellungen bildeten Daten der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) 2021 (für einen Überblick über die Erhebung und Ergebnisse der deutschen Teilstichprobe der Studie, die international unter dem Namen *Progress in International Reading Literacy Study* bekannt ist, siehe McElvany et al., 2023). IGLU fokussiert auf die Lesekompetenz von Viertklässler*innen und findet seit 2001 in einem fünfjährigen Zyklus statt. Darüber hinaus werden aber auch Hintergrundinformationen der Schüler*innen bezüglich ihrer familiären und schulischen Situation erhoben. Die aktuelle Erhebung wurde im Jahr 2021 durchgeführt und umfasst Daten von 4611 Schüler*innen (48.9% weiblich) der vierten Klassen von 252 Schulen. Das mittlere Alter der Kinder in der Stichprobe lag bei 10.40 Jahren ($SD = 0.53$). In Bezug auf den sozioökonomischen Hintergrund wies die Stichprobe einen mittleren HISEI (Highest International Socio-Economic Index of Occupational Status; Ganzeboom, 2010; Ganzeboom et al., 1992) von 52.79 ($SD = 20.65$) auf und liegt damit hinsichtlich dieses Merkmals auf einem ähnlichen, wenn auch leicht höheren, Niveau wie jenem anderer Large-Scale Assessment-Studien, die in jüngster Vergangenheit repräsentative Ergebnisse zu Grundschul- und Sekundarschulkindern berichtet haben (Lewalter et al., 2023; Stanat et al., 2022; Stanat et al., 2023). Weiterhin wiesen 38.3% der Schüler*innen einen Zuwanderungshintergrund auf (13.5% mit einem im Ausland geborenen Elternteil und 24.9% mit zwei im Ausland geborenen Elternteilen). Hinsichtlich dieses Merkmals weist die Stichprobe

den gleichen prozentualen Anteil auf, der jüngst in einer anderen repräsentativen Studie über Viertklässler*innen berichtet wurde (Henschel et al., 2022)¹.

Für die vorliegende Untersuchung wurde auf Schüler*innen fokussiert, die in Ländern der Bundesrepublik Deutschland zur Schule gingen, in denen am Ende der vierten Klasse eine Übergangsempfehlung ausgesprochen wird. Die finale Stichprobe umfasste daher 4188 Viertklässler*innen (49.0% weiblich) an 230 Schulen. Das mittlere Alter in der Stichprobe lag wie bei der Gesamtstichprobe bei 10.4 Jahren ($SD = 0.53$) und auch die übrigen demographischen Daten waren mit der Gesamtstichprobe vergleichbar (siehe Tabelle 1).

5.2. Instrumente

Gymnasialempfehlung

Die Schullaufbahneempfehlung wurde durch ein Item („Welche Empfehlung für die weiterführende Schule hat das Kind erhalten?“) erhoben. Die Lehrkräfte sollten dabei von verschiedenen Optionen die zutreffende ankreuzen. Die Variable wurde dichotomisiert, um die Unterscheidung zwischen einer Gymnasialempfehlung (1) und der Empfehlung für eine andere Schulform (0) herzustellen.

Schulnoten und Angaben zu Arbeits- und Sozialverhalten

Die Schulnoten flossen in Form von Angaben durch die Lehrkräfte in die Analyse ein. Es handelte sich um die Halbjahres-Zeugnisnoten der Schüler*innen in den Fächern Mathematik und Deutsch, die jeweils nach dem klassischen deutschen Notensystem anhand einer sechsstufigen Skala (1 = *Sehr gut*, 6 = *Ungenügend*) angegeben wurden. Weiterhin wurden Arbeits- und Sozialverhalten anhand der ersten vier Abstufungen dieser Notenskala angegeben (1 = *Sehr gut*, 4 = *Ausreichend*).

Lesekompetenz und kognitive Fähigkeiten

Der IGLU-2021-Lesekompetenztest umfasste informierende und erzählende Texte, an die sich Aufgaben im geschlossenen (Multiple Choice) und offenen Antwortformat anschlossen und auf unterschiedlichen Schwierigkeitsniveaus

1 Die berichteten demographischen Charakteristika der Stichprobe wurden auf Basis der imputierten Datensätze berechnet (s. Abschnitt „Datenanalyse“). Entsprechend können sie leicht von den im IGLU-2021-Berichtsband berichteten Werten abweichen.

verschiedene Aspekte des Leseverständnisses erfragten. Jedes Kind bearbeitete zwei dieser Texte mit den dazugehörigen Aufgaben (für eine ausführliche Beschreibung der eingesetzten Testinstrumente siehe Lorenz et al., 2023). Im Anschluss an die Auswertung der Aufgaben, die nach international einheitlichen Auswertungskriterien erfolgte, wurde der Lesekompetenztest international skaliert und es wurden individuelle Lesekompetenzwerte für die einzelnen Schüler*innen abgeleitet (von Davier et al., 2023). Dabei wurden für jedes Kind fünf Plausible Values berechnet, um Unsicherheiten in der Lesekompetenzschätzung zu berücksichtigen.

Für die Messung der kognitiven Fähigkeiten wurde die Subskala „figurale Analogien“ des kognitiven Fähigkeitstests (KFT, N2; Heller & Perleth, 2000) eingesetzt. Der Test besteht aus 25 Items, für deren Bearbeitung die Schüler*innen nach einer vierminütigen Einführung acht Minuten Zeit hatten (Lorenz et al., 2023). Die Items wurden mit einem 2-PL-IRT-Modell skaliert, aus dem personenspezifische WLE-Scores abgeleitet wurden. Die WLE-Reliabilität des skalierten Tests lag bei .821.

Familiärer Hintergrund

Der familiäre Hintergrund wurde durch den Zuwanderungshintergrund und den sozioökonomischen Status abgebildet. Dabei wurde der Zuwanderungshintergrund darüber operationalisiert, ob kein, ein oder beide Elternteile beziehungsweise Erziehungsberechtigten des Kindes im Ausland geboren waren. Es wurden zwei Dummy-Variablen für die Konstellationen gebildet, dass entweder ein oder beide Elternteile beziehungsweise Erziehungsberechtigten im Ausland geboren waren. Die Konstellation, dass kein Elternteil beziehungsweise kein*e Erziehungsberechtigte*r im Ausland geboren war, stellte die Referenzkategorie dar.

Der sozioökonomische Status wurde über den elterlichen beruflichen Status (*International Socio-Economic Index of Occupational Status*; ISEI) gemessen. Der ISEI (Ganzeboom, 2010; Ganzeboom et al., 1992) ist ein international vergleichbarer Index des Berufs, der Informationen über Bildung und Einkommen von Personen kombiniert. Die Skala des ISEI reicht theoretisch von 10 bis zu 90 Punkten, wobei höhere Punktzahlen einen höheren beruflichen Status anzeigen. Je Schüler*in wurde anhand der Angaben zu beiden Elternteilen beziehungsweise Erziehungsberechtigten der *höchste* vorliegende ISEI (HISEI) abgeleitet und für die Analysen genutzt.

Bildungsaspiration der Eltern

Die Bildungsaspiration der Eltern wurde anhand eines Items („Was glauben Sie: Welchen höchsten Schulabschluss wird Ihr Kind erreichen?“) erfasst. Die Eltern unterschieden dabei zwischen fünf hierarchisch angeordneten Optionen (1 = *Wird die Schule voraussichtlich ohne Abschluss verlassen*; 2 = *Hauptschulabschluss, Mittelschulabschluss, Berufsreife, Berufsbildungsreife*; 3 = *Realschulabschluss, Mittlerer Schulabschluss, Mittlere Reife, Qualifizierter Sekundarschulabschluss I*; 4 = *Fachhochschulreife*; 5 = *Abitur [Hochschulreife]*). Die Variable wurde dichotomisiert, um die Unterscheidung zwischen Abitur (1) und sonstiger Abschluss (0) herzustellen.

5.3. Datenanalyse

Die Analyse der Daten erfolgte mit Hilfe der Statistik-Software *R Studio* (Version 4.1.1; R Core Team, 2021). Insbesondere der geringen Rücklaufquoten hinsichtlich des Elternfragebogens und daraus folgenden hohen Anteilen an fehlenden Werten geschuldet (siehe Tabelle 1), wurden die fehlenden Werte auf den interessierenden Variablen in Vorbereitung auf die Analyse imputiert. Hierzu wurde das *R*-Paket *mice* (Version 4.1.3; van Buuren et al., 2015) verwendet. Basierend auf den für die Analyse relevanten Variablen, weiteren Hintergrundvariablen aus Schüler- und Elternfragebogen sowie Variablen aus der Perspektive der Lehrkräfte und Schulleitungen, die beispielsweise deren Einschätzungen zum elterlichen Engagement und der Zusammensetzung der Schülerschaft hinsichtlich Zuwanderungshintergrunds und sozioökonomischem Status umfassten, wurden 50 Datensätze erzeugt. Hinsichtlich der Anzahl der imputierten Datensätze orientierten wir uns an dem prozentualen Anteil fehlender Werte (White et al., 2011). Da jeder dieser Datensätze wiederum die fünf Plausible Values zur Lesekompetenz enthielt, wurden aus jedem Datensatz entsprechend fünf neue Datensätze generiert, die jeweils nur einen Plausible Value enthielten. Folglich wurden für die nachgeschaltete Analyse 250 Datensätze herangezogen. Um die Ergebnisse repräsentativ für die gesamte Population von Viertklässler*innen in Deutschland interpretieren zu können, wurden die Daten auf Ebene der Schüler*innen gewichtet. Darüber hinaus wurden unter Verwendung der Klassenzugehörigkeit als Cluster-Variable Random-Intercept-Modelle berechnet.

Die deskriptiven Statistiken wurden mit Hilfe des *R*-Pakets *BIFIEsurvey* (BIFIE et al., 2022) berechnet. Für die Beantwortung der Fragestellungen wurden vier aufeinander aufbauende Regressionsmodelle spezifiziert. Dabei diente die Funktion `glmer()` des *R*-Pakets *lme4* (Bates et al., 2015) zur Berech-

nung. Die abhängige Variable stellte jeweils die Gymnasialempfehlung durch die Lehrkräfte dar. Da es sich hierbei um eine dichotome Variable handelte, wurde eine logistische Regression gerechnet. Signifikante Koeffizienten im Modell bedeuten, dass eine Veränderung der unabhängigen Variablen mit einer Veränderung der Wahrscheinlichkeit einherging, eine Gymnasialempfehlung zu erhalten. Bei den in diesen Modellen ausgegebenen Koeffizienten handelt es sich um Logits (logarithmierte Odds Ratios), aus denen sich im Nachhinein durch die Funktion $e^{\text{Logit}} / (1 + e^{\text{Logit}})$ die Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, ableiten ließ. Dabei ist zu beachten, dass die Koeffizienten nicht einzeln in Wahrscheinlichkeiten umgerechnet werden können, sondern lediglich als Linearkombinationen unter Berücksichtigung des Intercepts, da die Veränderung von der Basis abhängt. In das erste Modell flossen lediglich die Schulnoten in Deutsch und Mathematik des letzten Halbjahres-Zeugnisses ein. In Modell 2 wurden zusätzlich das durch die Lehrkräfte benotete Arbeits- und das Sozialverhalten der Schüler*innen aufgenommen. Modell 3 berücksichtigt darüber hinaus die Lesekompetenz und die kognitiven Fähigkeiten der Schüler*innen als objektive, nicht direkt beobachtbare Leistungsindikatoren. Modell 4 erweiterte das vorherige Modell um Merkmale, die den familiären Hintergrund der Schüler*innen abbilden (Zuwanderungshintergrund und sozioökonomischer Status). In Modell 5 wurde schließlich zusätzlich die Bildungsaspiration der Eltern beziehungsweise Erziehungsberechtigten als möglicher Ursprung sekundärer Herkunftseffekte mit aufgenommen. Innerhalb der Modelle wurden die Schulnoten in Deutsch und Mathematik, sowie die Noten für das Arbeits- und das Sozialverhalten zentriert. Die Lesekompetenz, die kognitiven Fähigkeiten und der durch den HISEI gemessene sozioökonomische Status wurden standardisiert. Die Koeffizienten, die für jeden imputierten Datensatz zunächst einzeln berechnet wurden, wurden im Anschluss gemäß Rubin (1987) integriert.

Um die Modelle miteinander zu vergleichen, wurden neben dem Pseudo- R^2 nach McFadden (1972) die AIC- und BIC-Werte der Modelle herangezogen, bei denen kleinere Werte einen besseren Modellfit indizieren. Die fünf Modelle wurden zum einen für jeden imputierten Datensatz einzeln und zum anderen anhand eines gestapelten Datensatzes mit allen imputierten Datensätzen anhand dieser Werte verglichen. Dazu wurde für den Vergleich der AIC-Werte zusätzlich die Funktion `akaike.weights()` des R-Pakets *qpcR* (Spiess, 2018) genutzt.

6. Ergebnisse

6.1. Deskriptive Ergebnisse

Die Zusammenhänge zwischen den fokussierten Variablen entsprachen dem erwartbaren Muster. Wie Tabelle 1 zu entnehmen ist, hingen die Lesekompetenz sowie die kognitiven Fähigkeiten positiv mit der Gymnasialempfehlung zusammen, während die Schulnoten in den Fächern Deutsch und Mathematik sowie die Bewertung des Arbeits- und Sozialverhaltens (höhere Werte spiegeln negativere Bewertungen wider) negativ mit der Gymnasialempfehlung zusammenhingen. Weiterhin zeigten sich positive Assoziationen zwischen den Bildungsaspirationen sowie dem sozioökonomischen Status der Eltern und der Gymnasialempfehlung durch die Lehrkräfte. Hinsichtlich des zuwanderungsbezogenen Hintergrunds wies die Korrelation darauf hin, dass Kinder, die einen Zuwanderungshintergrund aufweisen, seltener eine Gymnasialempfehlung erhalten als Kinder, auf die dieses Merkmal nicht zutrifft.

Tabelle 1: Deskriptive Statistiken der für die Analyse relevanten Variablen und ihre Korrelationen.

	%/M (SD)	Mis- sing	Gym-E	DN	MN	AN	SN	LK	KF	ZH	HISEI
Gymnasial- empfehlung	43.32 %	4.60									
Deutschnote	2.57 (0.92)	6.87	-0.68*								
Mathenote	2.54 (0.98)	6.61	-0.64*	0.68*							
Note Arbeits- verhalten	2.19 (0.83)	1.30	-0.56*	0.64*	0.59*						
Note Sozial- verhalten	1.98 (0.69)	1.34	-0.31*	0.36*	0.29*	0.54*					
Lesekompe- tenz	525.36 (75.86)	0.00	0.53*	-0.58*	-0.51*	-0.46*	-0.29*				
Kognitive Fähigkeiten	-0.01 (1.01)	1.78	0.41*	-0.40*	-0.43*	-0.35*	-0.22*	0.45*			
Zuwande- rungshinter- grund	39.19 %	25.66	-0.18*	0.27*	0.18*	0.16*	0.09*	-0.25*	-0.17*		
HISEI	52.81 (20.60)	42.92	0.40*	-0.41*	-0.37*	-0.29*	-0.19*	0.39*	0.30*	-0.29*	
Bildungs- aspiration	50.98 %	42.53	0.59*	-0.51*	-0.50*	-0.43*	-0.25*	0.42*	0.32*	-0.03	0.40*

Anmerkungen. N = 4188. Missing = Anteil fehlender Werte vor Durchführung der Imputation; Gym-E = Gymnasialempfehlung; DN = Deutschnote; MN = Mathenote; AN = Note Arbeitsverhalten; SN = Note Sozialverhalten; LK = Lesekompetenz; KF = Kognitive Fähigkeiten; ZH = dichotomer Zuwanderungshintergrund (0 = Kein Zuwanderungshintergrund, 1 = Mindestens ein Elternteil im Ausland geboren).
*p < .05.

6.2. Vorhersage der Gymnasialempfehlung anhand von Leistungsindikatoren und familiären Hintergrundmerkmalen

Die Ergebnisse hinsichtlich der Fragestellungen können Tabelle 2 entnommen werden. In Bezug auf die erste Fragestellung zeigten die Ergebnisse von Modell 1, dass die Halbjahres-Zeugnisnoten in Deutsch und Mathematik mit der Gymnasialempfehlung der Lehrkräfte zusammenhängen ($\beta_{\text{Deutsch}} = -4.24$, $p < .001$; $\beta_{\text{Mathe}} = -3.02$, $p < .001$). Die anhand von Pseudo- R^2 quantifizierte Varianzaufklärung lag in diesem Modell bei 69.9%, was darauf hindeutete, dass die Noten in diesen beiden Schulfächern in hohem Maße mit der Gymnasialempfehlung der Lehrkräfte assoziiert sind. Bei Hinzunahme des Arbeits- und Sozialverhaltens als Prädiktoren (Modell 2) blieb dieser Zusammenhang bestehen. Darüber hinaus war die Note des Arbeitsverhaltens signifikant negativ mit der Gymnasialempfehlung der Lehrkräfte assoziiert ($\beta = -1.50$, $p < .001$), während die Note des Sozialverhaltens einen signifikant positiven Zusammenhang mit der Gymnasialempfehlung aufwies ($\beta = 0.27$, $p < .001$). Die Varianzaufklärung stieg in diesem Modell von 69.9% auf 72.1%. Zudem fielen die AIC- und BIC-Werte in Modell 2 niedriger aus als in Modell 1, sodass von einer Verbesserung des Modells durch Hinzunahme der zusätzlichen Prädiktoren auszugehen ist. Zusammenfassend ergaben die Analysen entsprechend, dass die Leistungsbewertungen der Schüler*innen durch die Lehrkraft (Schulnoten, Verhaltensweisen) einen hohen Anteil an Varianz der Gymnasialempfehlung erklären können. Dabei ging ein besser bewertetes Arbeitsverhalten über die Halbjahres-Zeugnisnoten in Deutsch und Mathematik hinaus mit einer höheren Wahrscheinlichkeit einher, eine Gymnasialempfehlung zu erhalten, während ein Anstieg des bewerteten Sozialverhaltens die Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, bei Berücksichtigung der anderen Noten verringerte.

Modell 3 (siehe Tabelle 2) beinhaltete neben den durch die Lehrkraft vorgenommenen Leistungsbewertungen der Schüler*innen auch die objektiven Leistungsmaße als Prädiktoren für die Gymnasialempfehlung. Die Ergebnisse in Bezug auf die zweite Fragestellung, die diesem Modell entnommen werden können, wiesen sowohl für die Lesekompetenz ($\beta = 0.64$, $p < .001$) als auch für die kognitiven Fähigkeiten ($\beta = 0.24$, $p < .001$) auf Zusammenhänge mit der Gymnasialempfehlung hin. Die Varianzaufklärung stieg in diesem Modell weiter auf 73.0%. Der Vergleich der AIC- und BIC-Werte ergab, dass die Aufnahme der objektiven Leistungsmaße zu einer Verbesserung des Modells führte, sodass diese auch bei Berücksichtigung der benoteten Leistungen und Verhaltensweisen der Schüler*innen einen zusätzlichen Erklärungswert für die Gymnasialempfehlung der Lehrkräfte zu haben scheinen.

In den weiteren Modellen wurden schrittweise die familiären zuwanderungsbezogenen und sozioökonomischen Hintergrundmerkmale (Modell 4) und die Bildungsaspiration (Modell 5) als Prädiktoren mit aufgenommen. Unter Kontrolle der Leistungsmerkmale und Verhaltensweisen der Schüler*innen, die schon vorher Teil des Modells waren, sowie dem HISEI als Indikator des sozioökonomischen Status wiesen zwar Kinder mit einem im Ausland geborenen Elternteil keine abweichende Wahrscheinlichkeit für eine Gymnasialempfehlung auf ($\beta_{ZH1} = 0.29, p = .050$), jedoch zeigte sich für Kinder mit zwei im Ausland geborenen Elternteilen eine signifikant höhere Wahrscheinlichkeit eine Gymnasialempfehlung zu erhalten ($\beta_{ZH2} = 0.31, p < .050$). Weiterhin war der sozioökonomische Hintergrund der Kinder signifikant positiv mit der Gymnasialempfehlung assoziiert ($\beta = 0.45, p < .001$). Die Varianzaufklärung in Modell 4 lag mit 73.5% ein wenig höher als in Modell 3 und auch die AIC- und BIC-Werte deuteten auf eine Zunahme der Vorhersagekraft des Modells bei Berücksichtigung der familiären Hintergrundmerkmale als zusätzliche Prädiktoren hin. In Bezug auf die dritte Fragestellung ergaben die Analysen zusammenfassend, dass Kinder aus sozioökonomisch privilegierten Familien unter Berücksichtigung individueller Leistungsmerkmale und Verhaltensweisen häufiger eine Gymnasialempfehlung erhielten als Kinder aus sozioökonomisch benachteiligten Verhältnissen. Weiterhin hatten Kinder mit zwei im Ausland geborenen Elternteilen unter Konstanthaltung der anderen Merkmale eine bessere Chance eine Gymnasialempfehlung zu erhalten als Kinder ohne Zuwanderungshintergrund.

Um die vierte Fragestellung zu beantworten, wurde in Modell 5 die Bildungsaspiration der Eltern mit aufgenommen. In Folge dessen wies der sozioökonomische Status weiterhin einen signifikanten Zusammenhang mit der Wahrscheinlichkeit auf, eine Gymnasialempfehlung zu erhalten ($\beta = 0.27, p < .010$). Jedoch wiesen in diesem Modell Kinder mit zwei im Ausland geborenen Elternteilen nicht mehr eine abweichende Wahrscheinlichkeit dafür auf, eine Gymnasialempfehlung zu erhalten ($\beta = -0.08, p = .559$). Die Bildungsaspiration selbst hing positiv mit der Wahrscheinlichkeit zusammen, eine Gymnasialempfehlung zu erhalten ($\beta = 1.77, p < .001$). Inhaltlich bedeuteten die Ergebnisse dieses Modells, das alle in dieser Studie betrachteten Prädiktoren berücksichtigte, dass bei einem Kind mit durchschnittlichen schulischen und objektiv gemessenen Leistungen, dessen Eltern einen durchschnittlichen sozioökonomischen Status aufwiesen und angaben, dass ihr Kind einen anderen Abschluss als das Abitur erreichen wird, die Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, bei 4.7%² lag. Lag der HISEI der Eltern

2 Die Wahrscheinlichkeit berechnete sich durch die Umrechnung des Intercepts anhand von $e^{-3.00}/(1 + e^{-3.00})$.

um eine Standardabweichung über dem Mittelwert, während die anderen Werte unverändert blieben, stieg die Wahrscheinlichkeit auf 6.1%³. Wenn die Eltern zusätzlich angaben, dass ihr Kind das Abitur erreichen wird, stieg die Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, weiter auf 27.7%⁴. Hatte ein Kind bei diesem familiären Hintergrund und dieser Bildungsaspiration der Eltern jeweils eine um einen Punkt bessere Note auf der entsprechenden Notenskala und wies um eine Standardabweichung höhere Leistungen in den objektiven Tests auf, stieg die Wahrscheinlichkeit für eine Gymnasialempfehlung auf 99.9%⁵.

Die Varianzaufklärung stieg in Modell 5 auf 75.5% an. Die Verringerung der AIC- und BIC-Werte von Modell 4 zu Modell 5 implizierte, dass die Bildungsaspiration zu einer bedeutsamen Verbesserung der Vorhersage der Gymnasialempfehlung beitrug. Zusammenfassend deuteten diese Ergebnisse darauf hin, dass zwar die Bildungsaspiration der Eltern selbst einen signifikanten Beitrag zur Aufklärung der Varianz in der Gymnasialempfehlung leistete, jedoch durch die Hinzunahme dieses Prädiktors Unterschiede, die durch den sozioökonomischen Hintergrund bedingt sind, nicht gänzlich erklärt werden konnten. Der sozioökonomische Status wirkte sich entsprechend auch dann noch positiv auf die Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, aus, wenn die durch die Lehrkraft vorgenommenen Leistungsbeurteilungen, die objektiven Leistungsmaße und die Bildungsaspiration der Eltern berücksichtigt wurden.

3 Die Wahrscheinlichkeit berechnete sich durch die Umrechnung des Intercepts und des Regressionsgewichts für den HISEI anhand von $e^{-3.00+0.27}/(1+e^{-3.00+0.27})$.

4 Die Wahrscheinlichkeit berechnete sich durch die Umrechnung des Intercepts und der Regressionsgewichte für den HISEI und die Bildungsaspiration anhand von $e^{-3.00+0.27+1.77}/(1+e^{-3.00+0.27+1.77})$.

5 Die Wahrscheinlichkeit berechnete sich durch die Umrechnung des Intercepts und der Regressionsgewichte für den HISEI, die Bildungsaspiration und die Leistungsmaße anhand von $e^{-3.00+3.53+2.68+1.46-0.46+0.53+0.21+0.27+1.77}/(1+e^{-3.00+3.53+2.68+1.46-0.46+0.53+0.21+0.27+1.77})$.

Tabelle 2: Vorhersage der Gymnasialempfehlung anhand von Leistungsindikatoren und familiären Hintergrundmerkmalen.

	Modell 1		Modell 2		Modell 3		Modell 4		Modell 5	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)	β	(SE)
Intercept	-1.89*	(0.21)	-1.98*	(0.22)	-2.00*	(0.22)	-2.15*	(0.23)	-3.00*	(0.25)
Deutschnote	-4.24*	(0.12)	-3.90*	(0.13)	-3.68*	(0.13)	-3.65*	(0.14)	-3.53*	(0.15)
Mathenote	-3.02*	(0.09)	-2.88*	(0.09)	-2.77*	(0.09)	-2.78*	(0.10)	-2.68*	(0.11)
Note Arbeitsverhalten			-1.50*	(0.08)	-1.45*	(0.08)	-1.51*	(0.08)	-1.46*	(0.10)
Note Sozialverhalten			0.27*	(0.06)	0.36*	(0.06)	0.40*	(0.06)	0.46*	(0.08)
Lesekompetenz ^a					0.64*	(0.05)	0.63*	(0.05)	0.53*	(0.06)
Kognitive Fähigkeiten ^a					0.24*	(0.04)	0.23*	(0.04)	0.21*	(0.05)
ZH 1: Ein Elternteil							0.29	(0.15)	0.10	(0.17)
ZH 2: Beide Eltern							0.31*	(0.12)	-0.08	(0.13)
HISEI ^a							0.45*	(0.08)	0.27*	(0.09)
Bildungsaspiration									1.77*	(0.14)
Zufallseffekte										
Streuung (s) auf Klassenebene	3.12		3.28		3.24		3.26		3.25	
Gütemaße										
AIC ^b	231785		215265		208180		203945		189171	
BIC ^b	231810		215303		208231		204015		189247	
Pseudo-R ²	0.699		0.721		0.730		0.735		0.755	

Anmerkungen. β = Logarithmierte Chancen/Odds Ratios (= Logits) eine Gymnasialempfehlung zu erhalten, ZH 1 = Zuwanderungshintergrund eines Elternteils (0 = *Kein Elternteil im Ausland geboren*, 1 = *Ein Elternteil im Ausland geboren*); ZH 2 = Zuwanderungshintergrund beider Elternteile (0 = *Kein Elternteil im Ausland geboren*, 1 = *Beide Elternteile im Ausland geboren*). Pseudo-R² wurde nach der Methode von McFadden (1972) berechnet.

^a z-standardisierte Variable.

^b Der Anschaulichkeit halber wurden die AIC- und BIC-Werte in dieser Tabelle über alle imputierten Datensätze gemittelt dargestellt. Der Modellvergleich erfolgte jedoch separat für jeden imputierten Datensatz und über einen gestapelten Datensatz.

* $p < .05$.

7. Diskussion

Der vorliegende Beitrag untersuchte anhand aktueller und repräsentativer Daten, welche Kriterien für die Gymnasialempfehlung, die eine Grundschullehrkraft ausspricht, eine Rolle spielen. Dabei war insbesondere von Interesse, ob familiäre Hintergrundmerkmale auch unter Berücksichtigung von diversen Leistungsmerkmalen und Verhaltensweisen mit der durch die Lehrkraft ausgesprochenen Gymnasialempfehlung zusammenhängen und ob die Bildungsaspiration der Eltern etwaige sozioökonomische oder zuwanderungsbezogene Disparitäten erklären kann.

Die Ergebnisse zeigten, dass die Halbjahres-Zeugnisnoten in Deutsch und Mathematik in hohem Maße mit der Gymnasialempfehlung zusammenhängen und entsprechend eine wichtige Rolle bei der Entscheidungsfindung der Lehrkräfte zu spielen scheinen, welche zukünftige Schulform ein Kind

ihrer Klasse besuchen soll. Die große Bedeutung der Zeugnisnoten für die Gymnasialempfehlung, konnte bereits in vielen vergangenen Untersuchungen bezeugt werden (Brändle & Weirich, 2023; Maaz & Nagy, 2009; Stubbe, Schaufelberger, et al., 2023; für einen Überblick siehe Glock et al., 2013). Die Analysen der vorliegenden Studie wiesen zudem darauf hin, dass die Noten für das Arbeits- und das Sozialverhalten über die fächerspezifischen Noten hinaus mit der Wahrscheinlichkeit zusammenhingen, eine Gymnasialempfehlung zu erhalten. Dabei fiel der Zusammenhang zwischen dem Arbeitsverhalten und der Gymnasialempfehlung erwartungsgemäß positiv aus (Anders et al., 2010; Klapproth et al., 2013; Lintorf et al., 2021; Stubbe et al., 2020) und blieb auch dann noch stabil, wenn die kognitive Fähigkeit, die Lesekompetenz und der Zuwanderungshintergrund der Schüler*innen sowie der sozioökonomische Status und die Bildungsaspiration der Eltern berücksichtigt wurden. Auch wenn das Arbeitsverhalten von Schüler*innen sich bereits in deren fächerspezifischen Noten widerspiegeln sollte, deuten diese Ergebnisse darauf hin, dass Lehrkräfte bestimmte Aspekte des Arbeitsverhaltens unabhängig von den Noten bewerten. Klapproth et al. (2013) fanden in ihrer Untersuchung, in der sie mehrere Facetten des Arbeitsverhaltens einbezogen, dass der Faktor „Zuverlässigkeit und Sorgfalt“ über die Schulnoten und objektiv gemessene Leistungen hinaus mit der Wahrscheinlichkeit, eine Empfehlung für eine höhere Schulform zu erhalten, zusammenhing. Ähnliches fanden Anders et al. (2010) für einen Faktor, der Facetten wie die Anstrengungsbereitschaft, den Fleiß oder die Pünktlichkeit abdeckte. Das Ergebnis der vorliegenden Untersuchung ist möglicherweise ebenfalls auf diese Facetten des Arbeitsverhaltens zurückzuführen und könnte zudem darauf hinweisen, dass Lehrkräfte sich an den Richtlinien für die Übergangsempfehlung orientieren, die in den Ländern der Bundesrepublik Deutschland festgelegt sind und in einigen Fällen das Arbeitsverhalten als eigenständiges Kriterium festgeschrieben haben (KMK, 2015). In Bezug auf das Sozialverhalten fanden wir zwar eine positive bivariate Korrelation mit der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten (siehe Tabelle 1), was mit den Ergebnissen bisheriger Studien einhergeht (Anders et al., 2010; Klapproth et al., 2013), jedoch wiesen die beiden Variablen unter Berücksichtigung der weiteren betrachteten Prädiktoren in allen Modellen einen signifikant negativen Zusammenhang miteinander auf. Eine mögliche Erklärung für dieses unerwartete Ergebnis könnte sein, dass die Aspekte des Sozialverhaltens, die die Lehrkräfte für die Übergangsempfehlung als relevant erachten, sich bereits in den fächerspezifischen Noten niederschlagen. Darüber hinaus zeigten unsere Ergebnisse eine hohe Korrelation mit dem Arbeitsverhalten auf (siehe Tabelle 1), sodass die Noten, die die Lehrkräfte für diese beiden Verhaltensweisen vergeben, eng

miteinander verknüpft zu sein scheinen. Anhand unserer Daten kann jedoch nicht abschließend geklärt werden, warum sich der Zusammenhang mit der Gymnasialempfehlung unter der Berücksichtigung der leistungsbezogenen Variablen umkehrt.

Weiterhin zeigten die Ergebnisse, dass auch für die Leistungsentwicklung relevante Fähigkeiten und Kompetenzen aus objektiven Tests, deren Ergebnisse den Lehrkräften nicht vorliegen, mit der Wahrscheinlichkeit zusammenhängen, eine Gymnasialempfehlung zu erhalten. Sowohl die figuralen kognitiven Fähigkeiten als auch die Lesekompetenz wiesen signifikante Zusammenhänge mit der Gymnasialempfehlung der Lehrkräfte auf, auch wenn die Halbjahres-Zeugnisnoten in Deutsch und Mathematik, das Arbeits- und das Sozialverhalten der Schüler*innen sowie ihr familiärer Hintergrund und die Bildungsaspiration ihrer Eltern berücksichtigt wurden. Vergangene Studien berichteten ähnliche Ergebnisse (Anders et al., 2010; Brändle & Weirich, 2023; Klapproth et al., 2013; Neugebauer, 2011; Stubbe et al., 2020; Stubbe, Schaufelberger, et al., 2023). Da den Lehrkräften jedoch die Testergebnisse der objektiven Leistungsindikatoren, die wir in der vorliegenden Studie nutzten, nicht zugänglich waren, bleibt die Frage offen, warum diese trotzdem mit der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, zusammenhängen. Eine mögliche Erklärung wäre, dass die Lehrkräfte Potenziale der Schüler*innen (kognitive Fähigkeiten und Lesekompetenz), die eine positive Leistungsentwicklung vermuten lassen, anhand des unterrichtlichen Geschehens wahrnehmen können und auf dieser Grundlage ihre Empfehlung für die weitere Schullaufbahn basieren. Auch wenn die kognitiven Fähigkeiten und die Lesekompetenz in bedeutendem Maße mit den Schulnoten zusammenhängen (siehe Tabelle 1), wäre es möglich, dass sich bestimmte Aspekte einer schulischen Bewertung entziehen, jedoch von den Lehrkräften für die Prognose der Befähigung für das Gymnasium herangezogen werden.

Hinsichtlich des Zuwanderungshintergrundes der Schüler*innen konnte bei simultaner Betrachtung der Leistungsmerkmale und des sozioökonomischen Hintergrundes gezeigt werden, dass Kinder mit einem im Ausland geborenen Elternteil keine andere Wahrscheinlichkeit aufwiesen, eine Gymnasialempfehlung zu erhalten, als Kinder ohne Zuwanderungshintergrund, während solche mit zwei im Ausland geborenen Elternteilen diesbezüglich eine höhere Wahrscheinlichkeit aufwiesen. Der Befund für Kinder mit einem im Ausland geborenen Elternteil ist insofern plausibel, als dass Kinder mit Zuwanderungshintergrund zwar seltener eine Gymnasialempfehlung erhalten und auf weniger leistungsorientierten weiterführenden Schulformen überrepräsentiert sind (Wößmann et al., 2023), dies aber zum Teil durch primäre Herkunftseffekte erklärt werden kann. So stammen Kinder mit Zuwande-

runghintergrund einerseits häufig aus Familien mit einem niedrigeren sozioökonomischen Status und weisen andererseits meist geringere schulische Leistungen auf als Kinder ohne Zuwanderungshintergrund, was mitunter darin begründet ist, dass sie zu Hause seltener Deutsch sprechen (Relikowski et al., 2010; Stubbe, Kleinkorres, et al., 2023; für einen Überblick siehe Dumont et al., 2014). Dies deckt sich auch mit den deskriptiven Ergebnissen dieser Studie, die zeigten, dass Kinder mit Zuwanderungshintergrund durchschnittlich niedrigere HISEI-Werte, geringere Lesekompetenzen und unvorteilhaftere Schulnoten aufwiesen (siehe Tabelle 1). Das Ergebnis, das sich für Kinder mit zwei im Ausland geborenen Elternteilen zeigte, knüpft hingegen an Befunde an, die für Kinder mit einem Zuwanderungshintergrund begünstigende sekundäre Herkunftseffekte berichteten. Diese Herkunftseffekte schlugen sich in bisherigen Studien zwar insbesondere im tatsächlichen Übergang auf die weiterführende Schule nieder, zeigten sich aber auch bereits in einer Bevorteilung bei der Übergangsempfehlung (Arnold et al., 2007; Gresch, 2012; Nishen et al., 2022) und konnten mit der Bildungsaspiration der Eltern in Verbindung gebracht werden (Becker & Gresch, 2016; Dumont et al., 2014; Astleithner et al., 2021). Dazu passend glichen sich in der vorliegenden Studie – unter Kontrolle leistungsbezogener Merkmale – für Kinder mit zwei im Ausland geborenen Elternteilen und Kinder ohne Zuwanderungshintergrund die Chancen auf eine Gymnasialempfehlung an, wenn die Bildungsaspiration der Eltern in das Modell mit aufgenommen wurde. Es muss jedoch berücksichtigt werden, dass kein bivariater Zusammenhang zwischen dem Zuwanderungshintergrund und der Bildungsaspiration gefunden wurde (siehe Tabelle 1). Wie beispielsweise Becker & Gresch (2016) berichteten, liegt eine starke Heterogenität der Bildungsaspiration der Eltern von Kindern mit Zuwanderungshintergrund je nach Herkunftsland vor. Insofern muss der Zusammenhang zwischen dem Zuwanderungshintergrund, der Bildungsaspiration der Eltern und der Gymnasialempfehlung differenzierter betrachtet werden.

Trotz der Berücksichtigung einer Vielzahl von leistungsbezogenen Variablen wiesen die Analysen auf einen inkrementellen Zusammenhang zwischen dem durch den HISEI gemessenen sozioökonomischen Status und der Gymnasialempfehlung hin. Dieser Befund konnte bereits in anderen Studien beobachtet werden (Maaz & Nagy, 2009; Stubbe, Schaufelberger, et al., 2023). Einige Studien fanden hingegen, dass der sozioökonomische Status bei Kontrolle anderer Variablen, insbesondere der Bildungsaspiration, nicht mehr mit der Gymnasialempfehlung zusammenhing (Brändle & Weirich, 2023; Dumont et al., 2019; Klapproth et al., 2013). Jedoch blieb der Zusammenhang im vorliegenden Beitrag auch dann noch bestehen, wenn die Bildungsaspi-

ration der Eltern, die ihrerseits einen signifikant positiven Zusammenhang mit der Gymnasialempfehlung aufwies, berücksichtigt wurde. Die Bedeutung der Bildungsaspiration der Eltern für die Gymnasialempfehlung der Lehrkräfte könnte beispielsweise damit erklärt werden, dass Lehrkräfte und Eltern in Austausch miteinander stehen und die Lehrkräfte entsprechend Einsichten über die elterlichen Präferenzen in Bezug auf eine bestimmte Übergangsempfehlung, sowie deren Einstellungen und Motivationen im Hinblick auf Bildung (vgl. z. B. Wert-Erwartungstheorien; Breen & Goldthorpe, 1997) gewinnen (Dumont et al., 2019). Im Rahmen eines solchen Austauschs wäre es möglich, dass Eltern unbewusst oder bewusst aktiven Einfluss auf die Schullaufbahneempfehlungen der Lehrkräfte nehmen (Dumont et al., 2014; Dumont et al., 2019). Dies ließe allerdings weiterhin die Frage nach differenzierten Merkmalen offen, in denen der sozioökonomische Hintergrund über die individuellen Leistungsmerkmale und Verhaltensweisen sowie die elterliche Bildungsaspiration hinaus noch in das Urteil der Lehrkräfte über eine Gymnasialempfehlung einfließt. Eine mögliche Erklärung für diesen Befund wäre, dass Lehrkräfte bestimmte Erwartungen an Schüler*innen unterschiedlichen familiären Hintergrunds haben und antizipieren, dass Kinder aus sozioökonomisch privilegierten Verhältnissen bessere elterliche Unterstützung erfahren und Ressourcen zur Verfügung gestellt bekommen, die sie für die Herausforderungen auf dem Gymnasium wappnen und ihnen bessere Erfolgchancen bieten (Dumont et al., 2014). Beispielsweise berichteten Gräsel et al. (2010) auf der Grundlage von Lehrkraftbefragungen, dass diese teilweise die Unterstützung aus dem Elternhaus in ihre Empfehlung mit einbeziehen. Mit den vorliegenden Daten kann diese Frage jedoch nicht abschließend beantwortet werden.

7.1. Stärken und Limitationen

Eine Stärke der vorliegenden Studie betrifft die umfassende Operationalisierung von Merkmalen, die für die Beurteilung, ob eine Schülerin oder ein Schüler für eine bestimmte weiterführende Schule geeignet ist, relevant sein sollten. Dabei ist hervorzuheben, dass auf Seiten der Merkmale, die den Lehrkräften zugänglich sind und deren Bewertung sie selbst vorgenommen haben, neben den Schulnoten auch das Arbeits- und das Sozialverhalten als Indikatoren berücksichtigt wurden. Weiterhin wurden die erreichten Leistungen von Schüler*innen in objektiven Leistungstests einbezogen, deren Ergebnisse den Lehrkräften nicht zugänglich waren. Der signifikante Einfluss, den die Leistungen in diesen Tests über die Noten, die die Lehrkräfte vergeben hat-

ten, hinaus besaßen, könnte ein Indiz dafür sein, dass die Lehrkräfte in ihre Schullaufbahnempfehlung auch die von den aktuellen Schulleistungen unabhängigen Leistungspotenziale der Schüler*innen mit einfließen lassen, die sie im Unterricht beobachten. Eine weitere Stärke bezieht sich auf die Stichprobe, die auf einer aktuellen und repräsentativen Erhebung von Viertklässler*innen in Deutschland basiert. Auf Grund des stetigen Wandels des deutschen Bildungssystems, gesellschaftlicher Dynamiken in den letzten Jahren sowie kontinuierlicher Bemühungen, Chancengleichheit herzustellen, ist es von großer Relevanz, die Kriterien, die Lehrkräfte für eine Gymnasialempfehlung heranziehen, unter Einbezug familiärer Hintergrundmerkmale anhand aktueller Daten zu beleuchten.

Bei der Interpretation der Ergebnisse der vorliegenden Studie müssen jedoch auch mehrere Limitationen berücksichtigt werden. Eine erste Limitation betrifft die verhältnismäßig geringe Beteiligung der Eltern und Erziehungsberechtigten an der Studie. Lediglich 55 % der Eltern und Erziehungsberechtigten von den 4611 teilnehmenden Schüler*innen füllten den Fragebogen mit Hintergrundinformationen zur familiären Situation aus. Entsprechend hoch waren die Anteile an fehlenden Werten bei den Variablen zur Bestimmung des sozioökonomischen Status und der Bildungsaspiration (siehe Tabelle 1). Dieser Einschränkung wurde mit einer Imputation von 50 verschiedenen Datensätzen begegnet. Die resultierende prozentuale Verteilung des Zuwanderungshintergrundes sowie der Mittelwert des sozioökonomischen Status (gemessen durch den HISEI) lagen nah an den Schätzungen anderer repräsentativer Stichproben (Lewalter et al., 2023; Stanat et al., 2022; Stanat et al., 2023), was auf eine adäquate Abschätzung dieser Variablen nach Durchführung der Imputation hindeutet. Nichtsdestotrotz muss diese Limitation bei der Interpretation der Ergebnisse berücksichtigt werden, da diese Variablen zentral für die Analysen waren. Eine weitere Limitation betrifft das querschnittliche Design der Studie. Hinsichtlich der Vorhersage der Gymnasialempfehlung durch die Bildungsaspiration der Eltern resultierte dadurch das Problem des zeitlichen Ablaufs. Wie bereits beschrieben wurde, wäre es möglich, dass Eltern im Austausch mit den Lehrkräften ihre Bildungsaspiration, im Sinne des von ihnen favorisierten Schulabschlusses für ihr Kind, durchscheinen lassen. Da die Untersuchung jedoch nach Abschluss des ersten Halbjahres der vierten Jahrgangsstufe durchgeführt wurde, was meist der Zeitpunkt ist, zu dem die Eltern die Übergangsempfehlung der Lehrkräfte mitgeteilt bekommen (Autor:innengruppe Bildungsberichterstattung, 2024), wäre ebenfalls möglich, dass die Aussage der Eltern über den Schulabschluss, den sie für ihr Kind erwarten, bereits durch die Empfehlung der Lehrkräfte gefärbt war. Zuletzt muss als Einschränkung angeführt werden, dass die in

dieser Studie berichteten Ergebnisse die Information unberücksichtigt ließen, ob in dem Bundesland, in dem die Schüler*innen jeweils zur Schule gingen, eine starke oder schwache Bindung an die Übergangsempfehlung bestand. Zwar wäre anzunehmen, dass sich dies insbesondere im Übergangsverhalten der Eltern bemerkbar macht (Gresch et al., 2010), jedoch wäre ebenfalls denkbar, dass Eltern mit einem hohen sozioökonomischen Status und hohen Bildungsaspirationen insbesondere in Bundesländern, in denen die Übergangsempfehlung verbindlich ist, versuchen, aktiv auf die Lehrkräfte einzuwirken, um sie zu einer Übergangsempfehlung für eine höhere Schulform zu bewegen. Die Befundlage hinsichtlich der Rolle der Verbindlichkeit der Übergangsempfehlung ist allerdings heterogen (Autor:innengruppe Bildungsberichterstattung, 2024), wobei jüngste empirische Evidenz eher nicht darauf hindeutet, dass die freie Wahl der Eltern über die weiterführende Schule zu einer Verschärfung sozioökonomischer Disparitäten führt (Esser & Hoenig, 2018; Bach, 2023).

7.2. Implikationen und Ausblick

Die Ergebnisse der aktuellen Studie deuteten insgesamt darauf hin, dass die schulischen und durch objektive Tests gemessenen Leistungen der Schüler*innen in hohem Maße mit der Wahrscheinlichkeit zusammenhängen, eine Gymnasialempfehlung zu erhalten. Dieser Befund spricht dafür, dass Lehrkräfte bei ihrer Übergangsempfehlung der Idee des deutschen Schulsystems, Schüler*innen leistungsgerecht zu gruppieren, nachkommen. Außerdem hing der Zuwanderungshintergrund bei Berücksichtigung der Leistungsmerkmale und anderer familiärer Hintergrundvariablen nicht mit dem Lehrkrafturteil über die Befähigung für das Gymnasium zusammen, was ebenfalls positiv zu bewerten ist, da es darauf hinweist, dass Lehrkräfte diesen für den Erfolg auf dem Gymnasium unwesentlichen Faktor nicht in ihr Urteil mit einbeziehen. Der sozioökonomische Status hingegen wies bei Kontrolle der Leistungsmerkmale weiterhin einen signifikanten Zusammenhang mit der Gymnasialempfehlung auf und auch die elterliche Bildungsaspiration war mit der Wahrscheinlichkeit assoziiert, von der Lehrkraft eine Gymnasialempfehlung zu erhalten. Auch wenn der Anteil der erklärten Varianz durch den sozioökonomischen Status und die Bildungsaspiration nicht sehr hoch war, so ist dies ein alarmierendes Ergebnis, zumal bereits die schulischen Leistungen in hohem Maße durch primäre Herkunftseffekte beeinflusst worden sein können (Dumont et al., 2014; Maaz & Nagy, 2009).

In zukünftigen Analysen sollte der Zusammenhang zwischen der Bildungsaspiration und der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, genauer untersucht werden. Idealerweise sollten dabei sowohl die Bildungsaspiration der Eltern als auch die Einschätzung der Lehrkräfte, welche Schulform sie für das Kind als geeignet erachten, über einen längeren Zeitraum untersucht werden, um die Stabilität dieses Merkmals in der Vorhersage der Gymnasialempfehlung berücksichtigen zu können und genauer aufzuschlüsseln, inwiefern sich Lehrkraft und Eltern in ihren Einschätzungen beeinflussen. Ein weiteres Desiderat in diesem Zusammenhang ist, die Dynamik zwischen dem Zuwanderungshintergrund und der Bildungsaspiration weiter aufzuklären. Eine tiefergehende Analyse unter Berücksichtigung des Grundes für die Zuwanderung und des Herkunftslandes könnte dabei helfen, die Motivation hinter der Bildungsaspiration genauer zu verstehen. So kann beispielsweise angenommen werden, dass Eltern, die aus ihrem Heimatland flüchten mussten, andere Erwartungen hinsichtlich des Abschlusses für ihr Kind hegen als solche, die ihr Heimatland verlassen haben, weil sie eine Verbesserung ihres sozialen Status anstreben (vgl. Astleithner et al., 2021; Kao & Tienda, 1995). Weiterhin kann sich die Bildungsaspiration von Eltern je nach deren Herkunftsland unterscheiden (Becker & Gresch, 2016).

Die vorliegenden Ergebnisse zeigten Zusammenhänge der kognitiven Fähigkeiten und der Lesekompetenz mit der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, was möglicherweise darauf zurückzuführen ist, dass Lehrkräfte diese Fähigkeiten und Kompetenzen unabhängig von den zu bewertenden Leistungen wahrnehmen und daraus die Lernentwicklung ableiten können. Zukünftige Studien sollten dabei alternative Methoden einbeziehen, um Leistungspotenziale zu messen. Beispielsweise könnte die Leistungsentwicklung der Schüler*innen über einen bestimmten Zeitraum mittels Lernverlaufsdagnostik (Souvignier, 2018) erfasst werden, um ihr Leistungspotenzial abzuschätzen. Diese Art der Diagnostik birgt nicht nur für die Forschung, sondern insbesondere für die Praxis Potenzial, da sie den Lehrkräften Leistungspotenziale greifbar macht und sie nicht implizit darauf schließen müssen. Über den Einblick in die Leistungsentwicklung der Schüler*innen hinaus wird den Lehrkräften dadurch ermöglicht, Kinder gezielt zu fördern. Testergebnisse aus der Lernverlaufsdagnostik könnten den Lehrkräften zudem als zusätzliches Mittel dienen, ihre Übergangsempfehlung vor den Schüler*innen und deren Eltern zu rechtfertigen.

Auch wenn die vorliegende Studie eine konzeptuelle Unterscheidung zwischen den schulischen Leistungen und Verhaltensweisen sowie objektiven Fähigkeiten und Kompetenzen einerseits und den familiären Hintergrundmerkmalen andererseits machte und deren Zusammenhänge mit der Gymna-

sialempefehlung unabhängig voneinander interpretiert wurden, ist zu beachten, dass auch Leistungs- und Kompetenzunterschiede bereits durch soziale Ungleichheiten hervorgerufen werden und als primäre Herkunftseffekte in die Schulnoten einfließen können (Dumont et al., 2014; Maaz & Nagy, 2009). Da unsere Studie die erhebliche Bedeutung der schulischen Leistungen für die Gymnasialempfehlung und folglich die spätere schulische Laufbahn nahelegt und damit frühere Befunde bestätigt, sollten im Sinne der Chancengleichheit und Bildungsgerechtigkeit, die das deutsche Bildungssystem als grundlegende Ziele verfolgt, Unterstützungsangebote für Kinder aus benachteiligten sozioökonomischen Verhältnissen geschaffen werden. Wößmann et al. (2023) schlagen in diesem Zusammenhang beispielsweise vor, frühkindliche Bildungsangebote für benachteiligte Kinder auszubauen, Familien benachteiligter Kinder bei der Erziehung zu unterstützen und Nachhilfeprogramme für benachteiligte Kinder kostenfrei anzubieten. Das *Startchancen-Programm*, dessen Umsetzung kürzlich vom Bundesministerium für Bildung und Forschung (BMBF) und den Ländern der Bundesrepublik Deutschland vereinbart wurde (BMBF, 2024), strebt an, Bildungserfolg und soziale Herkunft zu entkoppeln und bietet durch das ambitionierte Ziel, 4000 Schulen mit einem hohen Anteil an sozioökonomisch benachteiligten Kindern zu erreichen, einen vielversprechenden Ausblick. Ausgehend von den Befunden zu den Zusammenhängen des sozioökonomischen Status und der Bildungsaspiration mit der Wahrscheinlichkeit, eine Gymnasialempfehlung zu erhalten, sollten Lehrkräfte außerdem darauf aufmerksam gemacht werden, dass bei der Übergangsempfehlung weiterhin sozioökonomische Disparitäten vorliegen, die unabhängig von den schulischen Leistungen sind.

Literatur

- Anders, Y., McElvany, N., & Baumert, J. (2010). Die Einschätzung lernrelevanter Schülermerkmale zum Zeitpunkt des Übergangs von der Grundschule auf die weiterführende Schule: Wie differenziert urteilen Lehrkräfte? [The assessment of learning-relevant student characteristics at the time of transition from primary to secondary school: How differentiated are teachers' assessments?] In K. Maaz, J. Baumert, C. Gresch, & N. McElvany (Eds.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (pp. 313–330). Bundesministerium für Bildung und Forschung.
- Arnold, K.-H., Bos, W., Richert, P., & Stubbe, T. C. (2007). Schullaufbahnpräferenzen am Ende der vierten Klassenstufe [School career preferences at the end of the fourth grade]. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert, & R. Valtin (Eds.), *IGLU 2006: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 271–297). Waxmann.

- Astleithner, F., Vogl, S., & Parzer, M. (2021). Zwischen Wunsch und Wirklichkeit: Zum Zusammenhang von sozialer Herkunft, Migration und Bildungsaspirationen [Between desire and reality: The connection between social background, migration and educational aspirations]. *Österreichische Zeitschrift für Soziologie*, 46(3), 233–256. <https://doi.org/10.1007/s11614-021-00442-3>
- Autor:innengruppe Bildungsberichterstattung (2024). *Bildung in Deutschland 2024. Ein indikatorengestützter Bericht mit einer Analyse zu beruflicher Bildung* [Education in Germany 2024. An indicator-based report with an analysis of vocational education]. wbv Publikation <https://doi.org/10.3278/6001820iw>
- Bach, M. (2023). Heterogeneous responses to school track choice: Evidence from the repeal of binding track recommendations. *Economics of Education Review*, 95, 102412. <https://doi.org/10.1016/j.econedurev.2023.102412>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., Krivitsky, P. N., Tanaka, E., & Jagan, M. (2015). *Package 'lme4'*. [Computer software]. <https://cran.r-project.org/web/packages/lme4/lme4.pdf>
- Becker, B., & Gresch, C. (2016). Bildungsaspirationen in Familien mit Migrationshintergrund [Educational aspirations in families with a migration background]. In C. Diehl, C. Hunkler, & C. Kristen (Eds.), *Ethnische Ungleichheiten im Bildungsverlauf* (pp. 73–115). Springer VS. https://doi.org/10.1007/978-3-658-04322-3_3
- Biewen, M., & Tapalaga, M. (2017). Life-cycle educational choices in a system with early tracking and 'second chance' options. *Economics of Education Review*, 56, 80–94. <https://doi.org/10.1016/j.econedurev.2016.11.008>
- BIFIE, Robitzsch, A., & Oberwimmer, K. (2022). *BIFIEsurvey: Tools for survey statistics in educational assessment*. R package version 3.4-15. <https://CRAN.R-project.org/package=BIFIEsurvey>
- BMBF (2024, 2. Februar). *Vereinbarung zwischen Bund und Ländern zur Umsetzung des Startchancen-Programms für die Jahre 2024 bis 2034* [Agreement between the federal and state governments on the implementation of the Startchancen program for the years 2024 to 2034]. https://www.bmbf.de/SharedDocs/Downloads/DE/2024/blv-startchancen.pdf?__blob=publicationFile&v=4
- Böhmer, I., Hörstermann, T., Gräsel, C., Krolak-Schwerdt, S., & Glock, S. (2015). Eine Analyse der Informationssuche bei der Erstellung der Übergangsempfehlung: Welcher Urteilsregel folgen Lehrkräfte? [An analysis of the information search in the preparation of the transition recommendation: Which rule of judgment do teachers follow?]. *Journal for Educational Research Online*, 7(2), 59–81. <https://doi.org/10.25656/01:11490>
- Boudon, R. (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. John Wiley & Sons Inc.
- Brändle, T., & Weirich, S. (2023). Leistungsunabhängige Urteile? [Performance-independent judgments?] *Zeitschrift für Grundschulforschung*, 16(1), 153–172. <https://doi.org/10.1007/s42278-022-00160-z>
- Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials. Towards a formal rational action theory. *Rationality and Society*, 9(3), 275–305. <https://doi.org/10.1177/104346397009003002>
- Cromley, J. G. (2009). Reading achievement and science proficiency: International comparisons from the programme on international student assessment. *Reading Psychology*, 30(2), 89–118. <https://doi.org/10.1080/02702710802274903>
- Ditton, H. (2010). Der Beitrag von Schule und Lehrern zur Reproduktion von Bildungsungleichheit [The contribution of schools and teachers to the reproduction of educational inequality]. In R. Becker & W. Lauterbach (Eds.), *Bildung als Privileg: Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (pp. 247–275). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92484-7_9

- Dumont, H., Klinge, D., & Maaz, K. (2019). The many (subtle) ways parents game the system: Mixed-method evidence on the transition into secondary-school tracks in Germany. *Sociology of Education*, 92(2), 199–228. <https://doi.org/10.1177/0038040719838223>
- Dumont, H., Maaz, K., Neumann, M., & Becker, M. (2014). Soziale Ungleichheiten beim Übergang von der Grundschule in die Sekundarstufe I: Theorie, Forschungsstand, Interventions- und Fördermöglichkeiten [Social inequalities in the transition from elementary school to lower secondary level: theory, state of research, possibilities for intervention and support]. *Zeitschrift für Erziehungswissenschaft*, 17(2), 141–165. <https://doi.org/10.1007/s11618-013-0466-1>
- Dumont, H., Neumann, M., Becker, M., Maaz, K., & Baumert, J. (2013). Der Übergangsprozess von der Grundschule in die Sekundarstufe I vor und nach der Schulstrukturreform in Berlin: Die Rolle primärer und sekundärer Herkunftseffekte [The transition process from elementary school to lower secondary level before and after the school structure reform in Berlin: The role of primary and secondary effects of origin]. In K. Maaz, J. Baumert, M. Neumann, M. Becker, & H. Dumont (Eds.), *Die Berliner Schulstrukturreform: Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen* (pp. 133–207). Waxmann.
- Ehmke, T., & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb [Social background and skills acquisition]. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider, & P. Stanat (Eds.), *PISA 2009: Bilanz nach einem Jahrzehnt* (pp. 231–254). Waxmann.
- El-Mafaalani, A. (2014). Habitus-Struktur-Sensibilität – (Wie) kann ungleichheitssensible Schulpraxis gelingen? [Habitus-Structure-Sensitivity – (How) can inequality-sensitive school practice succeed?] In T. Sander (Ed.), *Habituussensibilität: Eine neue Anforderung an professionelles Handeln* (pp. 229–245). Springer VS. https://doi.org/10.1007/978-3-658-06887-5_10
- Esser, H. (2006). *Sprache und Integration: die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten* [Language and integration: the social conditions and consequences of migrants' language acquisition]. Campus Verlag.
- Esser, H., & Hoenig, K. (2018). Leistungsgerechtigkeit und Bildungsungleichheit [Achievement equity and educational inequality]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 70, 419–447. <https://doi.org/10.1007/s11577-018-0558-2>
- Feld, L. P., Doerr, A., Hirsch, P., & Sajons, C. (2017). Zuwanderung nach Deutschland von 1945 bis heute [Immigration to Germany from 1945 to the present day]. *Malteser Migrationsbericht*, 2017, 8–27. https://www.malteser.de/fileadmin/Files_sites/malteser_de_Relaunch/Angebote_und_Leistungen/Migrationsbericht/Malteser_Migrationsbericht_2017_es.pdf
- Ganzeboom, H. B. G. (2010). *A new International Socio-Economic Index [ISEI] of Occupational Status for the International Standard Classification of Occupation 2008 [ISCO-08] Constructed with Data from the ISSP 2002–2007; with an Analysis of Quality of Educational Measurement in SSP*. Vortrag auf der Annual Conference of International Social Survey Programme, Lissabon.
- Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). A Standard International Socio-Economic Index of Occupational Status. *Social Science Research*, 21(1), 1–56. [https://doi.org/10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B)
- Glock, S., Krolak-Schwerdt, S., Klapproth, F. & Böhmer, M. (2013). Prädiktoren der Schullaufbahneempfehlung für die Schulzweige des Sekundarbereichs I [Predictors of the school career recommendation for the school branches of the lower secondary level]. *Pädagogische Rundschau*, 67(3), 329–347.

- Gräsel, C., Krolak-Schwerdt, S., Nölle, I., & Hörstermann, T. (2010). Diagnostische Kompetenz von Grundschullehrkräften bei der Erstellung der Übergangsempfehlung. Eine Analyse aus der Perspektive der sozialen Urteilsbildung. Projekt Diagnostische Kompetenz [Diagnostic competence of primary school teachers in the preparation of transition recommendations. An analysis from the perspective of social judgment. Project Diagnostic competence]. In E. Klieme, D. Leutner, & M. Kenk (Eds.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (pp. 286–295). Beltz.
- Gresch, C. (2012). *Der Übergang in die Sekundarstufe I: Leistungsbeurteilung, Bildungsaspiration und rechtlicher Kontext bei Kindern mit Migrationshintergrund* [The transition to lower secondary level: achievement assessment, educational aspirations and the legal context for children with a migration background]. Springer.
- Gresch, C., Baumert, J., & Maaz, K. (2010). Empfehlungsstatus, Übergangsempfehlung und der Wechsel in die Sekundarstufe I: Bildungsentscheidungen und soziale Ungleichheit [Recommendation status, transition recommendation and the change to lower secondary level: Educational choices and social inequality]. In K. Maaz, J. Baumert, C. Gresch, & N. McElvany (Eds.), *Der Übergang von der Grundschule in die weiterführende Schule: Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (pp. 230–256). Bundesministerium für Bildung und Forschung.
- Heath, A., & Brinbaum, Y. (2007). Guest Editorial: Explaining Ethnic Inequalities in Educational Attainment. *Ethnicities*, 7(3), 291–304. <https://doi.org/10.1177/1468796.807080230>
- Heller, K. A., & Perleth, C. (2000). *KFT 4–12+R. Kognitiver Fähigkeitstest für 4. bis 12. Klasse, Revision* [KFT 4-12+R. Cognitive ability test for Grades 4 to 12, revision]. Beltz.
- Henschel, S., Heppt, B., Rjosk, C., & Weirich, S. (2022) Zuwanderungsbezogene Disparitäten [Immigration-related disparities]. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich & S. Henschel, (Eds), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann.
- Hunkler, C., & Tjaden, J. (2018). Die Ausbildungsentscheidungen von Migranten im stratifizierten deutschen Bildungssystem: zu optimistisch? [The educational decisions of migrants in the stratified German education system: too optimistic?] In E. Schilling (Ed.), *Verwaltete Biografien* (pp. 71–107). Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-20522-5_4
- Kao, G., & Tienda, M. (1995). Optimism And Achievement: The Educational Performance of Immigrant Youth. *Social Science Quarterly*, 76(1), 1–19. <https://www.jstor.org/stable/44072586>
- Kao, G., & Tienda, M. (1998). Educational Aspirations of Minority Youth. *American Journal of Education*, 106(3), 349–384. <https://doi.org/10.1086/444188>
- Klapproth, F., Glock, S., Krolak-Schwerdt, S., Martin, R., & Böhmer, M. (2013). Prädiktoren der Sekundarschulempfehlung in Luxemburg. Ergebnisse einer Large-Scale-Untersuchung. *Zeitschrift für Erziehungswissenschaft* [Predictors of the Tracking Recommendation in Luxembourg: Results of a Large-Scale Study], *Zeitschrift für Erziehungswissenschaft*, 16(2), 355–379. <https://doi.org/10.1007/s11618-013-0340-1>
- KMK (2019, 5. Dezember). *Bildungssprachliche Kompetenzen in der deutschen Sprache stärken* [Strengthening educational language skills in German]. Beschluss der Kultusministerkonferenz vom 05.12.2019. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2019/2019_12_05-Beschluss-Bildungssprachl-Kompetenzen.pdf

- KMK (2015, 19. Februar). *Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6* [Transition from elementary school to lower secondary schools and support, observation and orientation in Grades 5 and 6]. Informationsschrift des Sekretariats der Kultusministerkonferenz vom 19.02.2015. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf
- Kramer, R. T. (2017). „Habitus“ und „kulturelle Passung“ [“Habitus” and “cultural fit“]. In M. Rieger-Ladich & C. Grabau (Eds.), *Pierre Bourdieu: Pädagogische Lektüren* (pp. 183–205). Springer VS. https://doi.org/10.1007/978-3-531-18904-8_10
- Kuhn, A. (2023, 19. März). *Auf welche weiterführende Schule nach der Grundschule? – Das Deutsche Schulportal* [Which secondary school after elementary school? – The German School Portal]. Das Deutsche Schulportal. <https://deutsches-schulportal.de/bildungswesen/infografik-schuluebergang-auf-welche-schule-nach-der-grundschule>
- Lewalter, D., Diedrich, J., Goldhammer, F., Köller, O., & Reiss, K. (2023). *PISA 2022. Analyse der Bildungsergebnisse in Deutschland* [PISA 2022. Analysis of educational results in Germany]. Waxmann.
- Lintorf, K., van Ophuysen, S., & Osipov, I. (2021). Comparing Assessment Methods of Attribute Importance in Teachers’ Decisions. The Importance of Different Criteria for Tracking Recommendations after Primary School. *Education Sciences*, 11(10), 566. <https://doi.org/10.3390/educsci11100566>
- Liu, S., Wei, W., Chen, Y., Hugo, P., & Zhao, J. (2021). Visual-Spatial Ability Predicts Academic Achievement Through Arithmetic and Reading Abilities. *Frontiers in Psychology*, 11, 591308. <https://doi.org/10.3389/fpsyg.2020.591308>
- Lorenz, R., Frey, A., Trendtel, M., Ludewig, U., Schilcher, A., & McElvany, N. (2023). Ziele, Design, Instrumente und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2021) [Aims, design, instruments and implementation of the Progress in International Reading Literacy Study (PIRLS 2021)]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. Stubbe (Eds.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 27–52). Waxmann.
- Maaz, K., & Nagy, G. (2009). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems: Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte [The transition from primary to secondary school in the secondary school system: definition, specification and quantification of primary and secondary effects of origin]. In J. Baumert, K. Maaz, & U. Trautwein (Eds.), *Bildungsentscheidungen* (pp. 153–182). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92216-4_7
- McElvany, N., Lorenz, R., Frey, A., Goldhammer F., Schilcher, A., & Stubbe, T. (2023). *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* [IGLU 2021. Reading literacy of primary school children in international comparison and trend over 20 years]. Waxmann.
- McFadden, D. (1972). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press.
- Müller-Benedict, V. (2007). Wodurch kann die soziale Ungleichheit des Schulerfolgs am stärksten verringert werden? [How can social inequality in school success be reduced the most?] *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 59, 615–639. <https://doi.org/10.1007/s11577-007-0080-4>
- Neugebauer, M. (2011). Werden Jungen von Lehrerinnen bei den Übergangsempfehlungen für das Gymnasium benachteiligt? Eine Analyse auf Basis der IGLU-Daten [Are boys disadvantaged by female teachers when it comes to transition recommendations for grammar school? An analysis based on PIRLS data]. In A. Hadjar (Ed.), *Geschlechtsspezifische Bildungsungleichheiten*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92779-4_10

- Nishen, A. K., Corcoran, K., Holder, K., & Kessels, U. (2022). When Ethnic Minority Students are Judged as More Suitable for the Highest School Track. A Shifting Standards Experiment. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-021-00595-5>
- Paulus, W., & Blossfeld, H.-P. (2007). Schichtspezifische Präferenzen oder sozioökonomisches Entscheidungskalkül? Zur Rolle elterlicher Bildungsaspirationen im Entscheidungsprozess beim Übergang von der Grundschule in die Sekundarstufe [Class-specific preferences or socio-economic decision-making? The role of parental educational aspirations in the decision-making process during the transition from primary to secondary school]. *Zeitschrift für Pädagogik*, 53(4), 491–508. <https://doi.org/10.25656/01:4407>
- Pfost, M., Rausch, T., Schiefer, I. M., & Artelt, C. (2018). Zur Entwicklung von Gymnasiastinnen und Gymnasiasten ohne Gymnasialempfehlung [On the development of grammar school students without a grammar school recommendation]. *Zeitschrift für Erziehungswissenschaft*, 21(3), 511–534. <https://doi.org/10.1007/s11618-017-0787-6>
- Pröscholdt, M. V., Stumpf, E., & Schneider, W. (2011). Das Arbeitsverhalten in homogenen Begabtenklassen [Working behavior in homogeneous gifted classes]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43(2), 55–67. <https://doi.org/10.1026/0049-8637/a000035>
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Relikowski, I., Schneider, T., & Blossfeld, H.-P. (2010). Primäre und sekundäre Herkunftseffekte beim Übergang in das gegliederte Schulsystem: Welche Rolle spielen soziale Klasse und Bildungstatus in Familien mit Migrationshintergrund? [Primary and secondary effects of origin on the transition to the structured school system: What role do social class and educational status play in families with a migration background?] In T. Beckers, K. W. Birkelbach, J. Hagenah, & U. Rosar (Eds.), *Komparative empirische Sozialforschung: Anwendungsfelder und aktuelle Methoden in BestPractice-Studien* (pp. 143–167). VS Verlag für Sozialwissenschaften.
- Rohde, T. E., & Thompson, L. A. (2007). Predicting Academic Achievement With Cognitive Ability. *Intelligence*, 35(1), 83–92. <https://doi.org/10.1016/j.intell.2006.05.004>
- Rost, D. H., & Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen* [Consequences of achievement groupings]. Waxmann.
- Rubin, D. B. (1987). Multiple imputation for nonresponse in urveys. John Wiley & Sons. <https://doi.org/10.1002/9780470316696>
- Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction*, 18(2), 201–210. <https://doi.org/10.1016/j.learninstruc.2007.09.017>
- Schilcher, A., Glondys, M., & Wild, J. (2023). Leseunterricht in den Grundschulen in Deutschland [Reading lessons in primary schools in Germany]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. Stubbe (Eds.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 179–196). Waxmann.
- Slavin, R. E. (2011). Instruction based on cooperative learning. In R. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 344–360). Routledge.
- Souvignier, E. (2018). Computerbasierte Lernverlaufsdiagnostik [Computer-based learning progress diagnostics]. *Lernen und Lernstörungen*, 7, 219–233. <https://doi.org/10.1024/2235-0977/a000240>
- Spieß, A. N. (2018). *Package ‘qpcR’: Modelling and Analysis of Real-Time PCR Dat*. R package version 1.4-1. <https://CRAN.R-project.org/package=qpcR>

- Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* [IQB trends in student achievement. Competencies in the subjects German and mathematics at the end of year 4 in the third country comparison]. Waxmann. <https://doi.org/10.25656/01:26077>
- Stanat, P., Schipolowski, S., Schneider, R., Weirich, S., Henschel, S., & Sachse, K. A. (2023). *IQB-Bildungstrend 2022. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich* [IQB trends in student achievement 2022. Language skills at the end of year 9 in the third country comparison]. Waxmann. <https://doi.org/10.31244/9783830997771>
- Statistisches Bundesamt (2022). *Bevölkerung und Erwerbstätigkeit: Erwerbsbeteiligung der Bevölkerung – Ergebnisse des Mikrozensus zum Arbeitsmarkt (Endgültige Ergebnisse)* [Population and occupation: Labor force participation of the population – Results of the Mikrozensus on the labor market (Final results)]. https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Erwerbsstaetigkeit/Publikationen/Downloads-Erwerbsstaetigkeit/erwerbsbeteiligung-bevoelkung-endergebnisse-2010410207004.pdf?__blob=publication-File
- Stocké, V. (2007). Explaining educational decision and effects of families' social class position: An empirical test of the Breen-Goldthorpe Model of educational attainment. *European Sociological Review*, 23(4), 505–519. <https://doi.org/10.1093/esr/jcm014>
- Stubbe, T. C., Bos, W., & Euen, B. (2012). Der Übergang von der Primar- in die Sekundarstufe [The transition from primary to secondary school]. In W. Bos, I. Tarelli, A. Bremerich-Vos, & K. Schwippert (Eds.), *IGLU 2011: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 209–226). Waxmann.
- Stubbe, T. C., Bos, W., & Schurig, M. (2017). Der Übergang von der Primar- in die Sekundarstufe [The transition from primary to secondary school]. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E. Lankes, N. McElvany, T. C. Stubbe, & R. Valtin (Eds.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 235–250). Waxmann.
- Stubbe, T. C., Kasper, D., & Jentsch, A. (2020). Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen [School career preferences at the transition to secondary school and the connection with performance-related and social characteristics]. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky, & H. Wendt (Eds.), *TIMSS 2019 Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. (pp. 315–330). Waxmann.
- Stubbe, T. C., Kleinkorres, R., Krieg, M., Schaufelberger, R., & Schlitter, T. (2023). Soziale und migrationsbedingte Disparitäten in der Lesekompetenz von Viertklässlerinnen und Viertklässlern [Social and migration-related disparities in the reading skills of fourth graders]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. Stubbe (Eds.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 151–178). Waxmann.
- Stubbe, T. C., Schaufelberger, R., Krieg, M., Kleinkorres, R., & Schlitter, T. (2023). Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen [School career preferences at the transition to secondary school and the connection with performance-related and social characteristics]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. Stubbe (Eds.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 231–248). Waxmann.
- Timmermans, A. C., van der Werf, M. G., & Rubie-Davies, C. M. (2019). The interpersonal Character of Teacher Expectations: The Perceived Teacher-Student Relationship as an Antecedent of Teachers' Track Recommendations. *Journal of School Psychology*, 73, 114–130. <https://doi.org/10.1016/j.jsp.2019.02.004>

- van Buuren, S., Groothuis-Oudshoorn, K., Robitzsch, A., Vink, G., Doove, L., & Jolani, S. (2015). *Package 'mice'*. [Computer software]. <https://CRAN.R-project.org/package=mice>
- von Davier, M., Mullis, I. V. S., Fishbein, B., & Foy, P. (2023). *Methods and Procedures: PIRLS 2021 Technical Report*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/pirls2021/methods>
- van Leest, A., Hornstra, L., van Tartwijk, J., & van de Pol, J. (2021). Test- or Judgement-Based School Track Recommendations. Equal Opportunities for Students with Different Socio-Economic Backgrounds? *British Journal of Educational Psychology*, 91(1), 193–216. <https://doi.org/10.1111/bjep.12356>
- Wentzel, K. R., Jablansky, S., & Scalise, N. (2019). Do Friendships Afford Academic Benefits? A Meta-Analytic Study. *Educational Psychology Review*, 30, 1241–1267. <https://doi.org/10.1007/s10648-018-9447-5>
- Wentzel, K. R., Jablansky, S., & Scalise, N. R. (2021). Peer Social Acceptance and Academic Achievement: A Meta-Analytic Study. *Journal of Educational Psychology*, 113(1), 157–180. <http://dx.doi.org/10.1037/edu0000468>
- White, I. R., Royston, P., & Wood, A. M. (2011). Multiple Imputation Using Chained Equations: Issues and Guidance for Practice. *Statistics in Medicine*, 30, 377–99. <https://doi.org/10.1002/sim.4067>
- Wößmann, L., Schoner, F., Freundl, V., & Pfaehler, F. (2023). Der ifo-„Ein Herz für Kinder“-Chancenmonitor: Wie (un-) gerecht sind die Bildungschancen von Kindern aus verschiedenen Familien in Deutschland verteilt? [The ifo “Ein Herz für Kinder” opportunity monitor: How (un)fairly are the educational opportunities of children from different families in Germany distributed?] *ifo Schnelldienst*, 76(04), 29–47.

Obtaining Abitur Without Teacher Recommendation – The Role of Social Background, Opportunity Costs, and Motivation

Katharina Molitor, Paul Fabian & Justine Stang-Rabrig

Zusammenfassung

Schüler*innen erhalten am Ende der Grundschulzeit eine Übergangsempfehlung für eine weiterführende Schulform von ihren Lehrkräften, die prädiktiv für die besuchte Schulform ist, aber sozial verzerrt sein kann. Ein erheblicher Anteil Schüler*innen besucht eine höhere Schulform als empfohlen und ist dort erfolgreich. Die Bedingungen für einen erfolgreichen Gymnasialbesuch sind allerdings bislang nur teilweise erforscht. Auf der Grundlage theoretischer Annahmen sind die soziale Herkunft (SES) und Motivation zentrale Prädiktoren für den Bildungserfolg und es kann angenommen werden, dass sie sich zwischen Schüler*innen mit und ohne Gymnasialempfehlung (GE) unterscheiden. Wir untersuchen, ob es systematische Unterschiede in Bezug auf SES (elterliche Bildung/Berufsstatus), Opportunitätskosten (finanzielle Kosten, Lernstress) und Motivation (Erfolgserwartung, wahrgenommene elterliche Aspirationen, Wert von Bildung, leistungsbezogene Motivation) zwischen Schüler*innen mit und ohne GE gibt. Darüber hinaus haben wir untersucht, ob sich die Beziehungen zwischen SES, Motivation und dem Erreichen des Abiturs zwischen diesen Gruppen unterscheiden. Eine MANOVA ergab einen im Mittel höheren SES und höhere Motivation bei Abiturient*innen mit GE und einen höheren Lernstress bei Abiturient*innen ohne GE. Strukturgleichungsmodelle zeigten, dass die elterliche Bildung für Schüler*innen ohne GE und Lernstress und Erfolgserwartung für Schüler*innen mit GE relevanter waren. Implikationen für Forschung und Praxis werden diskutiert.

Schlagworte: Bildungserfolg, Bildungsungleichheit, Erwartungs-Wert-Modell, Lehrkräfteeinschätzungen, Motivation

Abstract

At the end of elementary school, German students receive an enrollment recommendation by their teachers which affects the attended secondary school type but can be biased by, for example, their socio-economic status (SES). A substantial share of students choose to attend and are successful in a higher-track school than recommended. However, it is unclear, which factors are of relevance for them being successful on academic school tracks. Based on theoretical assumptions, SES and motivational factors are core predictors of educational attainment and can be assumed to differ between students with and without academic-track enrollment recommendations (ATER). We investigated whether there are systematic differences in terms of SES (parental education/occupation), opportunity costs (financial cost, learning stress), and motivational factors (students' expectation of success, perceived idealistic parental aspirations, value of education, performance-related motivation) between students with and without ATER. Further, we analyzed whether the relations between SES, opportunity costs, motivational factors, and obtaining *Abitur* differed between these groups. A MANOVA revealed an on average higher SES and motivation among students obtaining Abitur with ATER and for students who obtained Abitur without ATER. Structural equation models showed that parental education was more relevant for students without ATER, while learning stress and expectation of success were more relevant for students with ATER. Implications for research and practice are discussed.

Keywords: Educational attainment, educational inequality, expectancy-value model, teacher assessment, motivation

1. Introduction

The transition from elementary to secondary school has been shown to contribute significantly to educational inequalities in Germany (e.g., Gibson & Koch, 2022; Lohmann & Groh-Samberg, 2010). A key characteristic of the German education system is the early division into different school types. By the end of elementary school, students receive enrollment recommendations (ER) from their teachers for a secondary school type, which influence the attended school track. Ideally, the ER is based on performance in core subjects (math, German), motivation, and school-related behavior. However, the prognostic quality and validity of ER was repeatedly discussed, since ER are not only linked to students' performance, motivation, and achievement-re-

lated behavior, but also to their social background (SES) (e.g., Dumont et al., 2014). Inaccurate teacher judgements (e.g., Urhahne & Wijnia, 2021) might influence the ER. Knowing that the ER has long-lasting effects on students' further educational career, it is highly problematic that some students are assessed inaccurately and thus cannot achieve their full potential. However, students can be successful in academic-track schools (ATS) even if they lack the corresponding ER (e.g., Pfost et al., 2018). Nevertheless, it is unclear and of crucial importance what characterizes and helps students who attend a more ambitious school track than recommended being successful at this school type. Besides family background variables such as SES, motivation is a starting point, as it has been shown to promote educational attainment (e.g., Ditton et al., 2019).

Thus, in our study, we aimed to increase knowledge about students who did not receive an academic-track enrollment recommendation (ATER), but still obtained *Abitur*. Therefore, we firstly compared these students with students obtaining/not obtaining *Abitur* with and without ATER in terms of mean levels of SES and opportunity costs (parental education and occupational status, financial burden and learning stress) and individual motivational factors (perceived idealistic parental aspirations, expectation of success, value of education, and performance-related motivation). Secondly, we investigated whether the relations between SES, opportunity costs, and motivational factors and obtaining *Abitur* differed between students with and without ATER. Expanding upon previous research, we were able to track students until they left school.

2. Theoretical Background

2.1. The Transition to Secondary School in Germany

German students typically transfer to secondary school after grade four¹. The transition is based on teachers' ER students receive at the end of elementary school². The ER is based on (the development of) students' performance in core subjects (math and German) as well as students' motivation in terms of effort and school-related behavior (Kultusministerkonferenz [KMK], 2015). ERs are given for the most prevalent German school types: *Hauptschule*,

1 Except for the federal states of Berlin and Brandenburg, where elementary school lasts six years.

2 Except for the federal state of Bremen, where no enrollment recommendation is given.

Realschule (typically leading to vocational training), and *Gymnasium* (allowing students to continue on to university).

In different federal states, the ER is more or less binding for students' actual school enrollment. Mostly³, the final decision about the school type in which to enroll lies with the parents. In other federal states⁴, the ER is more binding and it is only possible to differ from the ER when taking performance tests or probation periods (KMK, 2015). The ER should reduce educational inequality, since it represents a teacher judgment based on supposedly objective performance criteria and uniform standards and is intended to counteract the influence of differential aspirations depending on parents' SES (KMK, 2015). However, it is problematic that these standards are not defined in concrete terms; there are no uniform regulations across federal states regarding the subjects or cut-off grades to be considered (KMK, 2015; Valtin, 2020). In practice, no inequality-reducing effect of the ER was found (Roth & Siegert, 2015).

2.2. Inequalities Arising From the Transition From Elementary to Secondary School

The transition to secondary school is seen as a factor generating and reinforcing educational inequalities in Germany, because the division in different school types happens very early compared to other countries (Lohmann & Groh-Samberg, 2010). These inequalities arise due to low educational upward mobility in the German school system and have consequences for students' further life course, as the attended school track after elementary school is a relevant predictor for the obtained school-leaving certificate and long-term educational and occupational outcomes (Stephens et al., 2015).

Educational inequalities resp. more precisely the relations between SES and educational outcomes can be explained by Boudons' (1974) theory of primary and secondary effects of social stratification. Primary effects relate to differences in students' performance depending on family background variables such as SES. Students' SES determines the availability of non-monetary resources (social and cultural capital; Bourdieu, 1973), meaning that, from birth onwards students are exposed to different stimuli (Schindler, 2018). The available resources determine childrens' development, abilities and later on their educational performance and attainment (Passaretta et al.,

3 Berlin, Hamburg, Hesse, Lower Saxony, Mecklenburg-Western Pomerania, North Rhine-Westphalia, Rhineland-Palatinate, and Schleswig-Holstein.

4 Baden-Württemberg, Bavaria, Brandenburg, Saarland, Saxony, Saxony-Anhalt, and Thuringia.

2022). Secondary effects refer to differences in educational decisions and aspirations caused by certain family background variables such as SES due to differences in knowledge about the educational system and therewith related expectancies to succeed on different school types, values attached to education as well as cost-benefit considerations depending on the available economic capital (Bourdieu, 1973). Boudons' (1974) theory was extended in terms of tertiary effects, firstly concerning inequalities in terms of students' immigrant background (Esser, 2016; Gresch, 2012), but later on also in terms of social inequalities (Blossfeld et al., 2019; Helbig & Morar, 2017). Next to focusing familiar prerequisites determining performance and educational decisions, tertiary effects consider effects of third parties such as teachers as important agents of socialisation. It is assumed that teachers convey socially unequal educational opportunities via grades and ER teacher assessments and teachers' perception of students' likelihood to succeed, as there are differences in i.e. grades students receive depending on their SES, although their performance in standardized competence tests is equal (Neugebauer, 2010).

The above described mechanisms operate at the transition to secondary school in the context of the ER, which can be socially distorted due to some issues (e.g., Dumont et al., 2019; Tiedemann & Billmann-Mahecha, 2010). Because of the relations between students' performance with their SES, primary effects are prevalent at the transition. Research showed that high-SES students obtained better grades and were more likely to receive higher school-leaving certificates (e.g., Broer et al., 2019). The relations between SES and achievement are already present by the end of elementary school (Bloh & van Ophuysen, 2008). Since the ER is, among other aspects, based on students' performance, high-SES students received ATER more often (Dumont et al., 2014). This relation is reinforced by tertiary effects – as already mentioned above, high-SES students receive better grades than their low-SES peers when they are at the same level of standardized test performance (Bittmann, 2021; Neugebauer, 2010).

Secondary effects can be found in deviations from the ER, which can be expected if parental aspirations differ from teachers' ER. Deviations are socially selective because parental aspirations differ by SES (e.g. Spera et al., 2009). Conversely, it is also possible to enroll in a lower-track school than recommended (enrolling in a non-ATS with ATER). Downward deviations are mainly found among low-SES students due to expectation of success and opportunity costs (Bloh & van Ophuysen, 2008). Opportunity cost encompass financial factors, which are relevant in this context, as ATS involve higher costs in terms of i.e. the subsequent university education compared to vocational training. But also without deviating from the teacher recommendation

secondary effects are present in terms of knowledge of the educational system. Parents can exert influence by contacting teachers and influencing them to recommend a higher school type. High-SES parents are most likely to do so, as they have high aspirations and wish to maintain their social status. This is particularly true in federal states where ER are more binding, as the freedom of choice of school type is more limited there (Dumont et al., 2019). Furthermore, learning stress is an opportunity cost to be considered, which is prevalent due to worse performance among low-SES students implying a higher learning load.

Tertiary effects become apparent when taking a closer look at the ER given by teachers in dependence of students' SES. While ER for high- and low-performing students were similar regardless of SES, the case was different for mid-performing students (Neugebauer, 2010). One reason might be the anticipated level of support that parents can provide, i.e. teachers believe high-SES students have a higher likelihood of success because they can draw upon a social support network if they face difficulties at school (Benner et al., 2016; Ditton & Krüsen, 2006).

2.3. Motivation and Educational Outcomes

These explanations raise the question which individual factors besides SES and opportunity costs promote educational attainment. One important factor that is strongly associated with educational outcomes is students' motivation. Hence, different aspects of students' motivation have emerged as important predictors for educational attainment (Ditton et al., 2019; Wang & Finch, 2018). Moreover, motivational factors are also considered alongside performance when teachers make their ER. Thus, motivation is of particular interest when explaining why students' might be successful without ATER.

Wigfield and Eccles' (2000) expectancy-value model of achievement motivation provides a comprehensive explanation for the relations between motivational factors and educational outcomes. Core elements in this model directly affecting achievement-related choices are students' expectation of success and the subjective task value. However, a broad set of further factors next to family background and opportunity costs indirectly affects achievement-related choices and outcomes, including previous achievement and perceptions of socializers' beliefs, expectations, and attitudes. In the context of this paper, the important motivational or motivational-related constructs students' expectation of success, perceived parental aspirations, value of education, and performance-related motivation are investigated.

There is strong evidence emphasizing the importance of expectation of success for performance, even when controlling for previous achievement (e.g., Wiederkehr et al., 2015). Also an intervention study showed that increasing expectation of success had a positive effect on performance among students who perceived their performance as worse than their peers' (Durik et al., 2015). Students' expectations of what they are capable of might be influenced by the ER they receive (Neuenschwander & Niederbacher, 2021). Furthermore, research on students with and without ATER found a lower self-concept among students enrolled in ATS without ATER (Bittmann, 2021; Pfof et al., 2018). Being in a high performing environment and due to the big-fish-little-pond effect (Marsh, 1987), students without ATER might rate their abilities and likelihood to succeed as lower than their peers. Also lack of teachers' beliefs of them being able to obtain Abitur is of relevance (Gentrup et al., 2020).

Another factor related to educational attainment are students' perceptions of socializers' expectations. Perceived idealistic parental aspirations can be a source of extrinsic motivation for students (Šimunović and Babarović, 2020). It has been shown that children aim to meet their parents' expectations in order to avoid punishment or obtain rewards (Cheung & Pomerantz, 2012). The final transition decision is made by parents and depends on their aspirations – especially among students without ATER perceived parental aspirations could be of relevance, as parents actively defy teachers' ER.

Further, subjective task value in terms of attainment value (Wigfield & Eccles, 2000) is another important predictor of educational attainment. In STEM (science, technology, engineering, and mathematics) subjects ascribing a high value to these subjects led to better performance (Steinmayr et al., 2012). Also having positive attitudes and high value to education in general positively predicted educational outcomes (Verešová & Mala, 2016). Enrollment in a higher school track than recommended might influence the relevance of students' value of education for their attainment.

Alongside these factors, students' performance motivation understood as the wish to perform well in school has also emerged as a central predictor for educational attainment (Ditton et al., 2019; Wang & Finch, 2018). As this study examines attainment of Abitur among students with and without ATER – a clearly performance-related goal – performance motivation is of major relevance.

2.4. Differences Between Students With and Without ATER

Several studies have identified that about 10–15% of students attend a higher-track school than recommended. A study by Tiedemann-Mahecha (2010) found an upward deviation of around 15% for Lower Saxony, Pfof et al. (2018) found an upward deviation for around 10% of students in Bavaria and Hesse, and Dudas (2014) found around 12% upward deviations for Hamburg. Depending on the analysis sample (incl./excl. multi-track schools), Bittmann (2021) showed for Germany – based on data from the German National Educational Panel Study (NEPS) – an upward deviation of 12%–18%. Focusing students who attend ATS without ATER, we can see that about 70% of them are successful, while approximately 90% of students with ATER are successful (Tiedemann & Billmann-Mahecha, 2010).

Comparing high- and low-SES students, Dudas (2014) found that the accuracy of teachers' ER differed by students' SES. While 87% of high-SES students who were suitable for ATS received a corresponding ER, only 58% of low-SES students received one. This is an indicator for tertiary effects being at work at the transition to secondary school. Comparing students at ATS with and without ATER, Pfof et al. (2018) found that students without ATER had lower grades and standardized math and reading scores than their peers with ATER in Grade 7. They also had a higher risk of transferring to a lower-track school. These findings imply a higher learning load and thus learning-related stress for students without ATER.

Overall, it can be stated that – although a higher share of students who receive an ER for the school type in which they actually enroll are ultimately successful in this school type – a large share of students without a corresponding ER is successful as well (Tiedemann & Billmann-Mahecha, 2010). However, it should be noted that being successful was defined as reaching Grade 7 or 8 without a grade retention; there is no information about the school-leaving certificates students ultimately obtained. Furthermore, only individual German federal states were taken into account in the analyses (Dudas, 2014; Pfof et al., 2018; Tiedemann & Billmann-Mahecha, 2010).

3. Research Questions and Hypotheses

We sought to close the following research gaps: First, as presented before, students' SES, opportunity costs, and motivation are relevant predictors of educational attainment. However, less is known about how students with vs. without ATER deviate in these important educational attainment-related aspects and whether the influences of these factors on educational attainment

differ among students with vs. without ATER being an important group to focus on. Second, we were able to track students until they obtained their school-leaving certificate and investigated SES, opportunity costs, and motivational variables simultaneously. This leads us to the following research questions and hypotheses:

Question 1: Are there on average systematic differences in terms of SES (parental education and occupation), opportunity costs (financial cost and learning stress), and motivational factors (expectation of success, perceived idealistic parental aspirations, value of education, learning motivation) between students successfully deviating upwards from the ER (obtaining Abitur without ATER (Group 1)) and students following the given ER (students obtaining Abitur with ATER (Group 2), students not obtaining Abitur without ATER (Group 3)) and students deviating downwards from the given ER (students not obtaining Abitur with ATER (Group 4))?

In terms of SES we expect the following group differences: Students who obtained Abitur without ATER (Group 1) on average H1a1) have a lower SES (parental education and occupational status) than students obtaining Abitur with ATER (Group 2), H1a2) have a higher SES than students not obtaining Abitur without ATER (Group 3) and H1a3) students not obtaining Abitur with ATER (Group 4).

For opportunity costs we expect the following results: Students who obtained Abitur without ATER (Group 1) on average have higher scores on financial burden and learning stress than H1b1) students obtaining Abitur with ATER (Group 2), H1b2) students not obtaining Abitur without ATER (Group 3) and H1b3) students not obtaining Abitur with ATER (Group 4).

Regarding expectation of success and perceived idealistic parental aspirations, we expect differences as following: Students obtaining Abitur without ATER (Group 1) do on average H1c1) score lower on expectation of success and higher on perceived idealistic parental aspirations than students obtaining Abitur with ATER (Group 2), H1c2) score higher on expectation of success and higher on perceived idealistic parental aspirations than students not obtaining Abitur without ATER (Group 3) and H1c3) students not obtaining Abitur with ATER (Group 4).

For value of education and learning motivation we expected the groups to differ in the following ways: Students who obtained Abitur without ATER (Group 1) do on average value education more and have higher learning motivation than H1d1) students obtaining Abitur with ATER (Group 2), H1d2) students not obtaining Abitur without ATER (Group 3) and H1d3) students not obtaining Abitur with ATER (Group 4).

Question 2: How do the relations between SES (parental education and occupation), opportunity costs (financial cost and learning stress), and motivational factors (expectation of success, perceived idealistic parental aspirations, value of education, learning motivation), and obtaining Abitur differ between students with and without ATER?

We expect the following relations for students with and without ATER with obtaining Abitur:

Among students without ATER, obtaining Abitur is predicted H2a) more strongly by SES (parental education and occupational status), H2b) opportunity costs (financial burden and learning stress), H2c) equally strong by the motivational factors expectation of success and more strongly by perceived idealistic parental aspirations, H2d) more weakly by motivational factors value of education and learning motivation than among students with ATER.

4. Method

4.1. Data, Sample, and Procedure

We used data from the NEPS starting cohort Grade 5 (SC3) (NEPS Network, 2023)⁵. NEPS is a longitudinal, multi-cohort large-scale survey providing information on educational trajectories as well as student, parent, and teacher assessment. Students who changed schools or left the educational system were individually tracked and interviewed. Our sample consisted of students for whom information on the ER was available. This information was gathered from parents. We thus excluded $N = 1,328$ students without parent interview or missing information on the ER. Since no ER is given in Bremen, students from Bremen were not in the analyses as well. In line with previous conducted studies (Dudas, 2014; Pfof et al., 2018; Tidemann & Billmann-Mahecha, 2010), also students attending comprehensive schools were excluded. The first included measurement was winter 2012/13, at which point students were attending Grade 7. Our sample in Grade 7 thus comprised $N = 2,671$ students, with a mean age of 12.5 ($SD = 0.84$) years, whereof 49% were female. Student questionnaires and competence tests were conducted in paper and pencil interviews (PAPI) mainly between Grades 7 and 9, while parent surveys were interviewed via telephone (CATI-mode) in Grade 7.

5 This paper uses data from the National Educational Panel Study (NEPS; see Blossfeld & Roßbach, 2019). The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network.

4.2. Variables⁶

Educational Attainment

Educational attainment was measured by the highest school-leaving certificate obtained. The variable was binary-coded for our analyses (0 = *no Abitur*, 1 = *Abitur*) and was based on a variable providing information on the highest school-leaving certificate obtained at various measurement points. Approximately 61.8% of our sample left school with Abitur. This share is higher than among all German students (34%; Statistical Federal Office [Destatis], 2023).

Enrollment Recommendation

Information on ER were collected retrospectively from parents in Grade 7 – the later transition to secondary school in the federal states of Berlin and Brandenburg thus does not lead to methodological problems. This variable was coded in binary form (0 = *no ATER*, 1 = *ATER*). 62.4% of the students in our sample received ATER at the end of elementary school.

SES

Students' SES was operationalized using *parental education* and *occupational status*. This information was collected from parents in Grade 7. Information for both measures were available for the interviewed parent and their partner, out of which we considered the highest educational level. We used the International Standard Classification of Education (ISCED) (United Nations Educational Scientific and Cultural Organization [UNESCO], 1997) to operationalize parental education; the scale ranged from 0 = *early childhood education* to 9 = *doctoral degree or equivalent*. To operationalize parental occupational status, we used the International Socio-Economic Index of Occupational Status (ISEI), which ranges from 10 (subsistence farmers, fishers, hunters, and gatherers) to 90 (judges) (Ganzeboom, 2010).

Opportunity Costs

Financial burden was measured in Grade 8 by one item on a five-point Likert scale from 1 = *very difficult* to 5 = *very easy* asking students: "As long as you go to school, you can hardly earn your own money. Your parents pay most of

6 Scale parameters (means, standard deviations) for both groups can be found in Table 2, resp. in the supplementary material for control variables (Table A2).

what you need (e.g., school supplies, clothes). How hard would it be for your parents to cover these costs if you obtained an Abitur?”. *Learning stress* was also measured in Grade 8 with one item on a five-point Likert scale from 1 = *very low* to 5 = *very high*. Students were asked: “The different school-leaving qualifications require different amounts of effort. How much effort would it take you to acquire the Abitur?”.

Motivational Factors

Students’ motivation was measured by four distinct constructs. Students’ *expectation of success* in terms of passing Abitur was measured in Grade 8 by the following question: “Regardless of the qualifications that you can actually obtain at your school: How likely do you think it is that you could pass the Abitur examination?”. Responses were recorded on a five-point Likert scale from 1 = *very unlikely* to 5 = *very likely*. *Perceived idealistic parental aspirations*⁷ covering perceived parental future-related academic wishes of the parents for the child were assessed by students in Grade 9 with one item (Hoenig et al., 2016). Students were asked: “The following question deals with what your parents expect from you in school and for your professional career. What highest school-leaving qualification would your parents like you to have?” This variable was binary-coded for analyses (0 = *no Abitur*, 1 = *Abitur*). The *value of education* was measured with five items, e.g., “Having a good education is valuable in its own right” (Stocké, 2005). The items were assessed in Grade 7 and answered on a five-point Likert scale from 1 = *fully disagree* to 5 = *fully agree*. The reliability of this scale was satisfying for both groups (with ATER: $\alpha = .69$, without ATER: $\alpha = .71$). Students’ *performance-related motivation* was measured with four items, e.g., “I learn for school because I want to be one of the best.” (Schiefele et al., 2002) covering students willingness to actively and purposefully engage with learning content. These items were assessed in Grade 8 and responses were recorded on a five-point Likert scale from 1 = *fully disagree* to 5 = *fully agree*. The reliability of this scale was good for both groups (with ATER: $\alpha = .84$, without ATER: $\alpha = .82$).

Control Variables

A set of control variables was included in the analyses. Students’ *language spoken at home* needed to be considered, since immigrant students on the one hand tend to perform on average worse in school compared to students without an immigrant background but on the other hand show higher scores

7 Hereafter: parental aspirations.

on aspirations and motivation (McElvany et al., 2018). Responsible for this mismatch are lacks in the knowledge of the language of instruction (Organization for Economic Co-operation and Development [OECD], 2019a), which is why we included home language in the analyses. Measured in Grade 7, students were defined as German speakers when they only spoke German with both adults in their household (0 = *home-language German*, 1 = *home-language non-German*)⁸. We also controlled for the *bindingness of the enrollment recommendation*. We therefore coded the federal states by their degree of bindingness at the time this cohort of students transitioned to secondary school (school year 2010/11) (0 = *non-binding recommendation*, 1 = *binding recommendation*).

Furthermore, Wigfield and Eccles (2000) theorized that previous performance is important in explaining students' educational attainment. Thus, we included WLE scores for *math* and *reading competences* measured in Grade 7 in our analyses (Duchhardt & Gerdes, 2013; Haberkorn et al., 2012). Both measures were standardized to a mean of 0 and ranged from -3.66 to 4.81 (math competences) and -4.33 to 5.78 (reading competences) in our dataset. Additionally, students' *basic cognitive skills*, which represent a factor crucial for students' performance and educational attainment, were controlled for (Roth et al., 2015). A reasoning assessment (DGCF) that ranged from 0 to 12 was measured in Grade 7 (Haberkorn & Pohl, 2013). Also *gender* was regarded, as relations between gender and educational attainment have been demonstrated (OECD, 2019a). Information on students' gender were measured in Grade 7 (0 = *male*, 1 = *female*). Also, the attended *school type* is an important predictor when investigating obtaining Abitur, for which we controlled in binary form (0 = *non-ATS*, 1 = *ATS*). We controlled for the highest attended school type during secondary school time. This measure stems from various measurement points. Students attending comprehensive schools had to be excluded, as no identification of school-track was possible.

8 In Germany, a person is considered having an immigrant background if they or at least one parent were not born with German citizenship (Destatis, 2024). Based on this, there is various research that uses the country of birth instead of nationality; corresponding definitions can be found at the OECD or the European Commission (OECD/EU (2019b) or also at IGLU (Stubbe et al., 2023). A child has a immigrant background if he or she or at least one parent was born abroad. Similarly, the definition used here is that the child has a non-German language of interaction if they do not speak German exclusively with at least one parent. The share of pupils with "home language non-German" (14.1%) is comparable to the share of pupils with a migrant background based on the above definition (13.9%).

4.3. Analyses

We used Stata 17 (StataCorp LLC, 2021) for our analyses. To answer our first research question, we conducted mean value comparisons (MANOVA) for all included variables, comparing four groups of students: students obtaining Abitur without ATER (Group 1), students obtaining Abitur with ATER (Group 2), students not obtaining Abitur without ATER (Group 3) and students not obtaining Abitur with ATER (Group 4). We compared Group 1 against the three other groups. To answer our second research question, we estimated four structural equation models (SEM) comparing the two groups of students (students with ATER vs. without ATER) in terms of educational attainment, i.e. whether they obtained Abitur or not. Model fit in Model 4 was assessed using χ^2 , RMSEA, CFI (Shi et al., 2020), and R^2 (Models 1–3 were saturated). We included the set of factors one at a time in a stepwise process⁹. Taking the hierarchical structure of the data (students nested within classes nested within schools) into account was not possible on the class level, since classes are not consistent across measurement points in the NEPS data. However, we clustered standard errors by school identifiers and controlled for school type. Standardized coefficients are reported. Value of education and performance-related motivation were modelled latently in the analyses. All SEMs included all control variables and all variables were correlated with one another. Measurement coefficients and intercepts were specified to be equal for both groups; other parameters were allowed to vary. Sample size was held constant for all models. Although the outcome variable is binary, we decided to estimate linear models due to the better interpretability of linear regression coefficients compared to logistic regression coefficients and due to the fact that the error produced by this procedure is usually very small (Gomila, 2019).

5. Results

5.1. Descriptive Statistics

We included 2,671 students in our analyses. The majority of students in our sample obtained a school-leaving certificate matching with the received ER – 51.4% obtained Abitur with ATER (Group 2), 27.3% did not obtain Abitur with no ATER (Group 3). The groups deviating from the ER in obtaining Abitur

9 We estimated mediation models using the motivational factors as mediators. Although RMSEA values were good, these models had an insufficient fit in terms of CFI, which was 0.84 or less. Thus, we decided not to report the mediation models.

without ATER (10.3%) (Group 1) and not obtaining Abitur with ATER (11.0%) (Group 4) were almost equally large.

All included variables were significantly correlated with obtaining Abitur among students with and without ATER, except for performance-related motivation (Table 1, Table A1 in the supplementary for all variables incl. control variables). Parental education, occupation, and learning stress were likewise strongly correlated with obtaining Abitur. Among the motivational variables, expectation of success had the strongest correlation with obtaining Abitur in both groups, while parental aspirations were equally relevant for students without ATER.

Table 1: Intercorrelations Among all Analyzed Variables in our Sample for Students With (Upper Right) and Without (Lower Left) ATER

Variable	1	2	3	4	5	6	7	8	9
1 Abitur	–	.21	.20	.05	-.05	.34	-.12	.12	.04
2 PE	.21	–	.61	.07	-.05	.20	-.07	.03	-.06
3 OS	.20	.58	–	.07	-.05	.20	-.07	.03	-.06
4 CosFin	.12	.09	.11	–	.05	.15	.03	.09	.08
5 CosLea	.07	.04	.03	.20	–	-.21	-.10	-.02	-.05
6 EOS	.28	.12	.13	.16	.02		.22	.14	.19
7 ParAsp	.29	.10	.14	.12	.02	.36	–	.17	.11
8 VOE	.08	.02	.03	.02	-.02	.23	.22	–	.21
9 PM	.03	-.01	-.03	.11	-.02	.20	.11	.23	–

Note. ATER = academic-track enrollment recommendation; PE = parental educational level; OS = parental occupational status; CosFin = financial cost; CosLea = Cost of learning stress; EOS = expectation of success; ParAsp = parental aspirations; VOE = value of education; PM = performance-related motivation. Significant correlations ($p < .05$) are in bold.

5.2. Mean Comparison of Students With and Without ATER

To determine differences between students who obtained Abitur without ATER (Group 1) and students 1) who obtained Abitur with ATER (Group 2), 2) who did not obtain Abitur without ATER (Group 3), and 3) who did not obtain Abitur with ATER (Group 4), we conducted a MANOVA. The MANOVA revealed statistically significant differences between the analyzed groups (see Table 2, $F(48, 3897) = 43.32, p < .001$, Wilk’s $\Lambda = .285$).

Table 2: Mean Comparison of all Analyzed Variables (Except for Control Variables), Standard Deviations (*SD*) and *n* of Respondents by Enrollment Recommendation and School-leaving Certificate

Variable	Group 1		Group 2		$\Delta M p$	Group 3		$\Delta M p$	Group 4		$\Delta M p$
	<i>M</i> (<i>SD</i>)	<i>n</i>	<i>M</i> (<i>SD</i>)	<i>n</i>		<i>M</i> (<i>SD</i>)	<i>n</i>		<i>M</i> (<i>SD</i>)	<i>n</i>	
PE	6.85 (2.07)	302	7.72 (2.07)	1,502	<.001	5.64 (2.40)	798	<.001	6.54 (2.34)	320	.054
OS	56.99 (17.80)	298	66.35 (16.65)	1,492	<.001	48.58 (18.76)	789	<.001	56.97 (18.92)	316	.494
CosFin	4.23 (0.91)	251	4.11 (0.94)	1,278	.262	3.94 (1.14)	593	.003	3.97 (1.01)	251	.003
CosLea	4.66 (0.57)	267	4.47 (0.69)	1,376	<.001	4.53 (0.91)	622	.112	4.58 (0.67)	265	.732
EOS	3.26 (0.93)	269	4.06 (0.71)	1,379	<.001	2.62 (1.04)	619	<.001	3.31 (0.99)	263	.717
ParAsp	0.54 (0.50)	251	0.77 (0.42)	1,326	<.001	0.24 (0.43)	611	<.001	0.36 (0.48)	238	<.001
VOE	3.02 (0.86)	271	3.25 (0.84)	1,415	.010	2.86 (0.89)	676	.004	2.95 (0.90)	273	.304
PM	2.57 (0.86)	272	2.44 (0.64)	1,390	.202	2.52 (0.62)	643	.972	2.37 (0.61)	272	.173

Note. Group 1: Students obtaining Abitur without ATER, Group 2: Students obtaining Abitur with ATER, Group 3: Students not obtaining Abitur without ATER, Group 4: Students not obtaining Abitur with ATER. ATER = academic-track enrollment recommendation; PE = parental educational level; OS = parental occupational status; CosFin = financial cost; CosLea = Cost of learning stress; EOS = expectation of success; ParAsp = parental aspirations; VOE = value of education; PM = performance-related motivation.

In terms of SES, the average SES among students who obtained Abitur with ATER was higher than that of their peers without ATER, confirming H1a1 (Table 2). The difference was significant for parental education and occupational status. Comparing Groups 1 and 3, an on average higher mean was found for Group 1 for both SES measures, confirming H1a2 as well. Contradicting H1a3, no statistically significant group differences between students deviating upwards vs. downwards from the ER were found. For opportunity costs, statistically significant differences between Groups 1 and 2 were found for financial burden, as presumed in H1b1. Although – in line with H1b2 and H2b3 – the results indicate that financial costs are higher in Group 1 than in the Groups 3 and 4, no statistically significant differences were found. Students in Group 1 exhibited on average lower values on expectation of success but also parental aspirations than students in Group 2, partly confirming H1c1. Compared to Groups 3 and 4, there is a shift towards higher expectation of success and parental aspirations in Group 1. Except for the means for Group 1 and Group 4 in terms of expectations of success, the mean differences are significant, partly confirming H1c2 and H1c3. We also observed that the mean value of education was statistically significantly higher for Group 1 compared to Group 2, while no significant differences were found for performance motivation, contradicting H1d1. Compared to Group 3 value of education on

average is higher in Group 1. For performance motivation, again, no statistically significant differences were found. H1d2 thus can be partly confirmed. Comparing groups 1 and 4, contradicting H1d3 no group differences in terms of value of education and performance motivation were found.

5.3. Results of the Group Comparison via Structural Equation Models

To answer the question of whether relations between SES, opportunity costs, motivational factors and obtaining Abitur differed between students with and without ATER, we specified four structural equation models (see Table 3). In Model 1, significant positive relations between parental education and obtaining Abitur were found in both groups. The relation was significantly stronger for students without ATER ($\Delta\chi^2 = 15.3$, $\Delta df = 2$, $p < .001$), implying that for students without ATER parental education is more important for obtaining Abitur than for students with ATER. This partially supports H2a, as no significant relations for parental occupation were found in either group.

In Model 2, learning stress and financial burden were included in the model. Both variables did not significantly predict obtaining Abitur in either group. Comparing Models 1 and 2 in terms of R^2 , slightly more variance was explained ($R^2_{\text{Model 1}} = .221$, $R^2_{\text{Model 2}} = .230$).

In Model 3, we included expectation of success and parental aspirations. Both factors significantly positively predicted obtaining Abitur. While parental aspirations were an equally strong predictor in both groups, the positive impact of expectation of success was significantly more relevant for students with ATER ($\Delta\chi^2 = 6.5$, $\Delta df = 2$, $p = .035$), meaning that students beliefs in themselves is of particular importance in this group.

Table 3: Results of the Structural Equation Models for the Relations Between SES, Opportunity Costs, Motivation and Obtaining Abitur for Students With and Without ATER

Variable	Model 1		Model 2		Model 3		Model 4	
	β	SE	β	SE	β	SE	β	SE
	with ATER/ without ATER		with ATER/ without ATER		with ATER/ without ATER		with ATER/ without ATER	
PE	.07*/.16***	.03/.04	.07*/.16***	.03/.04	.05/.14***	.03/.04	.05/.15***	.03/.04
OS	.05/.06	.03/.04	.04/.05	.03/.04	.00/.03	.03/.04	.01/.02	.03/.04
Cost financial			-.01/.05	.02/.04	-.02/.02	.02/.04	-.03/.03	.03/.04
Cost learning			.01/.05	.03/.04	.06*/.05	.03/.04	.07**/.05	.03/.04
EOS					.18***/.11**	.03/.04	.18***/.10*	.03/.04
Parental aspirations					.13***/.11**	.03/.04	.12***/.07	.03/.04
VOE							.08*/.09*	.03/.04
PM							.01/.05	.03/.05
Home language	.06*/.05	.03/.03	.05/.05	.03/.03	.03/.01	.03/.04	.01/.01	.03/.04
Bindingness ER	-.06**/-.02	.02/.03	-.06**/-.02	.02/.03	-.04/-.03	.02/.03	-.03/-.01	.02/.03
Gender	.07**/.12***	.02/.03	.07**/.11***	.03/.03	.06*/.11***	.02/.03	.07*/.11**	.03/.04
School type	.43***/.19***	.02/.03	.45***/.20***	.02/.03	.39***/.17***	.02/.03	.36**/.21***	.03/.04
Cognitive abilities	.12***/.06	.03/.04	.11***/.04	.03/.04	.11***/.04	.03/.04	.10**/.05	.03/.04
Math competences	.03/.02	.03/.04	.04/.02	.03/.05	.00/.03	.03/.05	.02/.02	.04/.05
Reading competences	.08**/.12**	.03/.04	.07*/.11**	.03/.04	.06/.10*	.03/.04	.07*/.12**	.03/.05
RMSEA	-		-		-		.037	
CFI	-		-		-		.939	
df	-		-		-		302	
χ^2	-		-		-		839.5***	
R ²	.221		.230		.271		.274	
N	1,634/1,037		1,634/1,037		1,634/1,037		1,634/1,037	

Note. PE = parental educational level (0 = *early childhood education* to 9 = *doctoral or equivalent level*); OS = parental occupational status (10 = *subsistence farmers, fishers, hunters, and gatherers* to 90 = *judges*); Cost financial (1 = *very hard* to 5 = *very easy*); Cost learning (1 = *very low* to 5 = *very high*); EOS = expectation of success (1 = *very unlikely* to 5 = *very likely*); Parental aspirations (0 = *no Abitur*, 1 = *Abitur*); VOE = value of education (1 = *fully disagree* to 5 = *fully agree*); PM = performance-related motivation (1 = *fully disagree* to 5 = *fully agree*); Home language (1 = *home-language non-German*); Bindingness ER (1 = *binding recommendation*); Gender (1 = *female*); School type (1 = *ATS*); Cognitive abilities (0 to 12); Math competence (3.66 to 4.81); Reading competence (-4.33 to 5.78). RMSEA = root mean square error of approximation. CFI = comparative fit index. df = degrees of freedom. Standardized coefficients. Model identification was achieved by setting the loading of the first indicator to 1.

* $p < .05$. ** $p < .01$. *** $p < .001$. Coefficients significantly differing between groups are in bold.

H2c had to be rejected, as neither the stronger predictive power of parental aspiration for students without ATER nor the equally strong relations for expectation of success for both groups were found. After including expecta-

tion of success and parental aspirations, changes in the previously reported relationships were observed. The relation between parental education and obtaining Abitur became insignificant for students with ATER, even though the coefficient did not change much in size, while the relation for students without ATER remained similar – still indicating a positive relation between parental education and obtaining Abitur in this group. Furthermore, the relation between learning stress and obtaining Abitur became significant for students with ATER. Model 3 explained a higher amount of variance compared to Model 2 ($R^2_{\text{Model 2}} = .230$, $R^2_{\text{Model 3}} = .271$).

In Model 4, we included value of education and performance motivation. χ^2 , RMSEA, and CFI indicated a good model fit. The expected stronger relations between value of education, performance motivation and obtaining Abitur for students without ATER (H2d) could not be confirmed. Although value of education was a relevant positive predictor for obtaining Abitur, no group differences were found. Both groups benefitted likewise from a high value of education. Performance motivation was no significant predictor in either group. After including value of education and learning motivation, further changes in the previously reported relations were observed. The associations between parental aspirations and obtaining Abitur became insignificant for students without ATER – however, the group difference was not significant ($\Delta\chi^2 = 0.57$, $\Delta df = 2$, $p = .753$). Model 4 explained a slightly higher amount of variance compared to Model 3 ($R^2_{\text{Model 3}} = .271$, $R^2_{\text{Model 4}} = .274$).

6. Discussion

Using data from NEPS starting cohort Grade 5 (SC3), our study aimed to investigate whether students with or without ATER differed concerning their SES, opportunity costs, and core motivational factors. Moreover, we examined whether different factors are of relevance for obtaining Abitur among students with and without ATER. The results contribute to previous findings and go beyond them in terms of identifying differences in central characteristics as SES, opportunity costs, and motivation and identifying the different important factors for obtaining Abitur in different groups, that are simultaneously considered.

Also in our sample, a substantial share of students deviated from the ER and was successful on a higher school type. The share of students successfully attending ATS without ATER was lower than in previous research focusing students in Grade 7 and thus, also another outcome. When comparing with other studies tracking students until the end of secondary school time, our

results are in line with previous research (Bloh & van Ophuysen, 2008; Pfost et al., 2018).

Summarizing the findings concerning students' SES for both research questions, students with and without ATER who obtained Abitur differed significantly regarding SES. Students with ATER had a higher SES than their peers. As the conditions for enrollment in an ATS were more favorable for high-SES students, this finding is in line with H1a1 and previous research (Bloh & van Ophuysen, 2008). When predicting educational attainment (H2a), we observed a higher relevance of parental education for students without ATER, underlining the relevance of SES for their educational outcomes. This result indicates that parental education is particular importance for students without ATER to obtain Abitur, as this is usually associated with higher SES and can therefore mean that students who may not have received ATER due to lower performance at the end of primary school are dependent on support in the form of tutoring, which parents with higher SES, i.e. also higher education, are more likely to be able to finance. Concerning opportunity costs, we found the following results: Students obtaining Abitur with and without ATER did not differ statistically significantly in financial burden, but the results confirmed our assumptions in tendency. However, learning stress was perceived as higher among students who obtained Abitur without ATER. This might be explained via performance, which tends to be lower among low-SES students and students deviating upwards from the ER (Pfof et al., 2018), meaning that this special group of students (obtaining Abitur without ATER) may have to invest more time in learning in order to obtain Abitur (Broer et al., 2019). Comparing our reference group to students not obtaining Abitur in tendency higher financial cost was detected. The mean differences in terms of opportunity costs showed that secondary effects of social stratification (Boudon, 1974) are relevant in our sample. Although students without ATER perceived more learning stress, this factor was not more relevant for obtaining Abitur (H2b). Quite oppositely, we observed that higher learning stress positively predicted obtaining Abitur in this group. This – at first glance – counterintuitive finding might be explained by the personal importance of learning and educational success. Learning stress must not be perceived negatively; it is also possible that it motivates students to learn (Saklofske et al., 2012). A non-linear relation is also possible – learning stress might be a positive predictor up to a certain point, but might turn negative at higher levels.

The findings concerning students' expectation of success were in line with H1c1. Students might believe more in themselves after receiving ATER, as they perceive that their teachers, who are important agents of socialization, believing in them to be able to perform well (Gentrup et al., 2020). Students

not receiving ATER perceived themselves as capable of obtaining Abitur less often. As they also scored lower on competences than their peers, the lower scores on their expectations of success might be due to the big-fish-little-pond effect (Marsh 1987) resp. the little-fish-big-pond effect (Zell & Lesick 2021), as they are in a high performing learning environment. The expectations of success did not significantly differ between the two groups deviating from the ER – this can be interpreted as the ER shaping students' expectations and students' beliefs about being able to obtain Abitur (Neuenschwander & Niederbacher, 2021).

The results in terms of parental aspirations contradict the assumptions made in H1c1, as students, obtaining Abitur with ATER, perceived parental aspirations as higher. However, considering the bias towards high-SES students in our sample, an explanation could be that parents with high aspirations contacted teachers before giving ER in order to encourage ATER (Dumont et al., 2019), which reinforces the impression that secondary effects are evident at the transition. However, it is an open question whether parents alone made the enrollment decision or if the students also had an influence (e.g., Wohlkinger, 2014). In this case, it is plausible that parental aspirations were less relevant for students who obtained Abitur without ATER.

Expectation of success and parental aspirations positively predicted obtaining Abitur for both groups, but were more relevant for students with ATER, contradicting H2c. Thus, among more privileged students (students with ATER, who tend to have a higher SES), motivational factors seem to develop greater relevance. Also having received ATER might reinforce students' beliefs in themselves and thus promote expectations of success being more relevant in this group.

Contradicting H1d1, value of education was higher among students obtaining Abitur with ATER compared to students obtaining Abitur without ATER. As status maintenance – which is also a factor for educational decisions and part of secondary effects of social stratification – is a relevant factor for high-SES parents, this finding is in line with this (Dumont et al., 2019). As for expectation of success no group differences were found between the two groups deviating from the ER. Even though value of education was a significant predictor in both groups, contradicting H2d no group differences were found, meaning that value of education is a likewise relevant motivational predictor in both groups and thus underlines its general importance for educational attainment (Meyer et al., 2024).

Summarizing our findings, students with and without ATER differ in terms of the factors predicting Abitur. While students' SES (parental education) was more relevant for students without ATER, motivational factors

(expectation of success and parental aspirations) were more relevant for students with ATER. Educational attainment among students without ATER was thus primarily determined by their parents' education, implying that low-SES students without ATER might be doubly disadvantaged. Nevertheless, motivation in terms of value of education and expectation of success was a relevant predictor for educational attainment and promoting students' motivation is a key way to foster educational attainment, this applies to all students. When attempting to reduce educational inequality among students without ATER, it would be interesting to examine how their somewhat higher motivation can be translated into higher educational attainment. This could be especially relevant for students without ATER who did not deviate from their ER, even though they theoretically had the potential to obtain Abitur.

6.1. Limitations and Strengths

When interpreting the results, some limitations must be kept in mind. Although working with a large sample, many students were excluded from our analyses due to missing information on ER. Thus, students from the federal state of Bremen were excluded, as no enrollment recommendation was given. Furthermore, in line with previous research investigating and comparing outcomes of students with and without ATER, students attending multi-track schools were not part of our sample (Dudas, 2014; Pfost et al., 2018; Tidemann & Billmann-Mahecha, 2010). Due to the exclusions, our sample is biased towards a high share of students receiving ATER and obtaining Abitur. The results thus have to be interpreted in light of a sample with favorable outcomes, which might over- or underestimate the impact of some predictors. However, including students from Bremen and from multi-track schools might lead to problems in interpreting the results as well, as it is not clear what is actually measured, as i.e. no allocation to an academic resp. non-academic track is possible, and results might be biased due to this. Thus, we decided against imputing missing data.

Next, no control on the class level was possible, as we used NEPS data which do not measure the classroom context and associated effects (e.g. Ladd et al., 2009). Also due to data structure, it was not possible to analyze individual pathways of school-track changes, which is why we included attending an ATS as a general measure. One strength of our study is that, contrasting previous studies, data allowed us to observe students until the end of their school careers. This contributed to a deeper understanding of the school-leaving certificates students ultimately obtain and the mechanisms contributing to

the attainment of a higher school-leaving certificate among students without ATER. Another strength is that we applied a comprehensive approach with two variables measuring SES and opportunity cost and four core motivational factors to obtain a more detailed picture and detect possible starting points for supporting higher educational attainment among students without ATER.

6.2. Implications for Research and Practice

Further research should identify factors beyond parental education that promote educational attainment and educational inequalities among incorrectly assessed students. The focus on different subgroups could be a starting point for that. Moreover, further variables such as parental support or students' well-being (Diener et al., 2018) or teaching quality (Klieme & Rakoczy, 2008) could be included to explain the educational attainment of students without ATER. This would extend the focus to another important institution (school) or socialization agent (teachers) that is also highly relevant for students' educational outcomes. However, attending *Gymnasium* is the most common, but not the only way of obtaining Abitur. Especially when focusing low-SES children, these alternative, non-traditional ways should be further investigated. In this context comprehensive schools, which do have a higher horizontal permeability than school tracks in the highly stratified educational system are of interest for further research, as well as obtaining the qualification to take up tertiary education via vocational schools or occupational qualification. Furthermore, a focus should be set on how to improve teachers' evaluations that lead to ER.

Some implications for practice arise from our findings as well. On the individual level, as evident in the expectancy-value model (Wigfield & Eccles, 2000, Ditton et al., 2019), expectation of success is highly relevant for educational success. Thus, interventions to foster students' expectations as well as their beliefs in themselves might be a good option to support students without ATER in particular. Moreover, the value such students attach to education could be strengthened and might in turn lead to more equal opportunities in the educational system. On the institutional level, the point at which students are divided into separate tracks should be discussed critically, as many relevant prerequisites for educational attainment develop later than the end of elementary school, thus it is difficult to make predictions about students' future development (Valtin, 2020). Longer periods of comprehensive schooling or greater flexibility between school tracks could be solutions here. Furthermore, teachers do not receive long-term feedback on the accuracy of their ER.

Providing teachers with feedback on whether the ER they made were accurate could improve the quality of their predictions in the long run (Dudas, 2014).

References

- Benner, A. D.; Boyle, A. E.; Sadler, S. (2016): Parental Involvement and Adolescents' Educational Success: The Roles of Prior Achievement and Socioeconomic Status. *Journal of Youth Adolescence*, 45, 1053–1064. <https://doi.org/10.1007/s10964-016-0431-4>
- Bittmann, F. (2021). Academic track mismatch and the temporal development of well-being and competences in German secondary education. *Vienna Yearbook of Population Research*, 19, 467–502. <https://www.jstor.org/stable/27144766>
- Bloh, B., & van Ophuysen, S. (2008). Was bedingt die Wahl eines nicht empfohlenen höheren Bildungsgangs [What are the reasons for choosing a non-recommended higher education course]? *Zeitschrift Für Erziehungswissenschaft*, 11, 626–647. <https://doi.org/10.1007/s11618-008-0037-z>
- Blossfeld, H.-P., Blossfeld, G. J., & Blossfeld, P. N. (2019). Soziale Ungleichheiten und Bildungsentscheidungen im Lebensverlauf. Die Perspektive der Bildungssoziologie [Social inequality and educational decisions over the life course. An educational sociology perspective]. *Journal for Educational Research Online*, 11, 16–30. <https://doi.org/10.25656/01:16785>
- Blossfeld, H.-P., & Roßbach, H.-G. (Eds.). (2019). *Education as a Lifelong Process: The German National Educational Panel Study (NEPS)*. Springer Fachmedien.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. Wiley series in urban research. Wiley.
- Bourdieu, P. (1973). Cultural reproduction and social reproduction. In R. Brown (Ed.), *Knowledge, education, and cultural change* (pp. 71–84). Tavistock Publications.
- Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic Inequality and Educational Outcomes: Evidence from Twenty Years of TIMSS*. Springer International Publishing.
- Cheung, C., & Pomerantz, E. (2012). Why does parents' involvement enhance children's achievement? The role of parent-oriented motivation. *Journal of Educational Psychology*, 104(3), 820–832. <https://doi.org/10.1037/a0027183>
- Diener, E., Lucas, R. E., & Oishi, S. (2018). Advances and Open Questions in the Science of Subjective Well-Being. *Collabra. Psychology*, 4(1), 15. <https://doi.org/10.1525/collabra.115>
- Ditton, H., Bayer, M., & Wohlkinger, F. (2019). Structural and motivational mechanisms of academic achievement: A mediation model of social-background effects on academic achievement. *The British Journal of Sociology*, 70(4), 1276–1296. <https://doi.org/10.1111/1468-4446.12506>
- Ditton, H., & Krüsken, J. (2006). Der Übergang von der Grundschule in die Sekundarstufe I [The transition from primary to secondary schools]. *Zeitschrift Für Erziehungswissenschaft*, 9(3), 348–372. <https://doi.org/10.1007/s11618-006-0055-7>
- Duchhardt, C., & Gerdes, A. (2013). *NEPS technical report for mathematics – Scaling results of Starting Cohort 4 in ninth grade* (NEPS Working Papers No. 22). University of Bamberg, National Educational Panel Study. https://www.neps-data.de/Portals/0/Working%20Papers/WP_XXII.pdf
- Dudas, D.-F. (2014). *Die prognostische Validität der nichtverbindlichen Gymnasialempfehlung am Ende der Grundschulzeit* [The prognostic validity of the non-binding grammar school recommendation at the end of primary school] [Doctoral dissertation, TU Dortmund University]. Eldorado Repository TU Dortmund University. <https://eldorado.tu-dortmund.de/server/api/core/bitstreams/0466ad35-8fcf-420d-9407-7ace63300e03/content>

- Dumont, H., Klinge, D., & Maaz, K. (2019). The Many (Subtle) Ways Parents Game the System: Mixed-method Evidence on the Transition into Secondary-school Tracks in Germany. *Sociology of Education*, 92(2), 199–228. <https://doi.org/10.1177/0038040719838223>
- Dumont, H., Maaz, K., Neumann, M., & Becker, M. (2014). Soziale Ungleichheiten beim Übergang von der Grundschule in die Sekundarstufe I: Theorie, Forschungsstand, Interventions- und Fördermöglichkeiten [Social disparities at the transition into secondary school: Theoretical conceptions and empirical evidence]. *Zeitschrift Für Erziehungswissenschaft*, 17(2), 141–165. <https://doi.org/10.1007/s11618-013-0466-1>
- Durik, A., Shechter, O., Noh, M., Rozek, C., & Harackiewicz, J. (2015). What if I can't? Success expectancies moderate the effects of utility value information on situational interest and performance. *Motivation and Emotion*, 39, 104–118. <https://doi.org/10.1007/s11031-014-9419-0>
- Esser, H. (2016). Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten [Ethnic inequalities in education: mechanisms, findings, debates]. In C. Diehl, C. Hunkler, & C. Kristen (Eds.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (pp. 331–396). Springer VS. <https://doi.org/10.1007/978-3-658-04322-3>
- Ganzeboom, H., & Treiman, D. J. (2010). *Occupational status measures for the new international standard classification of occupations ISCO-08; with a discussion of the new classification*. VU Amsterdam. <http://www.harryganzeboom.nl/isol/isol2010c2-ganzeboom.pdf>
- Gentrup, S., Lorenz, G., Kristen, C., & Kogan, I. (2020). Self-fulfilling prophecies in the classroom: Teacher expectations, teacher feedback and student achievement. *Learning and Instruction*, 66(5), 101296. <https://doi.org/10.1016/j.learninstruc.2019.10129>
- Gibson, A., & Koch, K. (2022). Der Übergang von der Grundschule in die weiterführenden Schulen [The transition from elementary school to secondary school]. In T. Hascher, T.-S. Idel, & W. Helsper (Eds.), *Handbuch Schulforschung* (pp. 695–714). Springer VS. https://doi.org/10.1007/978-3-658-24729-4_32
- Gomila, R. (2019). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*, 150(4), 700–709. <https://doi.org/10.1037/xge0000920>
- Gresch, C. (2012). *Der Übergang in die Sekundarstufe I: Leistungsbeurteilung, Bildungsaspiration und rechtlicher Kontext bei Kindern mit Migrationshintergrund* [The transition to lower secondary level: Performance assessment, educational aspirations and the legal context for children with a migration background.] (1st ed.). VS Verlag für Sozialwissenschaften GmbH.
- Haberkorn, K., & Pohl, S. (2013). *Cognitive basic skills-Data in the scientific use file*. University of Bamberg, National Educational Panel Study. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/com_cogbasic2013_en.pdf
- Haberkorn, K., Pohl, S., Hardt, K., & Wigand, E. (2012). *NEPS technical report for reading – Scaling results of the starting cohort 4 in ninth grade* (NEPS Working Papers No. 16). University of Bamberg, National Educational Panel Study. https://www.neps-data.de/Portals/0/Working%20Papers/WP_XVI.pdf
- Helbig, M., & Morar, T. (2017). Warum Lehrkräfte sozial ungleich bewerten. Ein Plädoyer für die Etablierung tertiärer Herkunftseffekte im werterwartungstheoretischen Standardmodell der Bildungsforschung [Why teachers evaluate socially unequal. A plea for the establishment of tertiary origin effects in the standard value expectancy theory model of educational research]. (WZB Discussion Papers 2017-005). WZB. <https://hdl.handle.net/10419/173280>
- Hoenig, K., Pollak, R., Schulz, B., & Stocké, V. (2016). Social capital, participation in adult education, and labor market success: Constructing a new instrument. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Eds.), *Methodological issues of longitudinal surveys: The example of the National Educational Panel Study* (pp. 291–312). Springer Fachmedien. https://doi.org/10.1007/978-3-658-11994-2_17

- Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts [Empirical teaching research and subject didactics. Outcome-oriented measurement and process quality of teaching]. *Zeitschrift Für Pädagogik*, 54, 222–237. <https://doi.org/10.25656/01:4348>
- Kultusministerkonferenz [KMK]. (2015). *Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe)* [Transition from elementary school to lower secondary schools and support, observation and orientation in Grades 5 and 6 (called orientation level)]. Kultusministerkonferenz.
- Ladd, G. W., Herald-Brown, S. L., & Kochel, K. P. (2009). Peers and motivation. In K. R. Wenzel & A. Wigfield (Eds.), *Educational psychology handbook series. Handbook of motivation at school* (pp. 323–348). Routledge/Taylor & Francis Group.
- Lohmann, H., & Groh-Samberg, O. (2010). Akzeptanz von Grundschulempfehlungen und Auswirkungen auf den weiteren Bildungsverlauf [Acceptance of secondary school track recommendations and their effects on educational achievement]. *Zeitschrift Für Soziologie*, 39(6), 470–492. <https://doi.org/10.1515/zfsoz-2010-0603>
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79(3), 280–295. <https://doi.org/10.1037/0022-0663.79.3.280>
- McElvany, N., Ferdinand, H. D., Gebauer, M. M., Bos, W., Huelmann, T., Köller, O., & Schöber, C. (2018). Attainment-aspiration gap in students with a migration background: The role of self-efficacy. *Learning and Individual Differences*, 65, 159–166. <https://doi.org/10.1016/j.lindif.2018.05.002>
- Meyer, J., Scharf, J., Daumiller, M., & Hübner, N. (2024). How values relate to student achievement in upper secondary education: Integrating interdisciplinary perspectives on value beliefs in the school context. *Social Psychology of Education*, 27, 2559–2588. <https://doi.org/10.1007/s11218-024-09906-2>
- National Educational Panel Study Network [NEPS Network]. (2023). *Nationales Bildungspanel, Scientific Use File der Startkohorte Klasse 5* [National educational panel study, Scientific use file of the starting cohort Grade 5]. LIfBi Leibniz Institute for Educational Trajectories. <https://doi.org/10.5157/NEPS:SC3:12.1.0>
- Neuenschwander, M. P., & Niederbacher, E. (2021). Disparitäten in Anstrengungsbereitschaft und Leistung nach SES, Familiensprache und Geschlecht: Folgen von Sozialisation oder von Diskriminierung durch verzerrte Lehrpersonenerwartungen [Disparities in effort and achievement related to SES, family language and gender: Consequences of socialisation or of discrimination through biased teacher expectancies]. *Zeitschrift für Soziologie der Sozialisation und Erziehung*, 41(4), 449–466. <https://doi.org/10.3262/ZSE2104449>
- Neugebauer, M. (2010). Bildungsungleichheit und Grundschulempfehlung beim Übergang auf das Gymnasium: Eine Dekomposition primärer und sekundärer Herkunftseffekte [Educational inequality and teacher recommendations at the transition to upper secondary school. A decomposition of primary and secondary effects of social origin.]. *Zeitschrift Für Soziologie*, 39(3), 202–214. <https://doi.org/10.1515/zfsoz-2010-0303>
- Organization for Economic Co-operation and Development (2019a). *PISA 2018 results (Volume II): Where all students can succeed*. OECD Publishing. <https://doi.org/10.1787/b5fd-1b8f-en>
- Organization for Economic Co-operation and Development /European Union (2019b). *Zusammen wachsen: Integration von Zuwanderern: Indikatoren 2018* [Indicators of Immigrant Integration: Settling in 2018], OECD Publishing. <https://doi.org/10.1787/150be71f-de>
- Passaretta, G., Skopek, J., & van Huizen, T. (2022). Is Social Inequality in School-Age Achievement Generated before or during Schooling? A European Perspective. *European Sociological Review*, 38(6), 849–865. <https://doi.org/10.1093/esr/jcac005>

- Pfost, M., Rausch, T., Schiefer, I. M., & Artelt, C. (2018). Zur Entwicklung von Gymnasias-tinnen und Gymnasiasten ohne Gymnasialempfehlung [The academic competence and self-concept of upper-track students with and without an upper-track school recommen-dation]. *Zeitschrift Für Erziehungswissenschaft*, 21(3), 511–534. <https://doi.org/10.1007/s11618-017-0787-6>
- Roth, B., Becker, N., Romeyke, S., Schäfer, S., Krieger, F., & Spinath, F. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, 53, 118–137. <https://doi.org/10.1016/j.intell.2015.09.002>
- Roth, T., & Siegert, M. (2015). Freiheit versus Gleichheit [Freedom versus equality]? *Zeitschrift für Soziologie*, 44(2), 118–136. <https://doi.org/10.1515/zfsoz-2015-0205>
- Schiefele, U., Moschner, B., & Husstegge, R. (2002). *Skalenhandbuch SMILE-Projekt* [Scale manual SMILE project, Unpublished manuscript]. Bielefeld University. Department of Psychology.
- Saklofske, D. H., Austin, E. J., Mastoras, S. M., Beaton, L., & Osborne, S. E. (2012). Relationships of personality, affect, emotional intelligence and coping with student stress and academic success: Different patterns of association for stress and success. *Learning and Individual Differences*, 22(2), 251–257. <https://doi.org/10.1016/j.lindif.2011.02.010>
- Schindler, S. (2018). Soziale Ungleichheit im Bildungsverlauf [Social inequality in the educa-tional process]. In N. McElvany, W. Bos, H. G. Holtappels, J. Hasselhorn, & A. Ohle-Peters (Eds.), *Dortmunder Symposium der empirischen Bildungsforschung: Band 3. Bedingun-gen erfolgreicher Bildungsverläufe in gesellschaftlicher Heterogenität: Interdisziplinäre Forschungsbefunde und Perspektiven für Theorie und Praxis*. Waxmann.
- Shi, D., Maydeu-Olivares, A., & Rosseel, Y. (2020). Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(1), 1–15. <https://doi.org/10.1080/10705511.2019.1611434>
- Šimunović, M., & Babarović, T. (2020). The role of parents' beliefs in students' motivation, achievement, and choices in the STEM domain: a review and directions for future research. *Social Psychology of Education*, 23(3), 701–719. <https://doi.org/10.1007/s11218-020-09555-1>
- Spera, C., Wentzel, K. R., & Matto, H. C. (2009). Parental aspirations for their children's educational attainment: Relations to ethnicity, parental education, children's academic performance, and parental perceptions of school climate. *Journal of Youth and Adolescence*, 38(8), 1140–1152. <https://doi.org/10.1007/s10964-008-9314-7>
- StataCorp LLC. (2021). *Stata Statistical Software: Release 17* [Computer software].
- Steinmayr, R., Dinger, F. C., & Spinath, B. (2012). Motivation as a mediator of social disparities in academic achievement. *European Journal of Personality*, 26(3), 335–349. <https://doi.org/10.1002/per.842>
- Statistical Federal Office [Destatis] (2023). Genesis-Online, Tabelle 21111-0001.
- Stephens, N. M., Brannon, T. N., Markus, H. R., & Nelson, J. E. (2015). Feeling at home in col-lege: Fortifying school-relevant selves to reduce social class disparities in higher education. *Social Issues and Policy Review*, 9(1), 1–24. <https://doi.org/10.1111/sipr.12008>
- Stocké, V. (2005). *Einstellung zu Bildung: Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)* [Attitudes towards education: Compilation of social science items and scales (ZIS)]. ZIS – GESIS Leibniz Institute for the Social Sciences. <https://doi.org/10.6102/ZIS136>
- Stubbe, T., Kleinkorres, R., Krieg, M., Schaufelberger, R. & Schlitter, T. (2023). Soziale und migrationsbedingte Disparitäten in der Lesekompetenz von Viertklässerinnen und Viert-klässlern [Social and migration-related disparities in the reading skills of fourth graders]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. Stubbe, (Eds.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 151–177). Waxmann.

- Tiedemann, J., & Billmann-Mahecha, E. (2010). Wie erfolgreich sind Gymnasiasten ohne Gymnasialempfehlung? Die Kluft zwischen Schullaufbahnempfehlung und Schulförmwahl der Eltern [How successful are pupils in the german gymnasium without a recommendation for this school-type]? *Zeitschrift für Erziehungswissenschaft*, 13(4), 649–660. <https://doi.org/10.1007/s11618-010-0146-3>
- United Nations Educational Scientific and Cultural Organization. (1997). *International Standard Classification of Education (ISCED 1997)*. United Nations educational, scientific and cultural Organization.
- Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review*, 32, 100374. <https://doi.org/10.1016/j.edurev.2020.100374>
- Valtin, R. (2020). *Eine Zwei ist eine Drei ist eine Vier. Über die mangelnde Vergleichbarkeit von Zensuren und die Problematik der Grundschul-Empfehlungen* [A two is a three is a four. On the lack of comparability of grades and the problem of primary school recommendations]. <https://doi.org/10.25656/01:19226>
- Verešová, M., & Mala, D. (2016). Attitude toward School and Learning and Academic Achievement of Adolescents. In Z. Bekirogullari, M. Y. Minas, & R. X. Thambusamy (Eds.), *European Proceedings of Social and Behavioural Sciences: Vol. 16, Iceepsy 2016: Education and Educational Psychology: Proceedings of the 7th International Conference on Education and Educational Psychology Conference (ICEEPSY 2016), 11–15 October 2016, Rhodes, Greece* (pp. 870–876). Future Academy. <https://doi.org/10.15405/epsbs.2016.11.90>
- Wang, L., & Finch, H. (2018). Motivation variables mediate the relationship between socioeconomic status and academic achievement. *Psychology and Education: An Interdisciplinary Journal*, 55, 123–136.
- Wiederkehr, V., Darnon, C., Sebastien, C., Guimond, S., & Martinot, D. (2015). From social class to self-efficacy: internalization of low social status pupils' school performance. *Social Psychology of Education*, 18, 769–784. <https://doi.org/10.1007/s11218-015-9308-8>
- Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wohlkinger, F. (2014). *Die Rolle des Schölers bei der Wahl der weiterführenden Schule: Eine vergleichende Untersuchung von Grundschölern aus Bayern und Sachsen* [The role of the student in the choice of secondary school: A comparative study of primary school students from bavaria and saxony]. Springer VS. <https://doi.org/10.1007/978-3-658-06001-5>
- Zell, E., & Lesick, T. L. (2021). Taking Social Comparison to the Extremes: The Huge-Fish-Tiny-Pond Effect in Self-Evaluations. *Social Psychological and Personality Science*, 12(6), 1030–1038. <https://doi.org/10.1177/1948550620956535>

Supplementary Material

Table A1: Intercorrelations Among all Analyzed Variables in our Sample for Students With (Upper Right) and Without (Lower Left) ATER

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Abitur		.21	.20	.05	-.05	.34	-.12	.12	.04	.00	-.08	.07	.42	.24	.21	.22
2 PE	.21		.61	.07	-.05	.20	-.07	.03	-.06	-.19	-.00	-.00	.19	.11	.20	.18
3 OS	.20	.58		.07	-.05	.20	-.07	.03	-.06	-.10	-.01	-.01	.20	.11	.21	.19
4 CosFin	.12	.09	.11		.05	.15	.03	.09	.08	.04	-.01	.05	.07	.01	.01	.01
5 CosLea	.07	.04	.03	.20		-.21	-.10	-.02	-.05	.01	.07	.03	-.03	.03	-.12	-.07
6 EOS	.28	.12	.13	.16	.02		.22	.14	.19	-.02	-.07	.00	.25	.10	.23	.21
7 ParAsp	.29	.10	.14	.12	.02	.36		.17	.11	.07	-.04	.02	.27	.07	.09	.11
8 VOE	.08	.02	.03	.02	-.02	.23	.22		.21	.11	-.06	-.13	.15	-.00	.02	.01
9 PM	.03	-.01	-.03	.11	-.02	.20	.11	.23		.04	.02	-.09	-.03	-.05	.01	-.01
10 HL	-.00	-.23	-.24	.05	-.01	.12	.21	.17	.11		-.06	.03	.00	-.04	-.12	-.09
11 BindER	-.05	.05	.01	-.07	.05	-.06	-.07	-.06	-.03	-.01		-.07	.03	-.02	.08	.08
12 Gender	.09	-.06	-.04	.02	.01	-.01	.07	-.08	-.11	.06	-.03		-.01	.01	-.21	.09
13 STYPE	.27	.15	.17	.07	.03	.27	.28	.13	-.05	.04	-.07	-.05		.17	.23	.19
14 CogAb	.15	.13	.08	.05	.11	.14	.06	.00	-.14	-.10	.06	-.04	.17		.35	.27
15 MathC	.19	.19	.17	.02	.13	.20	.08	.03	-.10	-.17	.07	-.22	.23	.42		.45
16 ReadC	.23	.12	.13	.06	.06	.19	.13	-.01	-.19	-.08	.06	.07	.20	.31	.48	

Note. ATER = academic-track enrollment recommendation; PE = parental educational level; OS = parental occupational status; CosFin = financial cost; CosLea = cost of learning stress; EOS = expectation of success; ParAsp = parental aspirations; VOE = value of education; PM = performance-related motivation; HL = home language; BindER = bindingness of the enrollment recommendation; STYPE = school type; CogAb = cognitive abilities; MathC = math competences; ReadC = reading competences. Significant correlations ($p < .05$) are in bold.

Table A2: Mean Comparison of all Analyzed Variables, Standard Deviations (SD) and *n* of Respondents by Enrollment Recommendation and School-leaving Certificate

	Group 1		Group 2		$\Delta M p$	Group 3		$\Delta M p$	Group 4		$\Delta M p$
	<i>M</i> (SD)	<i>n</i>	<i>M</i> (SD)	<i>n</i>		<i>M</i> (SD)	<i>n</i>		<i>M</i> (SD)	<i>n</i>	
PE	6.85 (2.07)	302	7.72 (2.07)	1,502	<.001	5.64 (2.40)	798	<.001	6.54 (2.34)	320	.054
OS	56.99 (17.80)	298	66.35 (16.65)	1,492	<.001	48.58 (18.76)	789	<.001	56.97 (18.92)	316	.494
CosFin	4.23 (0.91)	251	4.11 (0.94)	1,278	.262	3.94 (1.14)	593	.003	3.97 (1.01)	251	.003
CosLea	4.66 (0.57)	267	4.47 (0.69)	1,376	<.001	4.53 (0.91)	622	.112	4.58 (0.67)	265	.732
EOS	3.26 (0.93)	269	4.06 (0.71)	1,379	<.001	2.62 (1.04)	619	<.001	3.31 (0.99)	263	.717
ParAsp	0.54 (0.50)	251	0.77 (0.42)	1,326	<.001	0.24 (0.43)	611	<.001	0.36 (0.48)	238	<.001
VOE	3.02 (0.86)	271	3.25 (0.84)	1,415	.010	2.86 (0.89)	676	.004	2.95 (0.90)	273	.304
PM	2.57 (0.86)	272	2.44 (0.64)	1,390	.202	2.52 (0.62)	643	.972	2.37 (0.61)	272	.173
HL	0.17 (0.38)	298	0.12 (0.32)	1,490	.024	0.18 (0.38)	785	.244	0.12 (0.32)	314	.002
BindER	0.34 (0.48)	267	0.49 (0.50)	1,376	<.001	0.40 (0.49)	798	.563	0.59 (0.49)	320	<.001
Gender	0.53 (0.50)	301	0.52 (0.50)	1,494	.670	1.44 (0.50)	790	.320	0.42 (0.49)	316	.020
CogAb	9.15 (2.13)	209	10.23 (1.56)	1,129	<.001	8.40 (2.34)	601	<.001	9.01 (2.36)	263	.530
MathC	0.72 (1.07)	256	1.66 (1.03)	1,389	<.001	0.27 (1.04)	631	<.001	1.05 (0.96)	256	.010
ReadC	0.78 (1.21)	255	1.66 (1.23)	1,388	<.001	0.18 (1.13)	630	<.001	0.89 (1.08)	257	.587

Note. Group 1: Students obtaining Abitur without ATER. Group 2: Students obtaining Abitur with ATER. Group 3: Students not obtaining Abitur without ATER. Group 4: students not obtaining Abitur with ATER. ATER = academic-track enrollment recommendation; *PE* = parental educational level; *OS* = parental occupational status; *CosFin* = financial cost; *CosLea* = Cost of learning stress; *EOS* = expectation of success; *ParASP* = parental aspirations; *VOE* = value of education; *PM* = performance-related motivation; *HL* = home language; *BindER* = bindingness of the enrollment recommendation; *CogAb* = cognitive abilities; *MathC* = math competences; *ReadC* = reading competences.

Addressing the Heterogeneity of Students' Attitudes Towards School: Uncovering School Alienation Profiles in German Secondary Schools

Alyssa Grecu, Sira Neuhaus, Sittipan Yotyodying & Nele McElvany

Zusammenfassung

Die Einstellung der Schüler*innen zur Schule wandelt sich im Laufe der Sekundarstufe erheblich. Veränderungen während der Adoleszenz sowie die sich im Bildungsverlauf verändernden Lernumgebungen prägen die schulischen Erfahrungen der Schüler*innen und können die Entfremdung vom Lernen, von Lehrpersonen und Mitschüler*innen verstärken. Je nach betroffenem Bereich führt Schulentfremdung zu unterschiedlichen problematischen sozialen und lernbezogenen Verhaltensweisen in und außerhalb der Schule. Dieser Beitrag befasst sich mit der Heterogenität der Einstellungen von Schüler*innen zur Schule und zielt darauf ab, empirische Evidenz für die bereichsspezifische Ausprägung von Schulentfremdung und ihren Zusammenhang mit individuellen Schülermerkmalen zu liefern. Querschnittsdaten von 223 Sekundarschüler*innen (Klassen 7 und 10) aus zwei deutschen Bundesländern wurden für latente Profilanalysen mit anschließender Korrelationsanalyse verwendet. Die Ergebnisse zeigten unterschiedliche Schulentfremdungsprofile für die Klassenstufen 7 und 10. Jungen, Schüler*innen mit Migrationshintergrund, Schüler*innen, die eine Haupt- oder Realschule besuchen, und Schüler*innen, die bereits eine Klasse wiederholt hatten, wiesen höhere Entfremdungsprofile auf. Die Implikationen der vorliegenden Studie werden diskutiert.

Schlagnworte: Schulentfremdung, Latente Profilanalyse, Sekundarschule, Deutschland, Adoleszenz

Abstract

Students' attitudes towards school change significantly, as they progress through secondary school. Considerable changes during adolescence as well as changing learning environments shape students' experiences at school and may increase students' alienation from learning, teachers, and classmates. Depending on the domain affected, school alienation leads to different problematic social and learning-related behaviors in and out of school. This paper addresses the heterogeneity of students' attitudes towards school and aims to provide empirical evidence on the domain-specific manifestation of school alienation and its relation with individual students' characteristics. Cross-sectional data comprising 223 secondary school students (Grades 7 and 10) from two German federal states were used for latent profile analysis followed by a correlational analysis. Results revealed distinct school alienation profiles for Grade 7 and Grade 10. Boys, students with immigrant background, students attending Hauptschule or Realschule, and students who had previously repeated a grade were found to show higher alienation profiles. Implications of the present study are discussed.

Keywords: School alienation, latent profile analysis, secondary school, Germany, Adolescence

1. Introduction

In recent years, there has been a surge of academic interest in the importance of socio-emotional characteristics for educational achievement. Scholars widely acknowledge that students' attitudes and emotions in school decline over the course of schooling (Archambault et al., 2009; Çağlar, 2013; Hascher & Hagenauer, 2010) hindering engagement in classroom learning-related activities and affecting educational outcomes (Becker, 2024). It must be noted that past research often focuses on positive constructs such as well-being, motivation or enjoyment in the school context. However, to adequately investigate the underlying processes of detachment from learning and school in general, concepts are needed that specifically capture the increase of negative attitudes and emotions. This is precisely what school alienation research addresses. Alienation can be described as the state or the experience of being isolated from a group or an activity to which one should belong or in which one should be involved (Mann, 2001). Applied to the school context, alienated students feel powerless to influence events, cannot attach importance to what

goes on in school and are estranged from the school community and its values (Dyran, 1980). This is particularly critical as students cannot avoid school and associated expectations due to compulsory education.

As students' attitudes towards school decrease, alienation may manifest in different ways, depending on its severity and the students' personalities. Some students might express their alienation more outwardly, while others might internalize their feelings. Accordingly, alienated students show disengagement from learning and deviant behavior, potentially resulting in problematic educational trajectories, up to and including dropping out of school (Avci & Celikkaleli, 2016; Brown et al., 2003; Studsrød & Bru, 2012; Sutherland, 2011).

Up to now, research on school alienation has taken up an international perspective (Hascher & Hadjar, 2018), while the German context has been under-explored. School alienation is a negative development at the individual level in the school environment over time and may be triggered by the institutional setting as well as by negative social experiences (Hascher & Hadjar, 2018). As characteristics of schooling and teaching differ between education systems, findings obtained in other education systems can only be transferred to a limited extent. Taking this as starting point, this article examines the domain-specific manifestation of school alienation and its relation with individual student characteristics in the German context by shedding light on the heterogeneity of students' attitudes towards school.

2. Theoretical Background

Since the mid-20th century, scholarly focus turned towards the phenomenon of school alienation (Bronfenbrenner, 1986; Brown et al., 2003; Finn, 1989; Mann, 2001; Mau, 1992; Seeman, 1959). Until recently, there was no precise definition of the essence of school alienation. The issue was thoroughly examined in a comprehensive review by Hascher and Hadjar (2018), who conceptualized school alienation by synthesizing existing theoretical approaches and empirical findings. According to them, school alienation is defined as a process by which students develop negative attitudes, including cognitive and affective components, and distance themselves from domains related to schooling. These negative attitudes might change over time and can solidify into a disposition. School alienation consists of cognitive and affective elements. Cognitive aspects refer to students' beliefs, knowledge, assumptions, and judgements with respect to the school environment. Affective aspects comprise students' emotions and feelings resulting from their experiences in school as well as whether they perceive the educational offer as useful. This process involves a decreasing

sense of belonging in terms of three school-related domains: (a) learning as the main objective of schooling, (b) teachers as educators and mentors, (c) classmates as the social peer community. The extent of school alienation can vary across one to all three affected domains, allowing for the capture of different degrees of alienation within each domain. The presented approach focuses on the three domains mentioned above, to which students' negative feelings and thoughts are related. This concept allows to distinguish between psychological processes in the development of alienation and enacted outcomes (Hascher & Hadjar, 2018). The latter includes behaviors that are considered to be expressions or consequences of school alienation such as disengagement or dropout, which are located outside the concept. Within educational research, several theoretical approaches exist to explain the development of school alienation. From a developmental perspective, stage-environment-fit theory (Eccles et al., 1993) posits that changes in school alienation can be explained by the degree of alignment or mismatch between students' developmental needs and the educational environment. This alignment influences their attitudes at different developmental stages and different types of educational environment are beneficial for different age groups.

As students' progress through secondary school, the gap between their needs and the school environment widens. They mature and their emotional, cognitive and social needs and personal goals evolve. Their needs for autonomy, relatedness, and competence also increase. Specifically, positive attitudes and motivation towards school and learning tend to decrease, while motivational conflicts between leisure activities and school tasks increase (Hascher & Hagenauer, 2010; Hofer, 2004). Elements known to mitigate school alienation by enhancing stage-environment-fit in secondary school are autonomy support, involvement, and structure (Eccles et al., 1993; Roeser et al., 2000). At the same time, school demands evolve including relationships to teachers and peers change (Eccles et al., 1993).

This change in the school environment interacts with the beginning adolescence, which is characterized by significant changes in students' attitudes, a critical factor that has been linked to increased school alienation (Calabrese & Seldin, 1986; Calabrese, 1987). Moreover, early adolescence presents students with the developmental challenge of balancing their desire for integration into one or more social groups, which can sometimes result in rejection or isolation (Newman & Newman, 2003). Given this context, some researchers view alienation as an inevitable and natural stage in adolescent development (Nutbeam et al., 1993; Sidorkin, 2004). Nevertheless, alienation from school may begin in the early grades and manifests over time as a result of processes of frustration and withdrawal (Steuer et al., 2024).

Against this background, secondary school students seem to be particularly vulnerable for experiencing mismatch regarding the fit of their psychological needs and school's demands. This in turn, may lead to a decrease of positive emotions and attitudes toward school resulting in school alienation (Schmid et al., 2021).

2.1. State of Research

Researchers agree that students' positive attitudes and emotions towards school and learning such as well-being decrease over the course of schooling (e.g., Govorova et al., 2020; Hascher & Hagenauer, 2020; Raccanello et al., 2019; Symonds & Galton, 2014). Even if there is evidence that negative attitudes towards school begin to develop in early years of schooling (Finn, 1989; Hascher & Hagenauer, 2010), the transition to secondary school is a particularly critical time when students lose their enjoyment and interest in learning (Archambault et al., 2009; Wigfield et al., 2006). In line with this, studies from several educational contexts show that school alienation increases during the course of secondary schooling (Calabrese & Seldin, 1986; Calabrese, 1987; Grecu, 2020; Scharf, 2021). Considering the domain-specific nature of school alienation, this increase is stronger in the two academic domains alienation from learning and alienation from teachers (Grecu, 2020; Scharf, 2021), which emphasizes domain-specific developmental trends. Moreover, the time point when school alienation increases, varies between education systems (Hadjar et al., 2021; Scharf, 2021). While there is a general increase in school alienation during secondary school, this is not universal. In fact, some students experience a decrease in this phenomenon at a later point in their secondary education (Hadjar et al., 2021).

The development of school alienation needs to be discussed from a multi-causal perspective, as it typically involves a variety of factors (Hascher & Hadjar, 2018). Therefore, the causes of school alienation are manifold and relate to students' individual characteristics, their socialization contexts and how they experience the learning environment and its agents.

Factors on the Individual Student Level

Regarding the question of which groups of students are particularly vulnerable to school alienation, the importance of gender, immigrant background, and social background is relevant. Referring to social capital theory (Bourdieu, 1983), this is commonly explained by varying opportunities to meet school's

requirements and culture. In this sense, students may develop alienation when they lack academic, social or cultural capital and cannot meet schools' expectations concerning behavior, social skills, and achievement (Grecu, 2020).

These studies consistently indicate significant gender differences, with boys experiencing higher levels of alienation than girls (Grecu, 2020; Hadjar & Lupatsch, 2010; Hadjar et al., 2015; Leduc & Bouffard, 2017; Morinaj et al., 2020). Stage-environment fit theory (Eccles & Midgley, 1989) offer one explanation for this, suggesting that schools are better at meeting the needs of girls than those of boys. As a result, boys find it more difficult to adapt to schools' expectations (Hadjar et al., 2015). In this sense, school alienation can be understood as boys' response to unmet needs at school at a behavioral level (Hadjar et al., 2015). Moreover, peer attitudes seem to be an important factor that causes boys to show deviant behavior in the classroom (Hadjar et al., 2015).

When considering immigrant background, the findings vary depending on the domain of alienation. For instance, Grecu (2020) found no differences in alienation from learning based on immigrant background. However, students with an immigrant background tend to perceive higher levels of alienation from teachers compared to their peers without an immigrant background (Hadjar et al., 2021; Morinaj et al., 2020; Scharf, 2021).

Students from low socioeconomic status (SES) backgrounds have been identified as the third risk group for school alienation, experiencing higher levels of alienation compared to their peers from high SES backgrounds (Hadjar & Lupatsch, 2010; Hadjar et al., 2015; Hascher & Hagenauer, 2010; Leduc & Bouffard, 2017). Furthermore, previous studies indicate that primary school students experience less alienation compared to those in secondary school (Çağlar, 2013; Hascher & Hagenauer, 2010). In stratified education systems, school alienation tends to intensify after the transition to secondary school, which is rooted in several challenges to adapt to the new school environment and its requirements (Grecu et al., 2019).

A recent study adopted a similar approach by examining the alienation profiles of secondary school students, focusing on the core characteristics of the multi-domain concept of school alienation within the context of Switzerland (Canton Berne) and Luxembourg (Hadjar et al., 2021). Results show country-specific alienation profiles that change partly over time. Overall, findings emphasize the necessity to recognize the context-sensitive characteristics of the school alienation concept resulting in different alienation profiles depending on country and grade. The alienation profiles between countries differ in regard to the extent of alienation expressing the varying strength of alienation. Moreover, school alienation profiles are differently stable over

time in the countries under study. While the Swiss alienation profiles are rather stable, Luxembourgish alienation profiles indicate more transitions. For example, 96% of the students who were marginally alienated in year 7 transitioned to a profile of moderate alienation from learning by year 9. In both countries, the most intense alienation profile contains a low share of students, which increases over time. By Grade 9, 26% of students in the Swiss sample belong to the moderate alienation profile, while in the Luxembourgish sample 22% of students show considerable alienation from learning, teachers, and classmates. Overall, this study validates the conceptualization of school alienation as a multidimensional construct (Hascher & Hadjar, 2018; Morinaj et al., 2017). It highlights the necessity of distinguishing between different domains, as the extent of alienation varies across the three domains.

Factors of the Social Environment

The role of socialization agents depends on school alienation domain and educational context. In Switzerland as well as in Luxembourg, peers appeared to have a strong impact on all three domains of alienation (Morinaj et al., 2023; Schmid et al., 2021). Positive attitudes towards school among peers as well as peer support were found to reduce alienation depending on the several domains (Morinaj et al., 2023). On the other hand, peers' negative attitudes towards school such as low academic aspirations or oppositional orientations affect the development of school alienation (Barber et al., 2005; Hadjar & Lupatsch, 2010; Hascher & Hagenauer, 2010; Murdock, 1999; Organization for Economic Co-operation and Development [OECD], 2003).

Support by parents has the potential to prevent school alienation in the social domains. In the Swiss sample, the results revealed that supportive attitudes of parents related to school in Grade 7 are positively related with alienation from teachers and classmates in Grade 8 (Morinaj et al., 2023). Parental attachment also influences students' adjustment to school, which in turn affects the development of school alienation (Kocayörük et al., 2014; Kocayörük & Şimşek, 2016). On the downside, students may adopt their parents' negative attitudes toward school, leading to increased conflicts with learning, teachers, and classmates. This ultimately heightens students' feelings of negativity and alienation (Hascher & Hadjar, 2018; Kocayörük & Şimşek, 2016).

Factors on the School Level

Previous studies confirm the crucial role of teachers in the classroom emphasizing the importance of teaching styles, teacher support, teacher justice, and the overall teacher-student-relationship in the context of the development of school alienation. The classroom setting is substantially shaped by the teachers and determines how students experience learning and spending time in school including both an academic as well as a social dimension (Hagenauer & Volet, 2014; Hascher & Hadjar, 2018; Pianta & Hamre, 2009; Wang & Fredericks, 2014). Not only the teacher-student-relationship (Çağlar, 2013; Hascher & Hagenauer, 2010; Legault et al., 2006; Schmid et al., 2021; Studsrød & Bru, 2012) but also teaching styles (Grecu et al., 2022) were found to affect students' alienation from teachers and from learning. An authoritative teaching style has a positive impact on students' attitudes, their integration in school as well as on their educational success (Baker et al., 2009; Myers & Pianta, 2008). However, an unfair teaching style fosters the development of alienation in the domains learning and teachers (Grecu et al., 2022; Marcin et al., 2020). Of particular importance is autonomy support by teachers, which also reduces alienation from learning (Hadjar et al., 2015; Hadjar et al., 2021; Morinaj et al., 2023) and increases student engagement (Patall et al., 2018). In line with that, negative error climate fosters alienation from teachers in primary school (Steuer et al., 2024).

Also school-specific structural characteristics, such as school rules, decision-making processes, and opportunities of student participation determine how students experience the school environment. These aspects profoundly influence the extent to which students feel empowered and involved in school life (Çağlar, 2013; Hascher & Hagenauer, 2010; Sutherland, 2011; Mahmoudi et al., 2015). Further aspects located at the macro and meso levels affecting school alienation relate to the country-specific structures and conditions in the education system (Hadjar & Gross, 2016) and the class composition. On the system level, the extent of stratification and the class composition in secondary school shape the learning environments, which affects students' experiences in school. However, findings are mixed. For example, in Luxembourg, a high share of students with low SES leads to lower alienation from teachers, while it increases in Switzerland. In the same vein, low levels of values of education are associated with higher levels of alienation (Scharf, 2021).

Research on school alienation in Switzerland and Luxembourg provides insights how school alienation varies by school type and level of stratification. The increase in school alienation during secondary school is a common development (Grecu, 2020). However, the extent of school alienation varies by school type. In the Luxemburgish education system, alienation from

learning is stronger in the highest academic school track than in the lower school tracks. Contrary to expectations, the lowest school track shows the least pronounced extent of alienation from learning (Grecu, 2020; Scharf, 2021). In Switzerland, the evidence shows that alienation from learning is higher in ability-homogenous classes than in classes that are organised in a more inclusive framework that allows for more heterogeneous class composition (Scharf, 2021).

The stronger alienation in Luxembourg's highly stratified education system of Luxembourg can be attributed to higher pressure to perform and more frequent school transitions compared to less segregated and more permeable education systems (Hadjar et al., 2021). Despite expectations from stage-environment fit theory that homogeneous learning environments should better meet students' needs, the intense pressure in the highest educational track, often reinforced by parental expectations, may lead to negative attitudes to school (Grecu, 2020; Scharf, 2021). In contrast, less segregated systems, with joint teaching and more diverse peer interaction, tend to have lower performance pressure (Scharf, 2021).

Consequences of School Alienation

What truly renders school alienation a problem are its manifold consequences for affected students. Recent empirical studies applying a longitudinal perspective revealed that school alienation leads to lower well-being (Dekel & Tuval-Mashiach, 2012; Morinaj et al., 2019; Morinaj & Hascher, 2019; Rayce et al., 2008) as well as to different learning and social behaviors (Morinaj et al., 2019). The latter includes disengagement from learning, deviant behavior and problematic educational trajectories and school drop-out (Archambault et al., 2009; Avci & Çelikkaleli, 2016; Brown et al., 2003; Studsrød & Bru, 2012; Sutherland, 2011).

Recent studies showed that each of the three domains of alienation is differently associated with various dimensions of well-being as well as with specific learning and social behaviors. While alienation from learning was found to have a negative impact on positive attitudes to school and school enjoyment, alienation from teachers as well as alienation from classmates have a negative impact on positive attitudes to school and contribute to social problems in school. Alienation from teachers negatively predicted positive attitudes to school and positively predicted worries and social problems in school (Morinaj & Hascher, 2019). Concerning the behavioral level, alienation from learning and from teachers is associated with lower participation in class.

Furthermore, alienation from teachers and from classmates contributes to subsequent delinquent behaviour (Morinaj et al., 2019).

Moreover, school alienation is reciprocally related with academic achievement. A longitudinal study focussing on secondary school students showed that higher levels of alienation from learning are associated with lower academic achievement expressed by a decrease of GPA at the subsequent grade (Morinaj et al., 2020). Again, domain-specific associations were found for each country under study. Additionally, at some time points of inquiry lower achievement was also found to predict alienation in specific domains (Morinaj et al., 2020).

3. Aims of the Study

This study aims to validate previous research on alienation patterns among secondary school students within the German context, building upon pioneering insights from Switzerland and Luxembourg. Given the limited knowledge about school alienation in the German school context and its relation to students' characteristics and their experiences at school, the present study aims to provide empirical evidence on the domain-specific manifestation of school alienation and its relation to individual students' characteristics.

This paper addresses the heterogeneity of students' attitudes towards school and raises the following research questions:

- a) which alienation profiles can be found among seventh and tenth graders of German secondary schools;
- b) how are the alienation profiles associated with individual characteristics, namely gender, immigrant background, school track, and grade retention?

4. Methods

4.1. Data Set and Sample Description

This study was used data from the Study on the Development of Political and Social Competence in Adolescence (EPKO: Alscher et al., 2023). Data were collected via an in-class questionnaire survey during winter 2020/2021 at German secondary schools. Using a convenience sampling approach, 152 seventh graders ($M = 12.48$ years old, $SD = 0.70$; 54.7% male; 44.6% with an immigrant background) and 98 tenth graders ($M = 16.12$ years old; $SD = 0.96$; 54.3% male; 46.3% with an immigrant background) from two federal states North

Rhine-Westphalia (65%) and Thuringia (35%), participated. The latter can be partly explained by a comparatively high share of students with an immigrant background in North Rhine-Westphalia (40% in 2021/22), which is also reflected in the sample and combines first and second generation immigrant background (Landesbetrieb Information und Technik Nordrhein-Westfalen, 2022).

In the sample of seventh graders, 56.6% of seventh graders attended schools with a Gymnasium track, while 43.4% were enrolled in other types of schools. In the sample of tenth graders, only 21.4% of tenth graders attended schools with a Gymnasium track, compared to 77.6% who were enrolled in other types of schools.

4.2. Measures

In this article, we used established scales to assess study variables. First of all, measures for school alienation will be described followed by measures for individual characteristics.

Measures for School Alienation

To assess the multidimensional aspects of school alienation, this study utilized four scales adapted from the School Alienation Scale (Morinaj et al., 2017). These scales were designed to capture varying facets of alienation within the school environment, including alienation from classmates, teachers, and learning (both affective and usefulness dimensions). Each scale consisted of multiple items, and participants responded to these items using a Likert scale ranging from 1 = *completely disagree* to 4 = *completely agree*. The specific number of items, an example item, and internal consistency for each scale are detailed in Table 1.

Table 1: Scales to Assess Multidimensional School Alienation

Construct	N _{item}	Example items	M/(SD)		α
			Grade 7	Grade 10	
Alienation from Classmates	4	In my class, I feel like someone who doesn't fit in.	1.94 (0.68)	1.88 (0.75)	.72
Alienation from Teachers	3	I can't trust my teachers.	1.86 (0.69)	1.86 (0.61)	.73
Alienation from Learning (affective)	3	The things we learn in school are boring.	2.51 (0.73)	2.64 (0.72)	.83
Alienation from Learning (usefulness)	3	I find it pointless what we have to learn in school.	1.80 (0.72)	2.21 (0.74)	.80

4.3. Measures for Individual Characteristics

In addition to the school alienation measures, the study incorporated various measures to assess students' individual characteristics. These measures were derived from established scales used within the EPKO study (Alscher et al., 2023). These additional measures encompassed a range of factors for student characteristics, including gender, immigrant background, school track, and grade retention. Table 2 provides detailed information on the construction, example items, scale range, and other relevant details for each of these measures.

Table 2: Scales to Assess Student Characteristics

Construct	N_{item}	Construction, example items	Characteristic value
Gender		Dichotomous variable: female or male	0 = <i>male</i> 1 = <i>female</i>
Immigrant background	1	Category based on information on the students' countries of birth and their parents, immigrant background: either first generation, second generation or 2.5 generation immigrant (i.e., students with one parent born abroad)	0 = <i>non-immigrants</i> 1 = <i>immigrants</i>
School track: Gymnasium or other	1	What type of school do you attend?	0 = other school type 1 = Gymnasium and other school types with a Gymnasium track
Grade retention	1	Have you ever been left behind or repeated a class?	0 = no grade retention 1 = grade retention

4.4. Analytical Procedure

As alienated students might be affected in different interrelated domains, experiencing alienation differently, a person-oriented approach (Bergman & Magnusson, 1997) is applied to analyze alienation profiles. The analytical procedure involves the following steps. First of all, a latent profile analysis (LPA) with Mplus 8.8 was carried out in order to examine school alienation profiles, thereby showing which domains of school alienation were affected. The goal was to see if there were different patterns in how students experience alienation. Several model fit indices were used to determine the best fitting model (e.g., Geiser, 2013; Jung & Wickrama, 2008): Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Adjusted Bayesian Information Criterion (aBIC), Consistent Akaike Information Criterion (CAIC), and entropy. Lower values for AIC, BIC, adjusted BIC, and CAIC indicate a better fit, while higher values for entropy ($> .80$) indicate a better fit. To compare the relative fit of different LPA models, we considered the Local Model (LM) and Adjusted Local Model (ALM) information criteria. Lower values of these criteria indicate a better model fit. Additionally, Likelihood

Ratio Test (LM-LRT), Lo-Mendall-Rubin Adjusted LRT Test (ALM-LRT), and Parametric Bootstrapped Likelihood Ratio Test (BLRT) were used to assess if increasing model complexity (more profiles) leads to a significantly better fit. A non-significant ($p > .05$) result from the LM-LRT, ALM-LRT, or BLRT indicates that the $k + 1$ -profile model is not significantly better than the k -profile model, so the k -profile model should be retained. In the second step, we conducted a correlational analysis, utilizing a cross-tabulation, to examine how the alienation profiles were related to individual characteristics (i.e., gender, immigrant background, school track, and grade retention).

5. Results

5.1. Latent Profile Analysis

In the present study, we performed a latent profile analysis (LPA) to identify distinct groups or profiles of students based on their levels of school alienation. Based on the results in Table 3, the best fitting model for seventh graders was a 4-profile model, while the best fitting model for tenth graders was a 3-profile model. This suggests that there are four distinct profiles of school alienation among seventh graders and three distinct profiles among tenth graders. The entropy values for the best fitting models were high (Grade 7: 0.81; Grade 10: 0.85), indicating that the profiles are well-separated and that the model accurately classified students into their respective profiles.

Table 3: Information Criteria and Entropies for LPA Models With Different Numbers of Profiles

Grade	N of Profile	AIC	BIC	aBIC	Entropy	LM-LRT	ALM-LRT	BLRT
7	1	1.256.012	1.279.936	1.254.620				
	2	1.213.748	1.252.623	1.211.484	0.671	0.1320	0.1408	0.0000
	3	1.194.991	1.248.819	1.191.858	0.880	0.1335	0.1417	0.0000
	4	1.177.565	1.246.345	1.173.561	0.900	0.1790	0.1902	0.0050
	5	1.170.378	1.254.111	1.165.503	0.850	0.6534	0.6645	0.0400
10	1	839.500	860.097	834.836				
	2	785.088	818.559	777.510	0.744	0.0201	0.0232	0.0000
	3	780.044	826.388	769.551	0.857	0.3345	0.3432	0.0650
	4	776.887	836.105	763.479	0.832	0.3127	0.3233	0.2000
	5	775.131	847.223	758.809	0.865	0.0867	0.0925	0.3500

Note. Best-fit model in bold.

The results of an LPA are presented in Figure 1 and 2, displaying distinct profiles of school alienation among seventh and tenth graders, respectively. The results of an LPA identifying four distinct profiles of school alienation

among seventh graders (Figure 1). Profile 1, the largest group (45%), was characterized by low alienation across all dimensions: classmates, teachers, and learning. Profile 2 (46%) exhibited moderate alienation from learning, while Profile 3 (7%) showed high alienation specifically from the perceived usefulness of learning. The smallest group, Profile 4 (2%), displayed high alienation across all dimensions, particularly from the affective aspect of learning, meaning they find learning boring or uninteresting. As can be seen in Figure 2, the results of an LPA for tenth graders, revealing three distinct profiles. Profile 1 (55%) represented students with low alienation across all dimensions, similar to Profile 1 in Grade 7. Profile 2 (43%) was characterized by moderate alienation from learning, mirroring Profile 2 in Grade 7. Profile 3 (2%) represented students with high alienation across all dimensions, similar to Profile 4 in Grade 7. Notably, the profile representing high alienation from the usefulness of learning (Profile 3 in Grade 7) was absent in the results in Grade 10, suggesting that this specific form of alienation might decrease as students progress through secondary school.

Figure 1: Profile Solutions for Seventh Graders

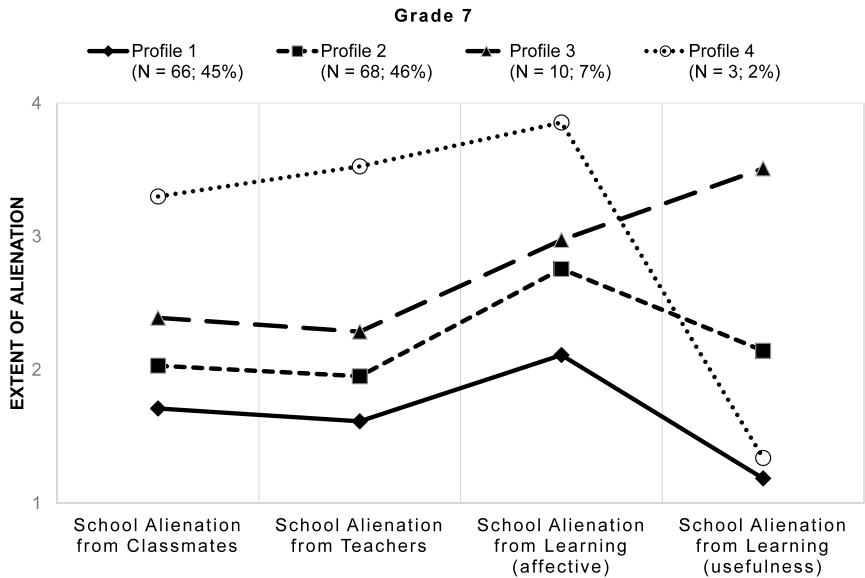
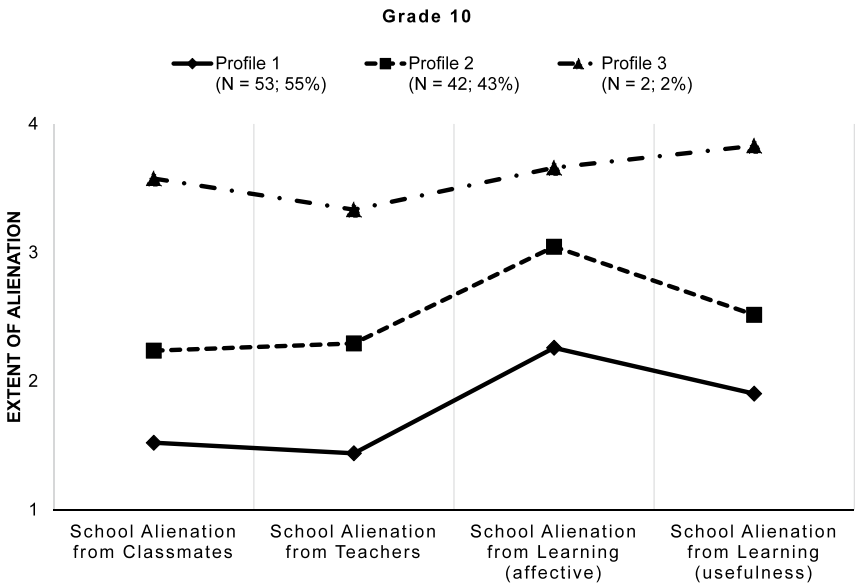


Figure 2: Profile Solutions for Tenth Graders



5.2. School Alienation Profiles Associated With Students’ Individual Characteristics

In this section, we performed a cross-tab analysis to examine the connections between school alienation profiles and students’ individual characteristics. In both groups of students (Grades 7 and 10), the results demonstrated weak positive correlations between alienation profiles and four aspects of students’ individual characteristics. Phi correlation coefficients (ϕ) of these associations ranged between 0.083 and 0.233. Table 4 shows in which ways the individual students’ characteristics regarding gender, immigrant background, school track, and grade retention differ between the alienation profiles in Grade 7 and 10. It is notable that students in the higher alienation profiles are predominantly male, come from immigrant backgrounds and have experienced at least one grade retention.

Table 4: School Alienation Profiles Associated With Students' Individual Characteristics

Profile (P)	Gender: Female	Immigrant Background	School Track: Gymnasium	Grade Retention
Grade 7				
P1: low alienation from classmates, teachers and learning	45.5%	46.2%	56.1%	16.7%
P2: moderate alienation from learning	50.7%	38.2%	60.3%	16.2%
P3: high alienation from learning (usefulness)	40.0%	80.0%	60.0%	40.0%
P4: high alienation from classmates, teachers and learning (affective)	0.0%	0.0%	33.3%	33.3%
Total	46.6%	44.1%	57.8%	18.4%
Phi coefficients	0.129	0.233	0.083	0.163
Grade 10				
P1: low alienation from classmates, teachers and learning	36.5%	48.1%	20.8%	49.1%
P2: moderate alienation from learning	36.6%	46.3%	23.8%	51.2%
P3: high alienation from classmates, teachers and learning	0.0%	0.0%	0.0%	100.0%
Total	35.8%	46.3%	21.6%	51.1%
Phi coefficients	0.109	0.137	0.085	0.144

Note. Valid sample included: 147 seventh graders (3.3% missing) and 97 tenth graders (1% missing).

6. Discussion

This study aimed to identify alienation profiles among seventh and tenth graders in German secondary schools and to examine the relation of domain-specific school alienation and individual students' characteristics, including gender, immigrant background, school track, and grade retention.

Regarding the alienation profiles identified among seventh and tenth graders, four distinct profiles emerged for seventh graders and three for tenth graders. Most seventh graders were categorized in profiles exhibiting low or moderate alienation. Thus, a substantial 91% of seventh graders were not severely alienated in any domain. However, a minority of seventh graders were classified into profiles representing high alienation, either from the perceived usefulness of learning (7%) or across all three domains (classmates, teachers, and learning). Similarly, among tenth graders, the majority exhibited low alienation across all dimensions and of moderate alienation from learning.

Although high levels of school alienation were relatively uncommon in both grades, the issue remains significant due to its potential long-term negative effects, such as a high risk of school dropout with associated disadvantages. These findings align with prior research in Switzerland and Luxembourg (Hadjar et al., 2021), which also found small sample sizes in high alienation profiles. However, our study reveals higher alienation from learning in our sample of German secondary students compared to the Swiss

and Luxembourgish contexts, which may reflect country-specific differences. Importantly, this study is the first to offer insights into alienation profiles among tenth graders, filling a gap left by previous research.

A notable finding regarding the development of high alienation profiles from seventh grade to tenth grade is that the perceived usefulness of learning decreases. It seems that tenth graders see less usefulness in the offered learning compared to seventh graders. Based on this finding, strategies to prevent alienation from learning can be derived by putting special interest on developing interesting learning content that resonates with students' everyday life.

As this study identified domain-specific alienation profiles for seventh and tenth graders, this research extends the domain specific school alienation construct by Hascher and Hadjar (2018). To date, our study and the study of Hadjar et al. (2021) are the only studies on school alienation profiles. Both studies emphasize the importance of alienation from learning, which is a central aspect of the school alienation construct. Therefore, we assume that alienation from learning plays a special role in relation to the interplay with the other alienation domains. Furthermore, both studies show country-specific alienation profiles that change partly over time. This emphasizes the necessity to recognize the context-sensitive characteristics of the school alienation concept.

The present study also set out with the aim of analyzing the relation of domain-specific school alienation with individual students' characteristics. In doing so, we focussed on common sociodemographic characteristics, including gender, immigrant background as well as on school track and grade retention. A correlational analysis revealed that more boys are represented in the higher alienation profiles. The finding that boys experience higher levels of alienation than girls confirms findings from previous studies which have also found substantial gender differences (Grecu, 2020; Hadjar & Lupatsch, 2010; Hadjar et al., 2015; Hascher & Hagenauer, 2010; Leduc & Bouffard, 2017; Morinaj & Hascher, 2019; Morinaj et al., 2020; Scharf, 2021). Moreover, the current study found that students with immigrant background are more frequently represented in the high alienation profiles. This finding aligns with studies showing that students with an immigrant background experience higher levels of alienation compared to native students (Hadjar et al., 2021; Morinaj et al., 2020; Scharf, 2021). While previous results have been mixed, our study reinforces the importance to consider immigrant background in school alienation research. This study found that students in lower school tracks, such as Realschule or Hauptschule (which prepare students for vocational training), are more frequently represented in high alienation profiles. This contrasts with previous research suggesting that school alienation is

highest in the high-achieving tracks (Grecu, 2020; Scharf, 2021). According to stage-environment fit theory, our findings may suggest that the highest track, Gymnasium, better addresses students' needs compared to lower tracks. However, given Gymnasium's highly competitive learning environment, which offers limited student autonomy, we rather assume that students in this track are more likely to adapt to the schools' demands by suppressing their personal needs.

Also students who have previously repeated a grade are more frequently represented in the high alienation profiles. Overall, this study confirms the importance of students' characteristics representing risk factors for developing school alienation. However, correlations between alienation profiles and gender, immigrant background, school track, and track retention are weak.

When interpreting these findings, the following limitations should be noted. First, the present study employed a cross-sectional research design which does not allow for causal interference to be drawn from the findings. Second, the sample of this study is not representative for all federal states in Germany. Nevertheless, all relevant school tracks from secondary school in North Rhine-Westphalia and Thuringia were included. Third, latent profile analysis face the limitation fit indices used to evaluate certain profile solutions can sometimes be contradictory. After thoroughly reviewing a range of indices, we selected the solutions with the most favorable values. Additionally, we took into account the sizes of the profile groups and whether these profiles were meaningful in light of conceptual considerations. Fourth, data were conducted during the COVID-19 pandemic. It remains unclear in which ways the specific conditions of homeschooling affect the alienation profiles and the allocation of students to these profiles.

The findings of this study have important implications both for research and practice in the field of empirical educational research. This study validates the conceptualisation of school alienation as a multidimensional construct as the level and domains affected differ over the course of schooling. Building up on Haschers and Hadjars (2018) conceptualisation of school alienation we further differentiated the domain alienation from learning. While Hascher and Hadjar (2018) combined usefulness and affective aspects under one domain, our analysis underscores the importance of distinguishing these two aspects. When applying the school alienation scale (Morinaj et al., 2017) to new educational contexts it is recommended to reflect on necessary adjustments within the given frame. Overall, this study provides initial insights into student characteristics in the context of school alienation. Future research is needed for deeper analyses focussing on individual and contextual factors to learn more about alienated students and their perceptions of the school environment.

Furthermore, the findings underscore the importance of informing both teachers and parents about the various expressions of school alienation to better recognize and address these issues early on. Integrating knowledge about school alienation into teacher training programs can equip educators with the skills needed to identify and mitigate alienation in their students effectively. Positive attitudes towards school and learning are also shaped by ongoing school development efforts. It is essential to prioritize students' needs when designing learning environments and engaging learning content. To foster students' autonomy, they must be actively involved in the decision-making processes development processes that shape their schools' development. Engaging students in this way not only enhances their sense of ownership but also ensures that the learning environment better aligns with their interests and needs.

Building on the findings of this study, several future research directions are suggested to further explore and address the issue of school alienation. Adopting a longitudinal perspective over the course of upper secondary school to investigate how school alienation develops over time and to explore the conditions under which students bonding to school increases again and levels of school alienation drop. Further studies are needed to investigate the specific conditions in both school and family environments that contribute to school alienation. This research could help identify protective factors, such as teaching or parenting styles, that may reduce the risk of alienation and support student well-being. Additionally, it would be valuable to examine how different school alienation profiles are linked to distinct student outcomes, such as well-being, academic achievement, mental health, and future educational or career trajectories. This could help tailor interventions and support strategies for specific student groups experiencing different forms of alienation.

Considering the overall findings, it is evident that addressing school alienation requires a multifaceted approach that includes both preventive and intervention strategies tailored to the specific needs of students. Fostering a more supportive and inclusive educational environment that promotes student engagement, well-being, and bonding is beneficial for students' educational achievement and the overall learning environment in school.

Acknowledgments: The authors gratefully acknowledge the support received.

References

- Alscher, P., Grecu, A., Ludewig, U., & McElvany, N. (2023). Towards the measurability of a competence model for school-based civic education. In N. McElvany, A. L. Grecu, R. Lorenz, M. Becker, C. Dignath, H. Gaspard, & F. Lauermann (Eds.), *50 Jahre Schulentwicklung – Leitthemen der empirischen Bildungsforschung* (pp. 230–261). Beltz Juventa.
- Archambault, I., Janosz, M., Fallu, J. S., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence*, 32(3), 651–670. <https://doi.org/10.1016/j.adolescence.2008.06.007>
- Avci, R., & Çelikkaleli, Ö. (2016). Peer victimization, trait anger and alienation as predictors of violence tendency in adolescents. *Journal of Educational Sciences Research*, 6(2), 151–167. <https://doi.org/10.12973/jesr.2016.62.9>
- Baker, J. A., Clark, T. P., Crowl, A., & Carlson, J. S. (2009). The influence of authoritative teaching on children's school adjustment. Are children with behavioural problems differentially affected? *School Psychology International*, 30(4), 374–382. <https://doi.org/10.1177/0143034309106945>
- Barber, B. L., Stone, M. R., Hunt, J. E., & Eccles, J. S. (2005). Benefits of activity participation. In L. M. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized Activities as Contexts of Development* (pp. 185–210). Erlbaum.
- Becker, M. (2024). Sozioemotionale Charakteristika und Bildung [Socio-emotional characteristics and education]. *Zeitschrift für Erziehungswissenschaft*, 1–38. <https://doi.org/10.1007/s11618-024-01217-8>
- Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, 9(2), 291–319. <https://doi.org/10.1017/S095457949700206X>
- Bourdieu, P. (1983). Ökonomisches Kapital – Kulturelles Kapital – Soziales Kapital [Economic – capital – Cultural capital – Social capital]. In R. Kreckel (Ed.), *Soziale Ungleichheiten* (pp. 183–198). Schwartz.
- Bronfenbrenner, U. (1986). Alienation and the four worlds of childhood. *The Phi Delta Kappan*, 67(6), 430–436.
- Brown, M. R., Higgins, K., Pierce, T., Hong, E., & Thoma, C. (2003). Secondary students' perceptions of school life with regard to alienation: The effects of disability, gender and race. *Learning Disability Quarterly*, 26(4), 227–238. <https://doi.org/10.2307/1593636>
- Çağlar, Ç. (2013). The relationship between the perceptions of the fairness of the learning environment and the level of alienation. *Eurasian Journal of Educational Research*, 50, 185–206.
- Calabrese, R. L. (1987). Adolescence: A growth period conducive to alienation. *Adolescence*, 22(88), 929–938.
- Calabrese, R. L., & Seldin, C. A. (1986). Adolescent alienation. *The High School Journal*, 69, 120–125.
- Dekel, R., & Tuval-Mashiach, R. (2012). Multiple losses of social resources following collective trauma: the case of the forced relocation from Gush Katif. *Psychological Trauma: Theory, Research, Practice and Policy*, 4(1), 56–65. <https://doi.org/10.1037/a0019912>
- Dynan, M. P. (1980). *Do schools care?* Education Department of Western Australia.
- Eccles, J. S., & Midgley, C. (1989). Stage-Environment Fit. In C. Ames & R. Ames (Eds.), *Research on Motivation in Education* (pp. 139–186), Academic Press.
- Eccles, J. S., Midgley, C., Wigfield, A., Miller, C., Flanagan, C., & Mac Iver, D. L. (1993). Development during adolescence. The impact of stage-environment fit on young adolescents. *American Psychologist*, 48(2), 90–101.
- Finn, J. D. (1989). Withdrawing from School. *Review of Educational Research*, 59(2), 117–142. <https://doi.org/10.3102/00346543059002117>
- Geiser, C. (2013). *Data analysis with Mplus*. Guilford Press.

- Govorova, E., Benítez, I., & Muñiz, J. (2020). How schools affect student well-being: A cross-cultural approach in 35 OECD Countries. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00431>
- Grecu, A. (2020). *Die Entwicklung der Entfremdung vom Lernen in der unteren Sekundarstufe in Luxemburg: Der Beitrag differenzieller schulischer Lern- und Entwicklungsmilieus* [The development of alienation from learning in lower secondary education in Luxembourg: The contribution of differentiated school learning and development environments]. Beltz Juventa.
- Grecu, A., Hadjar, A., & Simoes Loureiro, K. (2022). The role of teaching styles in the development of school alienation and behavioral consequences: A mixed methods study of Luxembourgish primary schools. *SAGE Open*, 12(2), 1–15. <https://doi.org/10.1177/21582440221105477>
- Grecu, A., Hascher, T., & Hadjar, A. (2019). Teachers' images of the ideal student as a marker for school culture and its role for school alienation during the transition from primary to secondary education in Luxembourg. *Studia Paedagogica*, 24(2), 85–108. <https://doi.org/10.5817/SP2019-2-4>
- Hadjar, A., Backes, S., & Gysin, S. (2015). School alienation, patriarchal gender-role orientations and the lower educational success of boys. A mixed-method study. *Masculinities & Social Change*, 4(1), 85–116. <https://doi.org/10.4471/MCS.2015.61>
- Hadjar, A., Grecu, A., Scharf, J., de Moll, F., Morinaj, J., & Hascher, T. (2021). Changes in school alienation profiles among secondary school students and the role of teaching style. *International Journal of Educational Research*, 105, Article 101697, 1–14. <https://doi.org/10.1016/j.ijer.2020.101697>
- Hadjar, A., & Gross, C. (2016). Introduction. In A. Hadjar & C. Gross (Eds.), *Education Systems and Inequalities* (pp. 1–9). Policy Press.
- Hadjar, A., & Lupatsch, J. (2010). The lower educational success of boys. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62(4), 599–622. <https://doi.org/10.1007/s11577-010-0116-z>
- Hagenauer, G., & Volet, S. E. (2014). Teacher-student relationship at university: An important yet under-researched field. *Oxford Review of Education*, 40(3), 370–388. <https://doi.org/10.1080/03054985.2014.921613>
- Hascher, T., & Hadjar, A. (2018). School alienation – Theoretical approaches and educational research. *Educational Research*, 60(2), 171–188. <https://doi.org/10.1080/00131881.2018.1443021>
- Hascher, T., & Hagenauer, G. (2010). Alienation from school. *International Journal of Educational Research*, 49(6), 220–232. <https://doi.org/10.1016/j.ijer.2011.03.002>
- Hascher, T., & Hagenauer, G. (2020). Swiss adolescents' well-being in school. *Swiss Journal of Educational Research*, 42(2), 367–390. <https://doi.org/10.24452/sjer.42.2.5>
- Hofer, M. (2004). Schüler wollen für die Schule lernen, aber auch anderes tun. Theorien der Lernmotivation in der Pädagogischen Psychologie [Students want to learn for school, but also want to do other things. Theories of learning motivation in educational psychology]. *Zeitschrift für Pädagogische Psychologie*, 18(2), 79–92. <https://doi.org/10.1024/1010-0652.18.2.79>
- Jung, T., & Wickrama, K. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302–317. <https://doi.org/10.1111/j.1751-9004.2007.00054.x>
- Kocayörük, E., & Şimşek, Ö. F. (2016). Parental attachment and adolescents' perception of school alienation: The mediation role of self-esteem and adjustment. *The Journal of Psychology*, 150(4), 405–421. <https://doi.org/10.1080/00223980.2015.1060185>
- Kocayörük, E., Uzman, E., & Mert, A. (2014). How the attachment styles associated with student alienation: The mediation role of emotional well-being. *International Journal of Progressive Education*, 10(3), 34–46.

- Landesbetrieb Information und Technik Nordrhein-Westfalen (2022). *40,0 Prozent der Schülerinnen und Schüler in NRW hatten im Schuljahr 2021/22 eine Zuwanderungsgeschichte* [40.0 percent of students in NRW had an immigrant background in the school year of 2021/22]. Retrieved October, 09, 2024 from <https://www.it.nrw/400-prozent-der-schuelerinnen-und-schueler-nrw-hatten-im-schuljahr-202122-eine-18029>
- Leduc, C., & Bouffard, T. (2017). The impact of biased self-evaluations of school and social competence on academic and social functioning. *Learning and Individual Differences*, 55, 193–201. <https://doi.org/10.1016/j.lindif.2017.04.006>
- Legault, L., Green-Demers, I., & Pelletier, L. (2006). Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of Educational Psychology*, 98(3), 567–582. <https://doi.org/10.1037/0022-0663.98.3.567>
- Mahmoudi, H., Brown, M. R., Saribagloo, J. A., & Dadashzadeh, S. (2015). The role of school culture and basic psychological needs on Iranian adolescents' academic alienation. *Youth and Society*, 1–21. <https://doi.org/10.1177/0044118X15593668>
- Mann, S. J. (2001). Alternative perspectives on the student experience: Alienation and engagement. *Studies in Higher Education*, 26(1), 7–19. <https://doi.org/10.1080/03075070020030689>
- Marcin, K., Morinaj, J., & Hascher, T. (2020). The Relationship between Alienation from Learning and Student Needs in Swiss Primary and Secondary Schools. *Pädagogische Psychologie*, 34(1), 35–49. <https://doi.org/10.1024/1010-0652/a000249>
- Mau, R. Y. (1992). The validity and devolution of a concept. *Adolescence*, 27(107), 731–741.
- Morinaj, J., de Moll, F., Hascher, T., Hadjar, A., Grecu, A., & Scharf, J. (2023). School alienation among adolescents in Switzerland and Luxembourg: The role of parent and peer supportive attitudes toward school and teacher autonomy support. *Youth & Society*, 55(2), 187–212. <https://doi.org/10.1177/0044118X211043902>
- Morinaj, J., Hadjar, A., & Hascher, T. (2020). School alienation and academic achievement in Switzerland and Luxembourg: A longitudinal perspective. *Social Psychology of Education*, 23(2), 279–314. <https://doi.org/10.1007/s11218-019-09540-3>
- Morinaj, J., & Hascher, T. (2019). The longitudinal relation between student well-being and academic achievement. In L. Castelli, J. Marcionetti, A. Plata, & A. Ambrosetti (Eds.), *Well-being in education systems* (pp. 273–277). Hogrefe Editor.
- Morinaj, J., Marcin, K., & Hascher, T. (2019). School alienation, student learning and social behavior in challenging times. In S. A. Karabenick & T. C. Urdan (Series Eds.) and E. N. Gonida & M. S. Lemos (Guest Eds.), *Advances in Motivation and Achievement Series (Vol. 20): Motivation in Education at a Time of Global Change: Theory, Research, and Implications for Practice* (pp. 205–224). Emerald Group Publishing Limited. <https://doi.org/10.1108/S0749-742320190000020010>
- Morinaj, J., Scharf, J., Grecu, A., Hadjar, A., Hascher, T., & Marcin, K. (2017). School alienation. A construct validation study. *Frontline Learning Research*, 5(2), 36–59. <https://doi.org/10.14786/flr.v5i2.298>
- Murdock, T. B. (1999). The social context of risk: Status and motivational predictors of alienation in middle school. *Journal of Educational Psychology*, 91(1), 62–75. <https://doi.org/10.1037/0022-0663.91.1.62>
- Myers, S. S., & Pianta, R. C. (2008). Individual and contextual influences on student-teacher relationships and children's early problem behavior. *Journal of Clinical Child & Adolescent Psychology*, 37(3), 600–608. <https://doi.org/10.1080/15374410802148160>
- Newman, B. M., & Newman, P. R. (2003). *Development through Life*. (8th ed.). Wadsworth.
- Nutbeam, D., Smith, C., Moore, L., & Bauman, A. (1993). Warning! Schools can damage your health: Alienation from school and its impact on health behaviour. *Journal of Paediatrics and Child Health*, 29(s1), 25–30. <https://doi.org/10.1111/j.1440-1754.1993.tb02256.x>

- Patall, E. A., Hooper, S., Vasquez, A. C., Pituch, K. A., & Steingut, R. R. (2018). Science class is too hard: Perceived difficulty, disengagement, and the role of teacher autonomy support from a daily diary perspective. *Learning and Instruction*, 58, 220–231. <https://doi.org/10.1016/j.learninstruc.2018.07.004>
- Pianta, R. C., & Hamre, B. K. (2009). Classroom processes and positive youth development: Conceptualizing, measuring, and improving the capacity of interactions between teachers and students. *New Directions for Youth Development*, 121, 33–46. <https://doi.org/10.1002/yd.295>
- Raccanello, D., Brondino, M., Moè, A., Stupnisky, R., & Lichtenfeld, S. (2019). Enjoyment, boredom, anxiety in elementary schools in two domains: Relations with achievement. *The Journal of Experimental Education*, 87(3), 449–469. <https://doi.org/10.1080/00220973.2018.1448747>
- Rayce, S. L., Holstein, B. E., & Kreiner, S. (2008). Aspects of alienation and symptom load among adolescents. *The European Journal of Public Health*, 19(1), 79–84. <https://doi.org/10.1093/eurpub/ckn105>
- Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *The Elementary School Journal*, 100(5), 443–471. <https://doi.org/10.1086/499650>
- Scharf, J. (2021). *Bildungswerte und Schulentfremdung. Schulische Kontexteffekte in Luxemburg und im Schweizer Kanton Bern* [Educational values and school alienation. School context effects in Luxembourg and the Swiss canton of Bern]. Beltz Juventa.
- Schmid, A., Morinaj, J., & Hascher, T. (2021). On the relation between school alienation and social school climate. *Schweizerische Zeitschrift für Bildungswissenschaften*, 43(3), 451–463. <https://doi.org/10.25656/01:23787>
- Seeman, M. (1959). On the meaning of alienation. *American Sociological Review*, 24(6), 783–791. <https://doi.org/10.2307/2088565>
- Sidorkin, A. M. (2004). In the event of learning: Alienation and participative thinking in education. *Educational Theory*, 54(3), 251–262. <https://doi.org/10.1111/j.0013-2004.2004.00018.x>
- Steuer, G., Grecu, A. L., & Morinaj, J. (2024). Error climate and alienation from teachers: A longitudinal analysis in primary school. *The British Journal of Educational Psychology*, 1–14. <https://doi.org/10.1111/bjep.12659>
- Studsrod, I., & Bru, E. (2012). Upper secondary school students' perceptions of teacher socialization practices and reports of school adjustment. *School Psychology International*, 33(3), 308–324. <https://doi.org/10.1177/0143034311412841>
- Sutherland, A. (2011). The relationship between school and youth offending. *Social Policy Journal of New Zealand*, 37, 51–69.
- Symonds, J. E., & Galton, M. (2014). Moving to the next school at age 10–14 years: an international review of psychological development at school transition. *Review of Education*, 2(1), 1–27. <https://doi.org/10.1002/rev3.3021>
- Wang, M.-T., & Fredericks, J. A. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Development*, 85(2), 722–737. <https://doi.org/10.1111/cdev.12138>
- Wigfield, A., Byrnes, J. P., & Eccles, J. S. (2006). Development during early and middle adolescence. In P. A. Alexander & P. H. Winne (Eds.), *The Handbook of Educational Psychology*, (pp. 87–113). Lawrence Erlbaum Associates.
- Organization for Economic Co-operation and Development. (2003). *Student engagement at school. A sense of belonging and participation: Results from PISA 2000*. OECD Publishing. <https://doi.org/10.1787/9789264018938-en>

Kulturre sponsives Schulleitungs- handeln als Bedingung erfolgreicher Bildungsprozesse von Schüler*innen aus marginalisierten Herkunftsmilieus – Ein erster Kartierungsversuch für den deutschen Sprachraum

Christine Becks*, Susanne Julia Czaja* &
Esther Dominique Klein*

* Geteilte Erstautorinnenschaft

Zusammenfassung

Für den Bildungserfolg von Schüler*innen aus marginalisierten Herkunftsmilieus erscheint es bedeutsam, dass Lehrkräfte über kulturre sponsive Handlungskompetenzen verfügen. Vor diesem Hintergrund werden im englischsprachigen Raum Potenziale eines kulturre sponsiven Schulleitungshandelns (KRS�; Khalifa et al., 2016) diskutiert. In der deutschsprachigen Schulleitungsforschung spielt der Ansatz bislang jedoch keine Rolle. Der Beitrag bietet einen Überblick über die Forschung zu KRS� auf Basis eines systematischen Literature Reviews und diskutiert Adaptionismöglichkeiten für den deutschsprachigen Kontext. Die Analyse zeigt, dass bei einer Übertragung des Ansatzes fehlende systematische Unterstützungsstrukturen für Schulleitende, die flachen Hierarchien in Schulen sowie eine systemisch legitimierte Defizitperspektive auf Schüler*innen aus marginalisierten Herkunftsmilieus in Deutschland berücksichtigt werden sollten.

Schlagworte: Kulturre sponsives Schulleitungshandeln, Schulen an sozialräumlich benachteiligten Standorten, Führung, Systematic Literature Review

Abstract

The educational success of students from marginalized communities is influenced by educators' abilities to think and act in culturally responsive ways. In anglophone countries, scholars therefore discuss the potential of culturally responsive school leadership (CRSL; Khalifa et al., 2016). However, this approach has not yet found its way into school leadership research in Germany. This contribution provides an overview of the research on CRSL based on a systematic literature review, and discusses possible adaptations for school leadership in Germany. The analysis indicates that in transferring the approach, we must consider the lack of systematic support structures for school leaders, the flat hierarchies in schools, and a systemically legitimized deficit perspective on students from marginalized communities in Germany.

Keywords: Culturally responsive school leadership, schools serving disadvantaged communities, leadership, systematic literature review

1. Einleitung

Der Verlauf von Bildungsprozessen von Schüler*innen aus marginalisierten Herkunftsmilieus ist in hohem Maße durch die Praxis der Schulen geprägt, die sie besuchen (z. B. Palardy, 2008). Sie erfahren häufiger als Schüler*innen aus privilegierten Herkunftsmilieus Diskriminierung und Abwertung, und machen in der Folge multiple Scheiternserfahrungen im Bildungssystem. Neben familiären Ressourcen sind schulische Strukturen sowie Handlungspraxen bedeutsam, die auf der historischen Marginalisierung bestimmter Schüler*innengruppen aufbauen und auch von den schulischen Akteur*innen reproduziert werden.

Um marginalisierende Praktiken aufzubrechen, verfolgen Ansätze und Forschungsstränge von *Cultural Responsiveness* seit Anfang der 1990er Jahre das Ziel, schulisches Handeln responsiv mit Blick auf die Lebenswelt und Erfahrungen der Schüler*innen zu gestalten (z. B. Ladson-Billings, 1992, 1994; Moll & Gonzalez, 1994; Gay, 2000, 2002). Wenngleich in der deutschsprachigen Schulforschung und -praxis ein Bewusstsein für diese Herausforderung besteht, werden Theorieansätze aus der internationalen Bildungsforschung, die zu einer Systematisierung kulturre sponsiver schulischer Strategien genutzt werden können, bislang nur vereinzelt rezipiert (z. B. Ialuna, Civitillo & Jugert, 2024; Herzog-Punzenberger et al., 2020).

Der Beitrag stellt vor diesem Hintergrund den Ansatz des kulturreponsiven Schulleitungshandelns (Culturally Responsive School Leadership; Khalifa, 2018; Khalifa et al., 2016) vor. Im Anschluss bietet der Beitrag einen Überblick über die Forschung zum Thema durch ein systematisches Literature Review (Hallinger, 2013) und diskutiert Adaptionismöglichkeiten für den deutschsprachigen Kontext.

2. Begründungslinien für ein Kulturreponsives Schulleitungshandeln (KRSL)

In der Raumsoziologie beschreibt der Begriff der Marginalisierung das ‚an den Rand drängen‘ bestimmter gesellschaftlicher Gruppen im Kontext zunehmender Segregation, insbesondere in urbanen Räumen (z. B. Hertel, 2021). Damit ist häufig eine Ausgrenzung hinsichtlich des Zugangs zu knappen Ressourcen und der Teilhabe am gesellschaftlichen Leben verbunden.

Übertragen auf den Schulbereich verstehen wir unter marginalisierten Herkunftsmilieus all jene Herkunftsmilieus, die systematisch weniger an Bildung partizipieren bzw. von Bildungsangeboten profitieren und deren Bedürfnisse an schulisches Handeln, etwa aufgrund geringerer für Schule relevanter Ressourcen, Passungsproblematiken sowie Diskriminierungs- und Abwertungserfahrungen, in den institutionellen Rahmenbedingungen von Schule wenig oder keine Berücksichtigung finden. Von Marginalisierung betroffen sind vor allem Schüler*innen aus sozioökonomisch stark benachteiligten Milieus, aus migrantischen Milieus sowie *Students of Color*, wobei auch weitere Gruppen einbezogen werden können.

Während in Schulleistungsstudien insbesondere der direkte Zusammenhang zwischen den Herkunftsmilieus der Schüler*innen und ihren Kompetenzen fokussiert wird, zeigt die internationale Befundlage, dass Bildungsprozesse von Schüler*innen aus marginalisierten Herkunftsmilieus in hohem Maße auch durch die Praxis der Schulen geprägt sind (Palardy, 2008).

2.1. Vom Achievement Gap zum Education Debt

Mehr als zwei Jahrzehnte nach dem Erscheinen der ersten Schulleistungsstudien haben sich Disparitäten zwischen Schüler*innen aus marginalisierten und privilegierten Milieus in Bezug auf Kompetenzentwicklung, Bildungsabschlüsse und *Well-Being* kaum verändert (z. B. Organization for Economic Co-operation and Development [OECD], 2019). Die empirisch beobachtbaren

Bildungsdisparitäten verdeutlichen beispielsweise, dass das meritokratische Prinzip an seine Grenzen stößt, wenn Schüler*innen aus marginalisierten Herkunftsmilieus unabhängig von ihren Fähigkeiten strukturelle Nachteile erfahren. Angesichts der hohen Zahl von Schüler*innen aus solchen Milieus, die Mindeststandards nicht erreichen, stellt sich auch die Frage, ob es Schule gelingt, allen Mitgliedern der nachwachsenden Generation jene Kompetenzen zu vermitteln, die für die Reproduktion der Gesellschaft notwendig sind. Schließlich ist zu hinterfragen, inwieweit die sozialen Bildungsdisparitäten die integrative Kraft der Schule beeinträchtigen, wenn geringere Bildungschancen auch geringere Chancen der gesellschaftlichen Teilhabe bedeuten. Insofern stellen anhaltende Bildungsdisparitäten nicht nur ein Problem hinsichtlich der individuellen Bildungschancen der Schüler*innen dar, sondern rücken auch die Frage ins Zentrum, inwiefern Schule in der Lage ist, die an sie gestellten gesellschaftlichen Funktionen (vgl. Fend, 1980) zu erfüllen. Damit lässt sich schultheoretisch begründen, dass die Reduzierung von Bildungsdisparitäten (auch) in der Verantwortung von Schule liegt.

Angesichts vergleichbarer Befunde der letzten vier Jahrzehnte in den USA legte Ladson-Billings (2006) eine Neukonzeption des *Achievement Gaps* als *Education Debt* vor, wodurch vermeintlich individuelle Kompetenzunterschiede in Leistungsvergleichen als Ergebnisse historischer, wirtschaftlicher, sozio-politischer und moralischer Ungleichheit erkannt werden. So werden die vermeintlichen, individuellen ‚Defizite‘ der Schüler*innen nicht als deren individuell zu verantwortende Unfähigkeit verstanden, sondern als Spiegelbild klaffender gesellschaftlicher, generationaler Benachteiligungen rekonfiguriert. Damit entsteht ein neuer Blick auf Leistungsunterschiede: Als *Education Debt*, also als Bildungsschuld des Schulsystems gegenüber dem Subjekt, rückt schulisches, pädagogisches Handeln in den Fokus der Verantwortung, um ebendiese gesellschaftlichen, generationalen Benachteiligungen zu ringen und sie im Kontext von Schule bestmöglich zu reflektieren und abzuschwächen. Diese Forderung erreicht auch Forschung und Politik, die durch die Zuwendung zu Leistungsunterschieden statt *Education Debt* jene sozialen Trennlinien replizieren und vertiefen, denen sie entgegenzustehen haben (Salmen, 2019).

2.2. Der schulische Beitrag zur Reproduktion des Achievement Gap

Welche Rolle Schule bzw. schulische Praxis bei der Entstehung des *Achievement Gaps* einnimmt, wird kontrovers diskutiert. Sendzik et al. (2023) verweisen beispielweise auf die Perspektive der Bildungssoziologie, nach der

Schulen die Benachteiligungen marginalisierter sozialer Herkunftsmilieus nicht aufheben, sondern bestenfalls kompensieren können. Diese Sichtweise gehe davon aus, dass die familiäre Herkunft „zusammengenommen eine solch hohe Determinationskraft [entfaltet], dass es – trotz der moderierenden Wirkung von Schule – letztendlich zu einer intergenerationalen Vererbung von Bildung (und Bildungsarmut) kommt“ (S. 179).

Zugleich legen Befunde nahe, dass die Reduktion der Rolle von Schule auf Kompensation (vermeintlich oder tatsächlich) bestehender ‚Defizite‘ vermutlich zu kurz greift. So gibt es zahlreiche Forschungsergebnisse, die belegen, dass die kollektiven Selbstwirksamkeitserwartungen von Lehrkräften an Schulen an sozialräumlich benachteiligten Standorten (SsbS) durch die Herkunftsmilieus der Schüler*innen beeinflusst werden und diese Erwartungen sich entscheidend auf die Kompetenzentwicklung der Schüler*innen auswirken (z. B. Bandura, 1993). Weitere Befunde zeigen, dass Lehrkräfte im Schnitt geringere Erwartungen an Schüler*innen aus marginalisierten Herkunftsmilieus haben, was sich wiederum in deren tatsächlichen Leistungen widerspiegelt (z. B. de Boer et al., 2018).

Diese Effekte beschränken sich jedoch nicht nur auf die Einschätzung der individuellen Kompetenzen der Schüler*innen, sondern umfassen auch explizite oder implizite Annahmen über die Bedeutung ihrer Herkunft. Schüler*innen aus sozioökonomisch benachteiligten Milieus müssen beispielsweise deutlich höhere Leistungen erbringen, um eine Gymnasialempfehlung zu erhalten, als Schüler*innen aus privilegiierteren Herkunftsmilieus (z. B. Stubbe et al., 2023). Auch bei der Bewertung von Schüler*innenleistungen spielen herkunftsbezogene Vorannahmen eine Rolle (z. B. Bonefeld & Dickhäuser, 2018).

Erkenntnisse aus der Schulentwicklungsforschung, die sich explizit auf SsbS beziehen, können darlegen, dass es durchaus SsbS gibt, deren Schüler*innen überdurchschnittlich erfolgreich sind, und dass hierfür insbesondere die kollektiven innerschulischen Handlungspraxen bedeutsam sind (Klein, 2018). Die spezifischen schulischen Merkmale und Handlungspraxen, die in diesem Zusammenhang relevant sind, werden jedoch in *Large Scale Studies* in der Regel nicht erfasst und sind teilweise mit quantitativen Ansätzen auch nur begrenzt messbar, sodass die Frage, welchen Beitrag sie zum Abbau von Bildungsdisparitäten leisten können, bislang kaum quantitativ untersucht wurde (Sendzik et al., 2023; vgl. hierzu auch Valencia, 2010).

Eine Ausnahme bildet eine Längsschnittstudie von Palardy (2008) aus den USA, die zeigt, dass an SsbS mehr als 50% der Varianz in der Leistungsentwicklung der Schüler*innen durch Unterschiede zwischen den Schulen erklärt werden konnten, während der Anteil an privilegierten Schulstandorten deutlich geringer war. Relevant ist in diesem Zusammenhang aber nicht

nur die höhere Wirksamkeit schulischen Handelns für Schüler*innen aus marginalisierten Herkunftsmilieus, sondern auch die Frage, welche Handlungspraxen diese Unterschiede erzeugen. In der Studie von Palardy (2008) erwiesen sich beispielsweise die fachliche Qualifikation der Lehrkräfte und das Angebot von Kursen mit hohen Anforderungen als bedeutsam. Zudem gab es einen signifikanten Zusammenhang zwischen der Leistungsentwicklung und der Wahrnehmung der Schüler*innen, dass das *discipline management* der Schule ‚fair‘ sei. Bestimmte Interaktionspraktiken zwischen Lehrenden und Lernenden an SsBS – etwa ein restriktives Disziplinmanagement oder die Absenkung von Leistungsanforderungen – scheinen also Schüler*innen zusätzlich zu benachteiligen.

2.3. Deficit Thinking

Der Education Debt äußert sich entsprechend auch im Blick von Lehrkräften auf die von institutionalisierten Normalitätsvorstellungen vermeintlich abweichenden Schüler*innen, indem Lehrkräfte z. B. zum ultimativen Attributionsfehler (Pettigrew, 1979) tendieren: Bei der Attribution von als ‚abweichend‘ wahrgenommenem Verhalten oder geringeren Leistungen von Schüler*innen aus marginalisierten Herkunftsmilieus überschätzen sie die Bedeutung der Herkunft und unterschätzen die Wirkmacht ihres eigenen Handelns sowie situative Faktoren (Wang & Hall, 2018). Valencia (2010) prägte den Begriff des *Deficit Thinking*: Misserfolge werden auf internale ‚Defizite‘ der Schüler*innen (Verhalten, Motivation, sprachliche Fähigkeiten) attribuiert, welche ihrer familiären Herkunft zugeschrieben werden. Schulisches Handeln wird v. a. als Kompensation der ‚Defizite‘ betrachtet, nicht als Teil der Reproduktion von Bildungsungleichheit.

Racherbäumer (2017) differenziert auf der Basis ihrer Befunde zwei Typen von Defizitorientierungen. Einerseits finden sich in ihrer Studie Schulen, in denen die Überzeugungen der Lehrkräfte zwar von Defizitorientierungen geprägt sind, aber von einem pädagogischen Ethos der Fürsorge flankiert werden. Die Lehrkräfte sehen sich in der Verantwortung, die wahrgenommenen Defizite der Schüler*innen zu kompensieren; höhere Erwartungen an die Schüler*innen werden gleichwohl als ungerecht empfunden und die Anforderungen werden zum Schutz der Schüler*innen reduziert: „Obgleich die Schülerinnen und Schüler fachlich und individuell gefördert werden, scheint außer Frage zu stehen, dass sie den inhaltlich-fachlichen Anschluss an die altersentsprechende Lerngruppe nicht erreichen werden oder können“ (Racherbäumer, 2017, S. 131). In einem zweiten Typ der Distinktion manifes-

tieren sich die Defizitorientierungen in einer Abwertung der Schüler*innen, der geringe Erfolg oder das Verhalten der Schüler*innen wird als Ausdruck einer geringeren ‚Beschulbarkeit‘ (Educability) interpretiert, wodurch pädagogisch unethisches Verhalten legitimiert wird. Die Pädagog*innen äußern eine „subjektiv legitimierte und limitierte Handlungsfähigkeit hinsichtlich des Bildungs- und Erziehungsauftrages von Schule“ (Racherbäumer, 2017; S. 131).

Quantitative Befragungen aus verschiedenen, auf SsbS ausgerichteten Forschungs- und Entwicklungsprojekten im deutschsprachigen Raum legen nahe, dass Defizitorientierungen an SsbS durchaus verbreitet sind, wobei es zum Zusammenhang zwischen dem Grad der Marginalisierung der Schüler*innen und der Stärke der Defizitorientierungen inkonsistente Befunde gibt (Klein & Bronnert-Härle, 2022; Drucks & Bremm, 2021; Pumpow et al., 2024).

Wie sich defizitorientierte Perspektiven auf Schüler*innen aus marginalisierten Herkunftsmilieus äußern, ist inzwischen zudem in einer Vielzahl qualitativer Studien illustriert worden. Diverse Autor*innen berichten von mehr oder weniger impliziten Geringschätzungen der Fähigkeiten der Schüler*innen und der Konstruktion der Schüler*innen als ‚Bedürftige‘ (z. B. Wellgraf, 2012; Fölker & Hertel, 2015; Fölker et al., 2016). Hertel (2014) etwa beschreibt punitive Strategien als Folge einer Kriminalisierung der Schüler*innen, die in der Wahrnehmung der Lehrkräfte „durch einen Strafenkatalog überhaupt nicht zu beeindrucken“ seien (S. 390; vgl. auch Hertel, 2021). In anderen Studien äußern sich Aushandlungen der Frage, inwiefern die Schüler*innen das Engagement ihrer Lehrkräfte ‚verdient‘ haben (Bronnert-Härle & Klein, 2025). Wellgraf (2012) dokumentiert, wie Lehrkräfte geringere Chancen auf einen Ausbildungsplatz von Schüler*innen aus migrantisch geprägten Milieus mit gruppenbezogenen Einschätzungen des Verhaltens von ‚Ausländern‘ legitimieren.

Dass hierbei oftmals weniger das konkrete Verhalten einzelner Schüler*innen, sondern vielmehr implizite oder explizite Überzeugungen über Auswirkungen der Herkunftsmilieus verhandelt werden, illustriert Kollender (2020): Die von ihr interviewten schulischen Akteur*innen etwa schreiben den Eltern ihrer Schüler*innen zu, sich auf der Unterstützung durch den Staat ‚auszurufen‘, was wiederum auf deren Kinder abfärbe; die Aussagen kulminieren schließlich in Forderungen nach Möglichkeiten für die Schule, den Eltern staatliche Unterstützungsmaßnahmen abzuverlangen (vgl. zu Defizitorientierungen gegenüber Eltern auch Eiden, 2018).

Defizitorientierungen erfüllen dabei für Schulen und Lehrkräfte zentrale Funktionen. Anhand einer längsschnittlich angelegten Fallstudie können Drucks und Bremm (2021) darlegen, dass Defizitorientierungen zum einen ein selbstwertdienliches Attributionsmuster in einem Kontext darstellen, in

dem es aufgrund der höheren Anforderungen für Lehrkräfte deutlich herausfordernder ist, sich selbstwirksam zu fühlen. Zum anderen zeichnen die Autor*innen nach, dass Defizitorientierungen für SsbS identitätsstiftend sein und ihnen einen Daseinszweck liefern können.

2.4. Vom Deficit Thinking zu kulturresponsiver Handlungskompetenz

Deficit Thinking basiert im Sinne des Moral Debt (Ladson-Billings, 2006, S. 8) auf unhinterfragten Überzeugungen, die das Handeln ontologisch legitimieren (Gottuck et al., 2021; Valencia, 2010). Zwar verweisen diverse Studien darauf, dass Lehrkräfte eine durchaus positive Einstellung gegenüber Diversität haben und Benachteiligungen abbauen wollen (z. B. Bello, 2020). Zugleich rekrutiert sich aber der Großteil der Lehrkräfte nicht aus marginalisierten Herkunftsmilieus (Neugebauer, 2013). Mit Blick auf das Merkmal Migration und *Race* etwa zeigen Studien aus dem deutschen Sprachraum, dass gerade weiße Lehrkräfte sich häufig als *color-blind* verstehen und strukturelle Benachteiligungen sowie Mythen über marginalisierte Herkunftsmilieus wenig hinterfragen (Civittillo et al., 2021).

Quantitative Befunde aus dem angloamerikanischen Raum legen nahe, dass solche unhinterfragten Mythen insbesondere bei Lehrenden ohne biografischen Bezug zur Marginalisierung Wirkungen auf das Handeln entfalten können. Die Ergebnisse deuten darauf hin, dass Lehrkräfte mit *color-blind attitudes* geringere Erwartungen an ihre Wirkmacht für den Bildungserfolg von Schüler*innen mit Migrationsgeschichte haben (Cadenas et al., 2021). Ferner belegen Studien positive Effekte auf Leistungsentwicklung, motivationale Orientierungen und Fähigkeitsselbstkonzepte von Schüler*innen aus marginalisierten Herkunftsmilieus, wenn sie von Lehrkräften aus der gleichen Community unterrichtet werden (Egalite & Kisida, 2018; Hwang et al., 2023), oder ihre weißen Lehrkräfte nicht-weiße Peers haben (Gershenson et al., 2023). Lehrkräfte *of Color* unterstützen und empfehlen Begabtenförderung für Students of Color zudem häufiger (Grissom et al., 2020; Hart & Lindsay, 2024).

In den 1990er Jahren verdichteten sich die bestehenden Forschungslinien im anglo-amerikanischen Raum zu einem Programm der kulturresponsiven und kulturelevanten Pädagogik und Didaktik, welche die spezifischen Erfahrungshorizonte marginalisierter Schüler*innengruppen systematisch aufgreift, um effektivere Zugänge zum Lehren und Lernen für diese Schüler*innen zu entwickeln (Paris, 2012; Gay, 2002). Beispielsweise machen Schüler*innen aus benachteiligten Herkunftsmilieus häufiger belastende Armutserfahrungen, die ihr Selbstbild und ihre Motivation beeinflussen können

(Merten, 2003). Students of Color und Schüler*innen mit nicht-deutscher Familiensprache erfahren häufiger Alltagsrassismus, der auch von Angehörigen der Schule ausgeübt wird (Hertel, 2024; Aikins et al., 2021; Dean, 2020; Wellgraf, 2012).

Damit zusammenhängende Einstellungen, Erwartungen und Verhaltensmuster der Schüler*innen sind für Lehrkräfte schwer nachvollziehbar und bleiben diffus, wenn sie selbst diese Erfahrungen nicht gemacht haben (Beckett & Wrigley, 2014). In der Folge können sie dazu tendieren, unterschiedliche Lebenswelten zu negieren (Civitillo et al., 2021). Gleichsam folgt daraus nicht, dass pädagogisches Handeln in Schulen benachteiligter Quartiere nur möglich ist, wenn Lehrpersonen und Schüler*innen ähnliche Lebenskontexte und Erfahrungen aufweisen. Vielmehr geht es darum, Unterrichtsinhalte didaktisch so aufzuarbeiten, dass sie lernendenzentriert – und damit auf die Bezugshorizonte der Lebenswelt der Lernenden angepasst – sind und in der Folge relevante Bedeutung für die einzelnen Lernenden entfalten können. Das benötigt einerseits differenzierte, didaktisch bedachte Blicke auf Lernunterschiede zwischen marginalisierten und nicht-marginalisierten Schüler*innen (z. B. Khalifa, 2016; Au, 2009) und andererseits schulische Führung, die die Verpflichtung zu einer „pluralistischen Vision der Gesellschaft“ (Au, 2009, S. 179) trägt und für alle Mitglieder der Schulgemeinschaft verbindlich macht.

2.5. Kulturresponsives Schulleitungshandeln

Bei der Förderung kulturresponsiver Handlungskompetenz von Lehrkräften, insbesondere an Schulen an sozialräumlich benachteiligten Standorten, kommt der Schulleitung eine besondere Rolle zu. Studien, die international etablierte Führungsansätze wie *Transformational Leadership* (Leithwood & Jantzi, 2006) zugrunde legen, zeigen deutlich, dass zielorientiertes Führungshandeln positive Effekte beispielsweise auf die motivationalen Orientierungen von Lehrkräften und, darüber vermittelt, auf die Kompetenzentwicklung der Schüler*innen haben kann (z. B. Sun & Leithwood, 2015). Allerdings beschränken sich diese Studien meist auf die generelle Feststellung, dass Schulleitungen Ziele setzen, ohne zu spezifizieren, welcher Art diese Ziele sein sollten.

Die zuvor aufgeschlüsselte Befundlage zu SsbS legt jedoch nahe, dass hier die inhaltliche Ausrichtung des Führungshandelns entscheidend ist: Weniger erfolgreiche SsbS sind oft durch eine ‚pessimistische‘ Schulkultur und eine Tendenz zur Externalisierung von Verantwortung geprägt (Potter et al., 2002). Es ist dabei davon auszugehen, dass eine geringe Selbstwirksamkeit der

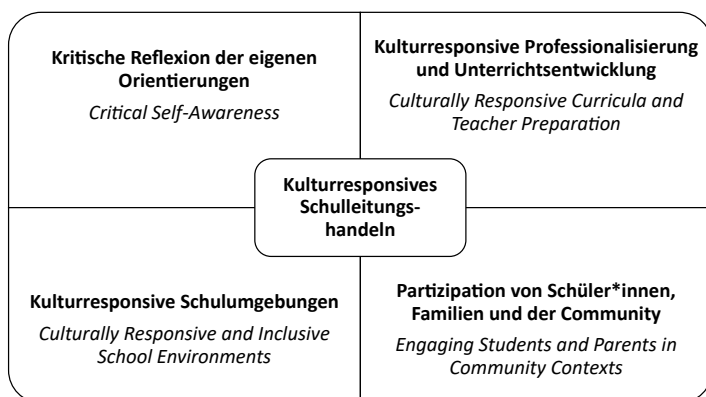
Lehrkräfte (Drucks & Bremm, 2021), erzeugt dadurch, dass die Lehrkräfte weniger Erfolge erfahren bzw. Erfolge schwieriger zu erreichen sind, als an anderen Standorten, zu einer Attribution von Ursachen auf das schulische Umfeld bzw. die Herkunftsmilieus der Schüler*innen führen. Da aber für die geringeren Bildungserfolge der Schüler*innen auch das Handeln der Schule maßgeblich ist (z. B. durch geringere Erwartungen und damit verknüpfte Handlungspraxen), kann eine verbesserte Wahrnehmung der eigenen Handlungsmöglichkeiten nur erreicht werden, wenn der Fokus auf die Veränderung dorthin verlagert wird. Eine zielorientierte Führung kann Veränderung also nur bewirken, wenn der Druck zur Anpassung nicht allein auf die Schüler*innen verlagert wird; das Ziel sollte auch normativ so gestaltet sein, dass Lehrkräfte ihre eigene Verantwortung reflektieren und ihr pädagogisches Handeln weiterentwickeln.

In zahlreichen internationalen und deutschsprachigen qualitativen Fallstudien erfolgreicher SsbS zeigt sich diesbezüglich eine gemeinsame Basis: Die Schulleitungen verfolgten meist eine Vision für ihre Schule, die die Schüler*innen als von der Gesellschaft marginalisiert begriff und das Ziel formulierte, ihnen zu ihrem Recht auf Bildung zu verhelfen. Dies bedeutete, dass die Schulen auch Verantwortung für den Abbau von Bildungsdisparitäten übernahmen, was u. a. durch Anerkennung der Lebenswelt der Schüler*innen und durch das Aufbrechen diskriminierender Praktiken gelang (Bronnert-Härle & Klein, 2025; Klein & Bronnert-Härle, 2022; Leithwood, 2021). Diese Visionen wurden in den Studien in der Regel durch systematische Strukturen unterstützt, die Lehrkräfte dazu befähigten, entsprechend zu handeln.

Insofern resümieren auch DeMatthews und Izquierdo (2020), dass kulturresponsive Führungspersonen von expliziten Werten und dem Engagement für marginalisierte Kinder geleitet sein müssen. Fehle diese normative Komponente, so laufe die Führungsperson Gefahr, Herausforderungen zu übersehen, benachteiligende Praktiken zu (re-)produzieren und dadurch Probleme noch zu verstärken (DeMatthews & Izquierdo, 2020).

Vor diesem Hintergrund wird im englischsprachigen Raum der Ansatz des *Culturally Responsive School Leadership* diskutiert, dessen Systematisierung großteils auf den Arbeiten von Khalifa et al. (2016) basiert. Anhand eines *Systematic Reviews* beschreiben Khalifa et al. (2016) vier relevante Handlungsdimensionen von KRSL, welche typische Kategorien etablierter Führungsansätze (z. B. Visionen, Unterstützung von Lehrkräften, Gestaltung der Organisation, vgl. Leithwood et al., 2006) auf die Frage zuspitzen, wann sie kulturresponsiv sind (vgl. Abbildung 1).

Abbildung 1: Dimensionen von KRSL



(adaptiert von Khalifa et al., 2016)

Kritische Reflexion der eigenen Orientierungen

Verschiedene qualitative Studien deuten darauf hin, dass Schulleitende ungünstige Perspektiven auf Schüler*innen haben können, die den Fokus ihres Handelns auf die Frage lenken, was (vermeintlich) an ihrer Schule *nicht* möglich ist (Bronnert-Härle & Klein, 2025; Drucks & Bremm, 2021; Racherbäumer, 2017). Weisen Schulleitende selbst Defizitorientierungen auf, dürfte es schwer sein, kulturreflexive Visionen und Ziele für die Schule zu formulieren und diese auch zu leben.

KRSL setzt insofern ein zielorientiertes Führungshandeln voraus, in dem die Führungsperson „critically self-reflects on leadership behaviors“ (Khalifa et al., 2016, S. 1283). Führungspersonen müssen sich ihrer eigenen Werte und Überzeugungen bewusst sein, um ein Bewusstsein für den Kontext zu entwickeln, innerhalb dessen sie agieren sowie auf dieser Basis eine Vision für eine kulturreflexive Schule entwickeln (Khalifa et al., 2016).

Kulturreflexive Professionalisierung und Unterrichtsentwicklung

Im deutschen Schulsystem sind weder Lehrpläne noch Lehrkräftebildung traditionell auf die Lebenswelt und Bedürfnisse marginalisierter Schüler*innen ausgerichtet, und angehende Lehrkräfte lernen kulturreflexives Handeln nicht als pädagogische Norm des Schulsystems kennen (z. B. Gottuck et al., 2021). Entsprechend sind Lehrkräfte häufig nicht auf Schüler*innen mit multiplen Scheiterns- und Diskriminierungserfahrungen vorbereitet, und erhalten durch Curricula und andere Rahmenvorgaben kaum Unterstützung.

Khalifa et al. (2016) betonen die Rolle von Führung für die kulturresponsive Professionalisierung der Lehrkräfte. Dies umfasst das Mainstreaming von Kulturresponsivität bei der Unterrichtsentwicklung, die Anregung für Lehrkräfte, ihr Handeln zu hinterfragen und eigene Überzeugungen zu reflektieren, sowie die Gestaltung von Strukturen und Angeboten, die es Lehrkräften ermöglichen, ihr eigenes Handeln kulturresponsiv weiterzuentwickeln.

Kulturresponsive Schulumgebungen

In Organisationen wird das kollektive Handeln der Menschen nicht nur durch die Überzeugungen einzelner Personen bestimmt, sondern insbesondere durch (bewusste oder unbewusste) Deutungen, die von den Organisationsmitgliedern geteilt werden (Schein, 2010). Sie entstehen, indem im gemeinsamen Handeln Deutungen des Handlungskontextes zwischen dem eigenen Idealbild von Schule und der Interpretation von Widersprüchen und Strukturproblemen entstehen und entscheiden darüber, welche Normen und Handlungspraxen sowohl aufseiten der Lehrkräfte, als auch bei den Schüler*innen als legitimiert wahrgenommen werden, und welche marginalisiert oder tabuisiert werden (Helsper, 2008; vgl. auch z. B. Helsper et al., 2017). An Schulen mit einem hohen Anteil an Schüler*innen aus marginalisierten Herkunftsmilieus finden diese Aushandlungsprozesse häufig vor dem Hintergrund von Entwicklungsdruck und einer sehr herausfordernden Aufgabe statt; insofern weisen qualitative Studien aus dem deutschen Sprachraum darauf hin, dass defizitorientierte Überzeugungen auch Bestandteil der geteilten Normen und Werte von SsBS sind (z. B. Drucks & Bremm, 2021). In einer Studie mit 945 Lehrkräften an 16 Schulen konnten Klein und Bronnert-Härle (2022) auch quantitativ aufzeigen, dass defizitorientierte Perspektiven auf die eigenen Schüler*innen in der Mehrheit der Schulen von einem Großteil der Lehrkräfte geteilt wurden.

Kulturresponsives Handeln setzt voraus, dass Handlungspraxen, die Schüler*innen aus armuts- oder migrantisch geprägten Herkunftsmilieus pathologisieren, abwerten oder marginalisieren, sowie die ihnen zugrundeliegenden kollektiven Deutungsmuster in der Schule identifiziert, expliziert und reflektiert werden. KRSL umfasst also eine eingehende Analyse der Schulumgebung und die Identifikation, Reflexion und Bearbeitung von Praktiken und Strukturen, welche die Marginalisierung von Schüler*innengruppen reproduzieren, und befördert im Gegenzug den Aufbau einer herkunftswertschätzenden, inklusiven Schulkultur (Khalifa et al., 2016).

Partizipation von Schüler*innen, Familien und der Community

Quantitative Befunde zu Defizitorientierungen gegenüber Schüler*innen im deutschsprachigen Raum (Klein & Bronnert-Härle, 2022; Pumptow et al., 2024) zeigen, dass Defizitorientierungen hier eher dem entsprechen, was Racherbäumer (2017) als Fürsorge bzw. Pathologisierung der Schüler*innen beschreibt; Distinktion und Abwertung der Schüler*innen zeigen sich dagegen nur in wenigen Schulen. Für die Bewertung der Eltern bzw. der Community der Schule liegen bislang keine vergleichbaren quantitativen Daten vor; die qualitativen Befunde lassen aber die Vermutung zu, dass diese vielfach eher in Richtung Distinktion und Abwertung tendiert – auch weil gegenüber den Eltern keine pädagogische Fürsorgepflicht besteht (z. B. Eiden, 2018; Kollender, 2022). Das soziale Umfeld der Schule wird von den schulischen Akteur*innen häufig als ursächlich für Probleme innerhalb der Schule wahrgenommen (z. B. Fölker et al., 2016). Entsprechend bestehen die Strategien von Schulen oftmals in einer Abschottung gegenüber den Familien bzw. dem Standort.

Kulturreponsives Führungshandeln umfasst nach Khalifa et al. (2016) die Fähigkeit, „to understand, address, and even advocate for community-based issues“ (S. 1282). Die Autor*innen beschreiben in ihrem Review einerseits Praktiken einer Partizipationskultur, die Schüler*innen und ihre Familien ernst nimmt, andererseits aber auch die systematische Öffnung der Schule gegenüber der lokalen Community.

3. Studiendesign

Für den Bildungserfolg von Schüler*innen aus marginalisierten Herkunftsmilieus ist es also zentral, dass die Pädagog*innen in Schule über kulturreponsive Handlungskompetenzen verfügen. Zugleich deuten Befunde aus SsbS darauf hin, dass Veränderungsprozesse hin zu kulturreponsivem Handeln die Schulen vor besondere Herausforderungen stellen. KRSL stellt eine Möglichkeit dar, diesen Herausforderungen zu begegnen; in der deutschsprachigen Schulleitungsforschung spielt der Ansatz bislang jedoch keine Rolle. Dementsprechend gibt es auch noch keine systematischen Analysen zu möglichen Handlungspraxen, die zu den Rahmenbedingungen des schulischen Führungshandelns in Deutschland passen, und entsprechend keine kontextadäquaten Wirkungsanalysen oder evidenzbasierten Handlungsempfehlungen.

Das Ziel des vorliegenden Beitrags besteht vor diesem Hintergrund zunächst in einer systematischen Zusammenfassung von KRSL-Handlungspraxen, sowohl von ‚good practice‘ als auch von ungünstigen Handlungspraxen. Im Anschluss sollen Bedingungen für die Übertragbarkeit dieser

Praxen auf den deutschsprachigen Kontext diskutiert werden. Der Beitrag versteht sich dabei als ‚Kartierungsversuch‘, der als Grundlage für nachfolgende Forschungs- und Entwicklungsprojekte dienen kann. Dazu bedient sich der Beitrag eines Systematic Literature Reviews (SLR) nach Hallinger (2013).

3.1. Datenquelle

In den Datenbanken ERIC, EBSCO, FIS Bildung, PubMed, Web of Science und ProQuest Educational Journal wurde zunächst mit der Verschlagwortung „Culturally Responsive School Leadership” und „Culturally Responsive Leadership” eine Liste aller hinterlegten Beiträge erstellt. Dabei wurden nur Beiträge berücksichtigt, die ab 2016 erschienen sind, da davon ausgegangen wurde, dass alle früheren Beiträge bereits in dem Literature Review von Khalifa et al. (2016) analysiert wurden.

Aus der Ergebnisliste wurden alle englisch- oder deutschsprachigen Beiträge ausgewählt, die (1) aus Peer Review Journals stammten, (2) bei denen es sich dem Abstract nach um empirische (qualitative oder quantitative) Originalbeiträge handelte, und (3) die einen inhaltlichen Bezug zu Handlungspraxen im Kontext von KRSL aufwiesen. Diese Beiträge wurden anschließend vollständig gelesen und dabei zunächst dahingehend geprüft, ob sie für das SLR relevant sind. Dabei führten drei Gründe zum Ausschluss von Beiträgen:

- Der Beitrag hatte keinen ausreichenden (mindestens impliziten) Bezug zum KRSL-Ansatz.
- Im Fokus standen nicht Handlungspraxen und deren Wirkungen, oder Handlungspraxen von anderen Führungspersonen (z. B. District Administration, Teacher Leaders) als der Schulleitung.
- Es handelte sich nicht um einen Originalbeitrag (z. B. Literature Reviews, keine eigene empirische Erhebung), es gab kein systematisches Erhebungsverfahren (z. B. Erfahrungsberichte einzelner Schulleitungen) oder das methodische Vorgehen wurde mangelhaft beschrieben.

Nach dieser ersten Analyse wurden wiederum weitere Beiträge ausgeschlossen. In das SLR wurden schließlich 30 Beiträge aufgenommen.

3.2. Analyse der Beiträge

Die Beiträge wurden in MAXQDA übertragen. Im ersten Teil des Kodiervorgangs wurden zunächst zwei Aspekte erfasst: Erstens, inwiefern KRSL implizit oder explizit als Rahmenmodell genannt wurde und mit welchen weiteren Theorieansätzen KRSL verknüpft wurde, und zweitens wurde das jeweilige methodische Vorgehen der Studien kodiert. In einem zweiten Teil des Kodiervorgangs wurden die Studienergebnisse nach den vier Handlungspraxen *Self-Awareness*, *Curricula and Teacher Preparation*, *School Environments*, *Engaging Students and Parents* kodiert. Eine letzte induktive Kodierung ergänzte weitere Handlungspraxen sowie Veränderungen aufseiten der Schüler*innen, Lehrpersonen und anderen Personengruppen.

4. Ergebnisse

Nachfolgend wird zunächst ein deskriptiver Überblick über die Beiträge gegeben, die in das Review eingegangen sind. Anschließend werden die darin berichteten Handlungspraxen entlang der vier Dimensionen des KRSL-Ansatzes zusammengefasst. Schließlich werden in den Studien berichtete Wirkungen von KRSL auf Lehrkräfte und Schüler*innen dargestellt.

4.1. Deskriptiver Überblick

Tabelle 1 bietet einen Überblick über die in die Analyse einbezogenen Beiträge. Jeweils eine Studie stammt aus Australien, Griechenland, Irland, Kanada, Korea, Malta, Neuseeland und den Philippinen. Alle anderen Studien berichteten Daten, die in den USA erhoben wurden.

Die methodische Ausrichtung der überwiegenden Zahl der Beiträge ist qualitativ; in fast allen dieser Studien kommen Interviews mit verschiedenen Akteursgruppen zum Einsatz, die insbesondere in Fallstudiendesigns durch weitere Methoden, z. B. Beobachtungsverfahren oder Analyse von Artefakten, ergänzt werden. Lediglich zwei Beiträge nutzen ausschließlich quantitative Erhebungsverfahren (Cager & Garibaldi, 2022; Ham et al., 2020); drei weitere Studien nutzen sowohl qualitative als auch quantitative Verfahren (Patterson & Wang, 2024; Shannon-Baker et al., 2020; Torres Clark & Chrispeels, 2022). Die meisten Beiträge berichten über Querschnittstudien; vereinzelt handelt es sich bei den Fallstudien aber auch um längsschnittliche Studien, in denen einzelne Schulen über mehrere Jahre systematisch begleitet wurden (z. B.

DeMatthews & Izquierdo, 2020) oder über einen längeren Zeitraum aktiv mit den Schulen gearbeitet wurde (z. B. Ylimaki & Brunderman, 2019).

Mit Blick auf die theoretische Rahmung ist KRSL in einem Teil der Studien die einzige Grundlage. In anderen Studien wird der Ansatz mit anderen Theorieperspektiven oder Konstrukten verknüpft. Verschiedene Autor*innen verbinden KRSL z. B. mit anderen Führungsansätzen (z. B. Guo-Brennan & Guo-Brennan, 2021; Oliver, 2023; Patterson & Wang, 2024; Torres Clark & Chrispeels, 2022), pädagogischen Zugängen (z. B. DeMatthews & Izquierdo, 2020; Kayser et al., 2020; Mastrogiovanni & Freidus, 2023; Ylimaki & Brunderman, 2019) sowie rassismuskritischen Perspektiven (z. B. Burgess et al., 2023; Jayavant, 2016; Rodela & Rodriguez-Mojica, 2020).

Tabelle 1: Überblick über die berücksichtigten Beiträge

Autor*innen	Journal	Kontext	Methodik
Brooks & Brooks (2019)	International Journal of Leadership in Education	Philippinen	qualitativ
Brown et al. (2019)	Educational Management Administration & Leadership	Irland	qualitativ
Burgess et al. (2023)	The Australian Educational Researcher	Australien	qualitativ
Cager & Garibaldi (2022)	The Journal of Negro Education	USA	quantitativ
Cai (2023)	Teachers College Record	USA	qualitativ
Callahan et al. (2023)	Leadership and Policy in Schools	USA	qualitativ
Cotman et al. (2020)	Journal of Cases in Educational Leadership	USA	qualitativ
Crawford et al. (2022)	Journal of School Leadership	USA	qualitativ
DeMatthews & Izquierdo (2020)	Urban Education	USA	qualitativ
Ezzani (2021)	Journal of School Leadership	USA	qualitativ
Fletcher et al. (2019)	The Qualitative Report	USA	qualitativ
Gordon & Ronder (2016)	International Journal of Educational Reform	USA	qualitativ
Guo-Brennan & Guo-Brennan (2021)	Leadership and Policy in Schools	Kanada	qualitativ
Ham et al. (2020)	Multicultural Education Review	Korea	quantitativ
Hesbol et al. (2020)	Journal on Transforming Professional Practice	USA	qualitativ
Jayavant (2016)	Education Sciences	Neuseeland	qualitativ
Jones et al. (2019)	Journal of School Leadership	USA	qualitativ
Kayser et al. (2020)	Chapman University Digital Commons	USA	qualitativ
Mastrogiovanni & Freidus (2023)	Journal of Research on Leadership Education	USA	qualitativ
Narine et al. (2022)	Journal of Instructional Pedagogies	USA	qualitativ
Oliver (2023)	Journal of Education	USA	qualitativ
Patterson & Wang (2024)	School Leadership & Management	USA	Mixed Methods
Rodela & Rodriguez-Mojica (2020)	Educational Administration Quarterly	USA	qualitativ
Savvopoulos et al. (2024)	Journal of Educational Administration & History	Griechenland	qualitativ
Scribner et al. (2021)	NASSP Bulletin	USA	qualitativ
Shannon-Baker et al. (2020)	Journal of Educational Foundations	USA	Mixed Methods
Thompson et al. (2023)	Journal of Urban Learning, Teaching, and Research	USA	qualitativ
Torres Clark & Chrispeels (2022)	Journal of Educational Administration	USA	Mixed Methods
Vassallo (2022)	Improving Schools	Malta	qualitativ
Ylimaki & Brunderman (2019)	Education Sciences	USA	qualitativ

4.2. Handlungspraxen

Kritische Reflexion der eigenen Orientierungen

Diese Kategorie beschreibt Khalifa et al. (2016) als „an awareness of self and his/her values, beliefs, and/or dispositions“ (S. 1280). Die meisten Studien adressieren die Fähigkeit zur kritischen Selbstreflexion: Die beforschten Schulleitenden stellten durch einen klaren Wertebezug Vielfalt als Ressource in den Vordergrund; diese Perspektive erhöhte in den Studien das Selbstbewusstsein von Schulleitenden und Lehrkräften sowie ihre Motivation bei der Umsetzung von KRSL und kultursensibler Pädagogik (Brown et al., 2019; Gordon & Ronder, 2016; Jayavant, 2016; Narine et al., 2022; Patterson & Wang, 2024; Scribner et al., 2021). Studienergebnisse verweisen auf die Wichtigkeit, dass Schulleitende diese Werthaltung gegen Widerstände auf Verwaltungs-, Schul- und Personenebene aufrechterhalten (Brooks & Brooks, 2019; Brown et al., 2019; Fletcher et al., 2019; Rodela & Rodriguez-Mojica, 2020) und durch gezielte Maßnahmen der kritischen Reflexion von sozio-kulturellen Problemlagen und dem eigenen pädagogischen Handeln fördern. Dadurch können Vorurteile und potenziell ungünstiges Verhalten bewusst gemacht und abgebaut werden (Cager & Garibaldi, 2022; Crawford et al., 2022; Kayser et al., 2020; Mastrogiovanni & Freidus, 2023; Patterson & Wang, 2024).

Die berichteten Befunde deuten an, dass insbesondere Führungs- und Lehrkräfte, die selbst unterrepräsentierten Gruppen angehören, mehr Verständnis und Bewusstsein für sozio-kulturelle Herausforderungen aufweisen (DeMatthews & Izquierdo, 2020; Fletcher et al., 2019; Thompson et al., 2023) und sich eher in die Schüler*innen hineinversetzen und als Vorbilder dienen können (Cai, 2023; Rodela & Rodriguez-Mojica, 2020; Thompson et al., 2023; Vassallo, 2022). Kulturelle Orientierungen und deren Wirkung nicht zu reflektieren, birgt dagegen die Gefahr, auch im Umgang mit Schüler*innen diese Reflexion zu vernachlässigen (Cotman et al., 2020).

Kulturreresponsive Professionalisierung und Unterrichtsentwicklung

Insbesondere die Veränderung von Überzeugungen und die Einflussnahme auf das pädagogische Handeln von Lehrkräften wird von Schulleitenden vielfach als besonders schwierig wahrgenommen. In den Studien bestanden Führungspersonen vielfach darauf, dass Lehrkräfte ihre Überzeugungen und Handeln hinterfragten, und verbanden ihre Forderung im Zweifel auch mit Konsequenzen für Lehrkräfte, die an ungünstigen Handlungspraxen festhielten (DeMatthews & Izquierdo, 2020).

Um Lehrkräfte zu erreichen und ihnen KRSL in konkreter Anwendung auf die eigene Schule näherzubringen, nutzten Schulleitende z. B. praktisches Anschauungsmaterial (Rodela & Rodriguez-Mojica, 2020). Schulleitende förderten kulturelle Kompetenzen der Lehrkräfte durch Weiterbildung, individuelle Unterstützung und informelle (Führungs-)Gespräche (Fletcher et al., 2019; Guo-Brennan & Guo-Brennan, 2021; Oliver, 2023; Scribner et al., 2021), Nutzung wissenschaftlich fundierter Informationen zu Effekten sozio-kultureller Differenzen (Hesbol et al., 2020; Vassallo, 2022), und Anleitung zur Entwicklung einer selbstreflexiven Grundhaltung (Crawford et al., 2022; Kayser et al., 2020; Mastrogiovanni & Freidus, 2023; Patterson & Wang, 2024; Vassallo, 2022). Schulleitende unterstützten z. B. die Nutzung kultursensibler Lehrbücher sowie non-verbaler Kommunikation in Kontexten sprachlicher Diversität, um ein kulturresponsives Unterrichtshandeln der Lehrkräfte zu befördern (Brown et al., 2019; Narine et al., 2022).

Deutlich wird zudem gerade in dieser Kategorie die Bedeutung systematischen Führungshandelns. So verweisen die Studien darauf, dass Strategien einer transformationalen, offenen und selbstkritischen Führung sowie die Verteilung von Führung ein kulturresponsives Handeln der Lehrkräfte begünstigte (Cai, 2023; Callahan et al., 2023; DeMatthews & Izquierdo, 2020; Oliver, 2023; Scribner et al., 2021; Torres Clark & Chrispeels, 2022; Vassallo, 2022). Auch Freiräume für die Lehrkräfte wirkten sich positiv aus (Brown et al., 2019; Burgess et al., 2023; Crawford et al., 2022; Guo-Brennan & Guo-Brennan, 2021; Mastrogiovanni & Freidus, 2023; Oliver, 2023; Thompson et al., 2023; Vassallo, 2022). Gleichwohl zeigen die Studien auch, dass eine reflektierende, kulturell offene Grundhaltung nicht von allen Schulleitungen vorgelebt wird und manche Weiterbildungsangebote die Bedarfe der Lehrenden nicht treffen (Brooks & Brooks, 2019; Guo-Brennan & Guo-Brennan, 2021; Savvopoulos et al., 2024).

Daneben verweisen einzelne Studien auch auf eine gezielte Personalplanung durch die Einstellung und Förderung eines diversen Kollegiums (Cai, 2023; Oliver, 2023). Schulleitende nutzten zudem Unterrichtsbeobachtungen und Gespräche mit Schüler*innen für die Beurteilung von Lehrkräften (Scribner et al., 2021).

Kulturresponsive Schulumgebungen

Zur Identifizierung (potenziell) benachteiligender Strukturen und Praktiken auf organisationaler Ebene werden von den Schulleitenden z. B. Schüler*innen- und Leistungsdaten, Unterrichtsbeobachtungen oder Gespräche mit Lehrkräften, Eltern, Sozialarbeiter*innen, Klassenteams etc. herangezogen

(DeMatthews & Izquierdo, 2020; Vassallo, 2022; Ylimaki & Brunderman, 2019). DeMatthews und Izquierdo (2020) beschreiben in ihrer Fallstudie auch das Vorgehen einer Schulleiterin, welche neben den genannten Maßnahmen ein interdisziplinäres Team aus Mitarbeiter*innen etablierte, das evaluierte, ob die schulischen Ressourcen im Sinne der Förderung von Schüler*innen und ihrer Familien genutzt wurden.

Wurden benachteiligende Strukturen oder Praktiken evident, wurden diese im Rahmen von KRSL explizit adressiert. Beschriebene Maßnahmen umfassten einen lösungsorientierten Austausch ohne Schuldzuweisung, die Thematisierung im Rahmen regelmäßiger Treffen bis hin zu (rassismus-)kritischen Diskursen oder gar konflikthafte Gespräche mit Lehrkräften über mögliche Defizitperspektiven und fehlendes *Commitment* zu kulturresponsiven Schulzielen (DeMatthews & Izquierdo, 2020; Vassallo, 2022; Ylimaki & Brunderman, 2019). Die analysierten Führungspraxen fokussierten das Herstellen und Befördern von Reflexionsanlässen, um Ursachen marginalisierender Strukturen oder Praktiken und die Bedürfnisse und Potenziale der Schüler*innen zu verstehen sowie das Bewusstsein um besondere Herausforderungen für bestimmte Schüler*innengruppen zu schärfen oder individuelle sowie kollektive Werte, Überzeugungen und Vorurteile zu reflektieren (Cai, 2023; DeMatthews & Izquierdo, 2020; Guo-Brennan & Guo-Brennan, 2021; Thompson et al., 2023; Vassallo, 2022; Ylimaki & Brunderman, 2019). Neben individuellen Gesprächen sowie Sitzungen mit unterschiedlichen schulischen Akteursgruppen wurden auch *Professional Learning Communities* als Reflexionsgelegenheit genutzt (DeMatthews & Izquierdo, 2020; Ylimaki & Brunderman, 2019).

Die Auflösung benachteiligender Praktiken und Strukturen ist zumeist mit der Etablierung korrektiver oder kompensatorischer Maßnahmen verbunden, etwa der Schaffung neuer oder inklusiverer (außer-) unterrichtlicher Angebote (Jayavant, 2016; Oliver, 2023; Scribner et al., 2021). Für den Aufbau einer kulturresponsiven Schulumgebung erscheinen zudem transformationale Führungspraktiken bedeutsam. Die einbezogenen Studien betonen die Kommunikation und Etablierung einer geteilten Vision, welche Vielfalt, Wertschätzung, Gleichberechtigung und Inklusion sowie die Förderung sozialer Mobilität in den Vordergrund stellt und die Gestaltung schulischer Strukturen und Prozesse sowie den Ressourceneinsatz prägt (Cai, 2023; Guo-Brennan & Guo-Brennan, 2021; Oliver, 2023; Rodela & Rodriguez-Mojica, 2020). Indem die Schulleitenden kulturresponsive Werte in ihrem Handeln priorisierten sowie aktiv vorlebten, fungierten sie als Vorbilder, die Selbiges von der Schulgemeinschaft einfordern (Brown et al., 2019; Narine et al., 2022). Hierzu gehört auch, dass Schulleitende kulturresponsive Ziele nicht nur konsequent

schulintern, sondern auch gegenüber externen Anspruchsgruppen, etwa der Schulverwaltung, verteidigten (Rodela-Rodriguez-Mojica, 2020; Vassallo, 2022; DeMatthews & Izquierdo, 2020). Dass die Etablierung einer kulturre-sponsiven Schulumgebung durch schwierige sozialpolitische Rahmenbedin-gungen, einen häufigen Schulleitungswechsel und die Schulgröße erschwert oder gar verhindert werden kann, zeigt die philippinische Studie von Brooks und Brooks (2019).

Partizipation von Schüler*innen, Familien und der Community

Die Grundlage für die aktive Einbindung der Schüler*innen, ihrer Fami-lien sowie des schulischen Sozialraums bildet in den meisten Studien eine potenzialorientierte Einstellung der Schulleitenden: Herkunft, Sprache und bisherige Erfahrungen wurden bewusst als Stärke bzw. Mehrwert betrachtet und entsprechend gefördert, wertgeschätzt sowie in den Schulalltag inte-griert (Burgess et al., 2023; Guo-Brennan & Guo-Brennan, 2021; Oliver, 2023). Die Schulleitenden zeigten Verständnis, Empathie und eine hohe Einsatzbe-reitschaft für ihre Schüler*innen und deren Familien und sahen sich in der Verantwortung, Barrieren einer aktiven Einbindung in den Schulalltag zu identifizieren und abzubauen (DeMatthews & Izquierdo, 2020; Fletcher et al., 2019; Vassallo, 2022). Das kulturre-sponsive Führungshandeln war geprägt vom Aufbau vertrauensvoller Beziehungen zwischen Schulleitung, Schü-ler*innen und Familien sowie der Community (Rodela & Rodriguez-Mojica, 2020). Burgess et al. (2023) betonen in ihrer Studie, dass positive Beziehungen zu Schüler*innen den Beziehungsaufbau mit den Familien erleichtern können.

Mehrere Studien lassen Führungspraktiken erkennen, die mit der Inten-tion verbunden sind, durch den direkten Kontakt (z. B. Gespräche mit Eltern, Beförderung von Hausbesuchen durch Lehrkräfte, Kontakte mit Vertreter*in-nen der Community) mehr über die Schüler*innen, ihre Familien sowie (kul-turellen) Gemeinschaften zu erfahren, um daraus Strategien zur Unterstüt-zung abzuleiten (z. B. Ezzani, 2021; Hesbol et al., 2020).

Ein intensiver Kontakt zwischen schulischen Akteuren und Familien kann überdies helfen, den gegenseitigen Respekt, das Verständnis, gemeinsame Ziele sowie das Infragestellen vorurteilsbelasteter Annahmen zu befördern (Hesbol et al., 2020; Mastrogiovanni & Freidus, 2023; Thompson et al., 2023). Burgess et al. (2023) zeigen, dass sich vertiefte Kenntnisse der Schulleitung über die lokale Community in bedeutungsvolle kulturelle, soziale und lern-bezogene Aktivitäten transferieren ließ.

Kulturre-sponsive Schulleitende passten ihre Praktiken an die Bedürfnisse und Voraussetzungen der Schüler*innen und Eltern an, indem sie sprachli-

che, zeitliche, kulturelle oder hierarchische Barrieren abbauten. Dies erfolgte z. B. durch Dolmetscher*innen, direkte Ansprache, Nutzung von Vornamen, Angebote von Gesprächen in Gruppen statt Einzelgesprächen, eine Politik der „offenen Tür“ oder regelmäßige formelle sowie informelle Gesprächsanlässe zu verschiedenen Tageszeiten (Brown et al., 2019; Burgess et al., 2023; Guo-Brennan & Guo-Brennan, 2021; Mastrogiovanni & Freidus, 2023; Savopoulos et al., 2024).

Die Studien zeigen, dass Schulleitende Beratung und Unterstützung entweder durch schulische Ressourcen und Personal ermöglichten oder Eltern halfen, (bildungs-)relevante Ressourcen aus der Community zu erhalten (z. B. Sprachkurse, Rechtsberatung, medizinische Leistungen; Crawford et al., 2022; Rodela & Rodriguez-Mojica, 2020; Shannon-Baker et al., 2020). Ermöglicht wurde dies über den Kontaktaufbau der Schulleitung zu relevanten Einrichtungen und Akteuren aus dem schulischen Sozialraum und deren Einbezug bei schulischen Entwicklungen (Hesbol et al., 2020; Jones et al., 2019; Narine et al., 2022). In der Studie von DeMatthews & Izquierdo (2020) verstand die Schulleiterin es gar als die zentrale Aufgabe schulischer Führungskräfte und Akteure, „for connecting families with the resources they need to be fully engaged parents actively supporting their children’s education“ (S. 384).

4.3. Wirkungen auf Lehrkräfte und Schüler*innen

Quantitative Wirkungsstudien mit Blick auf den Bildungserfolg von Schüler*innen liegen aktuell zu KRSL nicht vor. In den Studien zeichnet sich aber ab, dass KRSL das Potenzial hat, professionelles Wissen, Überzeugungen, motivationale Orientierungen und Handlungspraxen von Lehrkräften zu beeinflussen. So verweisen die Studien etwa auf Kooperationsanlässe, die durch KRSL geschaffen wurden, z. B. mit Blick auf die gemeinsame Vorbereitung von Lehrpersonen und Schüler*innen auf deren Collegebewerbungen (Fletcher et al, 2019), Planung von bilinguaem Unterricht (DeMatthews & Izquierdo, 2020), Teamplanungssitzungen zu neuen Strategien oder in Fortbildungen zu kulturresponsiver Unterrichtsentwicklung (Burgess et al., 2023; Ezzani, 2021). Auch mit Blick auf die motivationale Orientierung der Lehrkräfte deuteten die Studien auf Veränderungen hin: im Rahmen einer KRSL-Initiative berichteten Lehrpersonen von einer gemeinsamen Anstrengung und sehr absichtsvollem, hohem Aufwand:

There was, however, a concerted effort to close achievement and opportunity gaps beyond performances on standardized assessments. One participant explained: „Every day, I am intentionally doing various things, various strategies, using curriculum outside my resources. I think about who I’m teaching in order to reach my students’ needs to grow them from dependent learners into independent learning today. But it’s extremely intentional and it’s a lot of effort. It’s not extra effort, but it’s definitely a lot of effort. (Kayser et al., 2020, S. 18)

Ähnlich erfolgreich nahmen Teilnehmende eine Lehrkräftefortbildung als Möglichkeit wahr, „to develop a shared understanding of cultural responsiveness and EL equity“ (Callahan et al., 2023, S. 115), was gleichwohl durch die Absenz dreier Schulleitungsmitglieder und die als daraus resultierend berichtete Diskrepanz zwischen den neu erarbeiteten Zielen des Kollegiums und der Schulleitung zunichte gemacht worden sei (Callahan et al., 2023). In einer weiteren Fallstudie wird dargestellt, wie die KRSL-Praktiken eines Schulleiters im Kollegium sichtbar und erlebbar wurden: „demonstrating concern and care for students laid a foundation of trust that allowed staff to identify issues in students’ lives that might undermine student engagement with the school community“ (Scribner et al., 2021, S. 159), die das Kollegium ermutigte, dem Vorbild der Schulleitung zu folgen und Regeln abzuschaffen, die als den Bedarfen dieser Schüler*innen nicht angemessen identifiziert wurden (ibid., S. 159). Besprechungen wurden in der Wahrnehmung der Schulmitglieder effizienter und kritischer angelegt und durchgeführt; im Ergebnis wurde das Engagement der Schulleitung für den Wandel von den Lehrkräften unmittelbar erlebt und dadurch mitgetragen: „The principal expended considerable energy reinforcing a school-wide commitment to cultural responsiveness. (...) This commitment to transformation was echoed by teachers“ (ibid., S. 164).

Ham et al. (2020) können in einer koreanischen Studie auch anhand quantitativer Daten einen positiven Zusammenhang zwischen KRSL und motivationalen Orientierungen der Lehrkräfte beschreiben. Durch Veränderungen im professionellen Lernen von Lehrkräften deuteten vereinzelte Studien auch auf Erfolge in der Reduktion von Leistungsunterschieden hin (Rodela & Rodriguez-Mojica, 2020).

Insgesamt sind die Forschungsergebnisse gleichwohl ambivalent. So zeigte sich, dass das Handeln der Lehrkräfte in kritischen Situationen wieder auf ‚Gewohntes‘ zurückfiel; statt Selbstreflexion, Unterrichtsentwicklung und Zusammenarbeit wurde in solchen Situationen Testvorbereitung diskutiert und negativ über die Schulleitung gesprochen (Callahan et al., 2023; Ylimaki & Brunderman, 2019).

Mit Blick auf Wirkungen auf die Schüler*innen verweisen Studien, in denen Schulleitende Praktiken des KRSL ausüben, auf eine hohe Resilienz sowie Schulmotivation der Schüler*innen, eine friedliche Schulkultur, gute Beziehungen zwischen den schulischen Akteuren und den Schüler*innen bzw. ihren Familien sowie sehr gute akademische Leistungen und Übergangsquoten an Hochschulen (Ezzani, 2021; Fletcher et al., 2019; Rodela & Rodriguez-Mojica, 2020; Scribner et al., 2021). Ein direkter Bezug lässt sich aufgrund der Studiendesigns jedoch zumeist nicht herstellen. So wurden häufig jene Schulen einbezogen, die sich bereits zur Stichprobenziehung durch besondere akademische Leistungen der Schüler*innenschaft oder gute Übergangsquoten an Hochschulen (als good practice-Beispiel) auszeichneten (Fletcher et al., 2019; Rodela & Rodriguez-Mojica, 2020). Eine Interventionsstudie von Ylimaki und Brunderman (2019), in der 71 Schulleitungsteams kulturreresponsive Reflexions- und Handlungspraxen erlernten sowie erprobten, kann positive Entwicklungen hinsichtlich der formativen Beurteilung der Schule feststellen. Aufgrund einer fehlenden Kontrollgruppe sind Kausalaussagen jedoch auch hier nur eingeschränkt möglich.

5. Diskussion

Das vorangegangene Review beschreibt systematische Strategien des KRSL an Schulen an benachteiligten Standorten. Nachfolgend wird diskutiert, welche Implikationen sich für eine Übertragung des Ansatzes auf den deutschen Kontext ergeben.

Die Komplexität der Handlungspraxen verdeutlicht, dass KRSL anspruchsvoll ist und Rahmenbedingungen erfordert, in denen die Schulleitenden spezifische Kompetenzen entwickeln und ausbauen können. Es wird erkennbar, dass Schulleitende Kompetenzen in der Personalführung sowie -entwicklung besitzen müssen, ebenso wie ein klares Verständnis der schulorganisatorischen Strukturen und Ziele. Auch scheint Wissen um Deutungs- und Bewertungsmuster sowie konkrete Strategien, z. B. Vorbildhandeln, in der eigenen Schule notwendig, die benachteiligende Praktiken systematisch identifizieren und beheben. Ein kritischer Blick auf Schule und ihre Einbettung im Schulsystem, als auch ein kritischer Blick auf das eigene Handeln innerhalb desselben scheinen unabdinglich. Vor dem Hintergrund, dass deutsche Schulleitende sich häufig nicht als Führungspersonen verstehen (Warwas, 2012) und (noch) seltener strategische Schulentwicklung betreiben (Klein, Bronnert-Härle, Boone, & Muslic, 2022), erscheinen die beschriebenen Herangehensweisen aus dem Review einerseits wesentlich und innovativ, andererseits auch kulturell ent-

fernt und von einem anderen Verständnis geprägt. Eine daraus sich ergebende These kann sein, dass KRSL für deutsche Schulleitende gewinnbringend und gleichsam herausfordernd wirken könnte. Folgt man der These, so gälte es, Schulleitende dabei zu unterstützen, sich KRSL zu nähern und die Ansätze in das eigene Handeln aufzunehmen. Das bedürfe einer systematischen Ausbildung des dafür notwendigen Wissens und der Kompetenzen sowie wirksamer Unterstützungsstrukturen in der pädagogisch-praktischen Anwendung von KRSL auf allen Ebenen der Schulleitungsverantwortung. Daraus ergibt sich die Frage, wie KRSL an Schulen etabliert werden könnte, sollten Schulleitende weder seine Notwendigkeit, noch seinen Mehrwert erkennen. Forschung aus den USA weist darauf hin, dass Akteure aus der Schulverwaltung, insbesondere aus der Schulaufsicht, eine wichtige Rolle als *Change Agents* spielen können, z. B. durch das Setzen von Zielen, das Anstoßen schwieriger Diskurse, die Verdeutlichung der Relevanz von KRSL und die Unterstützung der Etablierung kulturresponsiver Praktiken (Aguayo et al., 2023; DeMatthews et al., 2017; Maxwell et al., 2013). Entsprechend widmen sich diverse Studien auch der Frage, wie KRSL von Führungspersonen der Schulaufsicht ausgeführt werden kann (z. B. Banwo et al., 2022). Für die Übertragung von KRSL auf den deutschen Sprachraum kann daher als These folgen:

These 1: Ein KRSL-Ansatz für den deutschen Sprachraum darf sich nicht alleine auf Schulleitung konzentrieren, sondern muss systematisch die Rolle von Ausbildungs- und Unterstützungsstrukturen für Schulleitende mitdenken.

Neben den Kompetenzen der Schulleitenden werden deren zeitliche Ressourcen einerseits, und ihre Handlungsmacht andererseits bedeutsam. Schulleitende in den USA haben „the positional authority to disrupt, dislocate, and destabilize asymmetrical power relations and other marginalizing conditions within schools“ (DeMatthews & Izquierdo, 2020, S. 363). In den Studien werden zwar viele Handlungspraxen beschrieben, die sich als mikropolitische Führungspraxen (Gieske, 2013) fassen lassen und ohne Verfügungsmacht auskommen. Die letztliche Konsequenz, mit der die Schulleitungen in den Studien die Vision kulturresponsiven Handelns verfolgen, und zwar notfalls auch am Willen der Lehrkräfte vorbei, lässt sich aber nicht ohne Weiteres auf eine Schulleitung im „Machtverdikt“ (Jahn, 2017) deutscher Schulen übertragen. In den Studien aus dem US-Kontext finden sich verschiedene Hinweise auf die Bedeutung verteilter Führung; dies dürfte umso relevanter für den deutschen Kontext sein. Hieraus ergibt sich für eine Übertragung des Ansatzes auf den deutschen Sprachraum folgende These:

These 2: Ein KRSL-Ansatz für den deutschen Sprachraum darf sich nicht alleine auf Schulleitungen konzentrieren, sondern muss systematisch die Rolle von Lehrkräften mit Führungsaufgaben und die Bedeutung flacher Hierarchien mitdenken.

Zuletzt ergeben sich auch mit Blick auf das Schulsystem insgesamt bedeutende Hinweise für eine Übertragung des Ansatzes auf den deutschen Raum. Der Großteil der Studien zu KRSL stammt aus Ländern, die kein gegliedertes Schulsystem haben. Im Gegensatz dazu gibt es im deutschen Schulsystem in der weiterführenden Schule eine systemisch angelegte strukturelle Segregation, die historisch aus einer sozialen Trennung entstanden ist und auch weiterhin entlang sozialer Kategorien verläuft, aber trotz gegenläufiger Evidenz durch den Verweis auf das meritokratische Prinzip legitimiert wird. Das System stuft Schüler*innen in den unteren Bildungsgängen also selbst offiziell als ‚weniger begabt‘ ein und es gibt empirische Hinweise darauf, dass sich diese Kategorisierung auch in den Überzeugungen der Lehrkräfte widerspiegelt (Bronnert-Härle & Klein, 2025; Wellgraf, 2012). Insofern ergibt sich für KRSL im deutschen Sprachraum die folgende These zur Herausforderung:

These 3: Ein KRSL-Ansatz für den deutschen Sprachraum muss berücksichtigen, dass Widerstand gegen kulturre sponsives Handeln auch aus einer schulsystembedingten Legitimierung der Marginalisierung bestimmter Schüler*innengruppen im gegliederten Schulsystem speist, und entsprechende Vorschläge dafür machen, wie dieser Widerstand bei gleichbleibenden Systemzwängen bearbeitet werden kann.

In der Zusammenschau des Reviews und der vorgelegten Thesen wird deutlich, dass sich der Mehrwert von KRSL für den deutschen Raum sowohl als Ansatz als auch als Handlungsprogramm auf mehreren Ebenen darstellt: Die konzeptuelle Fundierung von KRSL greift jene Phänomene auf, die im deutschen Diskurs unter ‚Chancengleichheit‘ und Benachteiligung diskutiert werden, und tut dies mit einem ressourcenorientierten, pädagogisch-positiven Blick auf alle Akteursgruppen in Schule. Eine solche Entwicklung scheint angemessen und schlusslogisch vor dem Hintergrund der seit 20 Jahren gleichbleibenden empirischen Ergebnisse aus PISA und anderen Leistungsstudien. Die Bedingungen von akademischer Leistung in den Vordergrund zu stellen, ist der zentrale Gewinn von KRSL, und der die grundlegende Wendung von individuellen Leistungsunterschieden auf die Lernenden als Subjekte in ihrer Lebenswelt vollzieht.

6. Fazit

Angesichts der Tatsache, dass für den deutschsprachigen Raum die systematische Bildungsbenachteiligung bestimmter Schüler*innengruppen bereits seit Jahrzehnten empirisch nachgewiesen werden kann, ist es bemerkenswert, dass kulturresponsive Ansätze sowohl in der deutschen Bildungspraxis wie auch -forschung bislang kaum Beachtung erfahren. Studien, die die Wirkung kulturresponsiver Führungspraxen in den Blick nehmen, geben erste Hinweise darauf, dass sich diese positiv auf die Schulkultur, schulische Interaktionen und die Motivation marginalisierter Schüler*innengruppen auswirken können. Es ist daher anzunehmen, dass kulturresponsive Handlungspraktiken einen wichtigen Ansatzpunkt zum Abbau systematischer Benachteiligungen im Bildungssystem leisten können, wenn sie sowohl auf Ebene der Einzelschule, wie auch im gesamten Mehrebenensystem priorisiert und umgesetzt werden. Eine Übertragbarkeit auf den deutschsprachigen Raum kann unter Berücksichtigung der genannten Thesen erfolgen, bedarf allerdings weiterführender kontextsensibler Forschung, die sowohl Handlungsbedingungen als auch Wirkungen kulturresponsiver Führungspraxen systematisch in den Blick nimmt.

Literatur

- Au, K. (2015). Isn't culturally responsive instruction just good teaching? In W. C. Parker, (Ed.), *Social Studies Today: Research and Practice* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315726885>
- Aguayo, D., Good, M. W., Diem, S., Herman, K. C., Burke, J., Davis, T., Hall, K., London, C., & Reinke, W. M. (2023). Promoting District-Level Culturally Responsive Practices. *Educational Administration Quarterly*, 59(3), 471–506. <https://doi.org/10.1177/0013161X231161041>
- Aikins, M. A., Bremberger, T., Aikins, J. K., Gyamerah, D., & Yıldırım-Calıman, D. (2021). *Afrozensus 2020: Perspektiven, Anti-Schwarze Rassismuserfahrungen und Engagement Schwarzer, afrikanischer und afrodiasporischer Menschen in Deutschland* [Afrozensus 2020: Perspectives, anti-Black racism experiences and involvement of Black, African and Afro-diasporic people in Germany]. Each One Teach One (EOTO) e.V.
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, 28(2), 117–148. https://doi.org/10.1207/s15326985ep2802_3
- Banwo, B. O., Khalifa, M., & Seashore Louis, K. (2022). Exploring trust: culturally responsive and positive school leadership. *Journal of Educational Administration*, 60(3), 323–339. <https://doi.org/10.1108/JEA-03-2021-0065>
- Beckett, L., & Wrigley, T. (2014). Overcoming stereotypes, discovering hidden capitals. *Improving Schools*, 17(3), 217–230. <https://doi.org/10.1177/1365480214556419>
- Bello, B. (2020). *Diversitätsbezogene Einstellungen von Lehramtsstudierenden. Eine quantitative Untersuchung* [Diversity-related attitudes of preservice teachers. A quantitative study]. Springer VS.

- Bonefeld, M., & Dickhäuser, O. (2018). (Biased) Grading of Students' Performance: Students' Names, Performance Level, and Implicit Attitudes. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00481>
- Bronnert-Härle, H. L., & Klein, E. D. (2025). Leading with an Equity Vision as a Driver of Instructional Change at Schools Serving Marginalized Communities? Evidence from Four German Schools. In J. Hugo, C. Assmann, A. Schmidt, & D. Sehmisch, D (Eds.), *Educational Leadership and Children's Rights? Kinderrechte als Führungsaufgabe?* (pp. 23–40). Waxmann.
- Brooks, M. C., & Brooks, J. S. (2019). Culturally (ir)relevant school leadership: Ethno-religious conflict and school administration in the Philippines. *International Journal of Leadership in Education*, 22(1), 6–29. <https://doi.org/10.1080/13603124.2018.1503819>
- Brown, M., McNamara, G., O'Hara, J., Hood, S., Burns, D., & Kurum, G. (2019). Evaluating the impact of distributed culturally responsive leadership in a disadvantaged rural primary school in Ireland. *Educational Management Administration & Leadership*, 47(3), 457–474. <https://doi.org/10.1177/1741143217739360>
- Burgess, C., Fricker, A., & Weuffen, S. (2023). Lessons to learn, discourses to change, relationships to build: How Decolonising Race Theory can articulate the interface between school leadership and Aboriginal students' schooling experiences. *The Australian Educational Researcher*, 50(1), 111–129. <https://doi.org/10.1007/s13384-022-00546-z>
- Cadenas, G. A., Cisneros, J., Spanierman, L. B., Yi, J., & Todd, N. R. (2021). Detrimental Effects of Color-Blind Racial Attitudes in Preparing a Culturally Responsive Teaching Workforce for Immigrants. *Journal of Career Development*, 48(6), 926–941. <https://doi.org/10.1177/0894845320903380>
- Cager, B. E., & Garibaldi, A. M. (2022). Reviving the Post-Pandemic Achievement of African American Students Through Culturally Responsive Leadership. *Journal of Negro Education*, 91(3), 397–407.
- Cai, L. (2023). An Asian American Feminist Manifesto: Asian American Women Heads of Schools Embodying Culturally Responsive School Leadership. *Teachers College Record*, 125(7–8), 173–187. <https://doi.org/10.1177/01614681231209589>
- Callahan, R. M., Mayer, A. P., Johnson, A. H., & Ochoa, C. J. (2023). Exploring Organizational Leadership for English Learner Equity. *Leadership and Policy in Schools*, 22(1), 99–118. <https://doi.org/10.1080/15700763.2021.1916539>
- Civitillo, S., Juang, L. P., & Schachner, M. K. (2021). Stressing similarities or ignoring differences? Shedding light into different forms of color-evasive ideology with pre- and in-service teachers. *Zeitschrift für Erziehungswissenschaft*, 24(1), 135–153. <https://doi.org/10.1007/s11618-021-00995-9>
- Cotman, A. M., Enyioha, J. C., Guerra, P. L., & Holmes, A. L. (2020). A Counselor's Tale: The Need for Cultural Awareness. *Journal of Cases in Educational Leadership*, 23(2), 3–18. <https://doi.org/10.1177/1555458919884714>
- Crawford, E. R., Aguayo, D., & Valle, F. (2022). Culturally Responsive Leadership for Latinx Immigrant Students: Advocacy in a U.S.-Mexico Border Alternative School. *Journal of School Leadership*, 32(6), 586–611. <https://doi.org/10.1177/10526846221095746>
- Dean, I. (2020). *Bildung – Heterogenität – Sprache. Rassistische Differenz- und Diskriminierungs-verhältnisse in Kita und Grundschule [Education – Heterogeneity – Language. Relations of racist difference and discrimination in child daycare centers and elementary school]*. Springer VS.
- de Boer, H., Timmermans, A. C., & van der Werf, M. P. C. (2018). The effects of teacher expectation interventions on teachers' expectations and student achievement: narrative review and meta-analysis. *Educational Research and Evaluation*, 24(3–5), 180–200. <https://doi.org/10.1080/13803611.2018.1550834>

- DeMatthews, D., Izquierdo, E., & Knight, D. S. (2017). Righting past wrongs: A superintendent's social justice leadership for dual language education along the U.S.-Mexico border. *Education Policy Analysis Archives*, 25(1), 1–28. <https://doi.org/10.14507/epaa.25.2436>
- DeMatthews, D. E., & Izquierdo, E. (2020). Supporting Mexican American Immigrant Students on the Border: A Case Study of Culturally Responsive Leadership in a Dual Language Elementary School. *Urban Education*, 55(3), 362–393. <https://doi.org/10.1177/0042085918756715>
- Drucks, S., & Bremm, N. (2021). Funktionen von Defizitorientierungen von Lehrkräften im Kontext unterschiedlicher herausfordernder Lagen [Functions of deficit orientations of teachers in the context of different challenging circumstances]. In I. von Ackeren, H. G. Holtappels, N. Bremm, & A. Hillebrand-Petri (Eds.), *Schulen in herausfordernden Lagen – Forschungsbefunde und Schulentwicklung in der Region Ruhr. Das Projekt „Potenziale entwickeln – Schulen stärken“* (pp. 244–276). Beltz Juventa.
- Egalite, A. J., & Kisida, B. (2018). The Effects of Teacher Match on Students' Academic Perceptions and Attitudes. *Educational Evaluation and Policy Analysis*, 40(1), 59–81. <https://doi.org/10.3102/0162373717714056>
- Eiden, S. (2018). Orientierungen von Lehrkräften im Umgang mit „Schwererreichbarkeit“ von Eltern [Orientations of teachers in dealing with “hard-to-reach” parents]. *Die Deutsche Schule*, 110(4), 326–340. <https://doi.org/10.31244/ddS.2018.04>
- El-Mafaalani, A., Waleciak, J., & Weitzel, G. (2023). Tatsächliche, messbare und subjektiv wahrgenommene Diskriminierung [Actual, measurable and subjective perceived discrimination]. In A. Scherr, A. El-Mafaalani, & G. Yüksel (Eds.), *Handbuch Diskriminierung* (pp. 195–213). Springer VS. https://doi.org/10.1007/978-3-658-10976-9_10
- Ezzani, M. (2021). A Principal's Approach to Leadership for Social Justice: Advancing Reflective and Anti-Oppressive Practices. *Journal of School Leadership*, 31(3), 227–247. <https://doi.org/10.1177/1052684620908347>
- Fend, H. (1980). *Theorie der Schule* [Theory of the school] Urban & Schwarzenberg.
- Fletcher, E. C., Hernandez-Gantes, V. M., & Smith, C. (2019). This Is My Neighborhood: An Exploration of Culturally Relevant Agency to Support High School Latinx Students in an Urban Career Academy. *The Qualitative Report*, 24(12), 3238–3262. <https://doi.org/10.46743/2160-3715/2019.3984>
- Fölker, L., Hertel, T., & Pfaff, N. (2015). Schule ‚im Brennpunkt‘ – Einleitung [School ‘in focus’ – Introduction]. In L. Fölker, T. Hertel, & N. Pfaff (Eds.), *Brennpunkt(-)Schule. Zum Verhältnis von Schule, Bildung und urbaner Segregation*. Verlag Barbara Budrich.
- Fölker, L., Hertel, T., & Pfaff, N. (2016). Schulische Praxis und Schulentwicklung unter Bedingungen sozialräumlicher Segregation. Befunde einer fallvergleichenden Studie zu zwei Schulen in benachteiligten Stadtteilen [School practice and school development under conditions of socio-spatial segregation. Findings of a comparative case study of two schools in disadvantaged districts]. In BMBF (Ed.), *Steuerung im Bildungssystem. Implementation und Wirkung neuer Steuerungsinstrumente im Schulwesen* (pp. 161–176). BMBF.
- Gay, G. (2000). *Culturally Resposive Teaching: Theory, Research, and Practice*. Teachers College Press.
- Gay, G. (2002). Preparing for Culturally Responsive Teaching. *Journal of Teacher Education*, 53(2), 106–116. <https://doi.org/10.1177/0022487102053002003>
- Gershenson, S., Lindsay, C. A., Papageorge, N. W., Campbell, R., & Rendon, J. H. (2023). *Spillover Effects of Black Teachers on White Teachers' Racial Competency: Mixed Methods Evidence from North Carolina* [IZA Discussion Papers, No. 16258]. IZA – Institute of Labor Economics. <https://docs.iza.org/dp16258.pdf>
- Geven, S. (2024). Tracking Procedures and Criteria and the SES Bias in Teacher Track Recommendations. *American Educational Research Journal*. Advance online publication. <https://doi.org/10.3102/00028312241288212>

- Gieske, M. (2013). *Mikropolitik und schulische Führung. Einflussstrategien von Schulleitern bei der Gestaltung organisationalen Wandels* [Micropolitics and school leadership. Strategies of influence of school leaders in shaping organizational change]. Verlag Julius Klinkhardt.
- Gordon, S. P., & Ronder, E. A. (2016). Perceptions of Culturally Responsive Leadership Inside and outside of a Principal Preparation Program. *International Journal of Educational Reform*, 25(2), 125–153. <https://doi.org/10.1177/105678791602500202>
- Gottuck, S., Pfaff, N., & Tervooren, A. (2021). Questioning Cultural and Power Relations as well as Debates on Disability and Migration: Concepts for Contemporary Inclusive Teacher Education. In L. Heidrich, Y. Karakaşoğlu, P. Mecheril, & S. Shure (Eds.), *Regimes of Belonging – Schools – Migrations: Teaching in (Trans)National Constellations* (pp. 387–404). Springer VS.
- Grissom, J. A., Kabourek, S. E., & Kramer, J. W. (2020). Exposure to Same-Race or Same-Ethnicity Teachers and Advanced Math Course-Taking in High School: Evidence from a Diverse Urban District. *Teachers College Record*, 122(7), 1–42. <https://doi.org/10.1177/016146812012200709>
- Guo-Brennan, L., & Guo-Brennan, M. (2021). Leading Welcoming and Inclusive Schools for Newcomer Students: A Conceptual Framework. *Leadership and Policy in Schools*, 20(1), 57–75. <https://doi.org/10.1080/15700763.2020.1838554>
- Hallinger, P. (2003). Leading Educational Change: reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education*, 33(3), 329–352. <https://doi.org/10.1080/0305764032000122005>
- Hallinger, P. (2013). A conceptual framework for systematic reviews of research in educational leadership and management. *Journal of Educational Administration*, 51(2), 126–149. <https://doi.org/10.1108/09578231311304670>
- Ham, S.-H., Kim, J., & Lee, S. (2020). Which schools are in greater need of culturally responsive leaders? A pedagogical uncertainty management perspective. *Multicultural Education Review*, 12(4), 250–266. <http://doi.org/10.1080/2005615X.2020.1842653>
- Hart, C. M. D., & Lindsay, C. A. (2024). Teacher-Student Race Match and Identification for Discretionary Educational Services. *American Educational Research Journal*, 61(3), 474–507. <https://doi.org/10.3102/00028312241229413>
- Helsper, W. (2008). Schulkulturen – die Schule als symbolische Sinnordnung [School cultures – the school as a symbolic order of meaning]. *Zeitschrift für Pädagogik*, 54(1), 63–80. <https://doi.org/10.25656/01:4336>
- Helsper, W., Dreier, L., Gibson, A., Kotzyba, K., & Niemann, M. (2017). *Exklusive Gymnasien und ihre Schüler: Passungsverhältnisse zwischen institutionellem und individuellem Schülerhabitus* [Exclusive Gymnasium schools and their students: the fit between institutional and individual student habitus]. Springer VS.
- Hertel, T. (2014). Zwischen Normbruch, Widerstand und ambivalenter Affirmation [Between breaking norms, resistance and ambivalent affirmation]. In J. Hagedorn (Ed.), *Jugend, Schule und Identität: Selbstwerdung und Identitätskonstruktion im Kontext Schule* (pp. 385–402). Springer Fachmedien Wiesbaden.
- Hertel, T. (2021). *Entziffern und Strafen. Schulische Disziplin zwischen Macht und Marginalisierung* [Deciphering and punishing. School discipline between power and marginalization]. transcript Verlag. <https://doi.org/10.14361/9783839455234>
- Hertel, T. (2024). Ambivalente Disziplinarsubjekte: Schüler:in-Sein und die Erfahrung der Strafe am Beispiel des schulischen Trainingsraums [Ambivalent disciplinary subjects: Being a student and the experience of punishment using the example of the school training room]. *Diskurs Kindheits- und Jugendforschung*, 19(3), 321–340. <https://doi.org/10.3224/diskurs.v19i3.03>

- Herzog-Punzenberger, B., Altrichter, H., Brown, M., Burns, D., Nortvedt, G. A., SkHgm, G., Wiese, E., Nayir, F., Fellner, M., McNamara, G., & O'Hara, J. (2020). Teachers responding to cultural diversity: case studies on assessment practices, challenges and experiences in secondary schools in Austria, Ireland, Norway and Turkey. *Educational Assessment, Evaluation and Accountability*, 32(3), 395–424. <https://doi.org/10.1007/s11092-020-09330-y>
- Hesbol, K. A., Bartee, J. S., & Amiri, F. (2020). Activism in Practice: The Influence of a Rural School Leader's Beliefs and Practices in Disrupting Historical Patterns of Underachievement in Traditionally Marginalized Students. *Impacting Education*, 5(2), 33–43. <https://doi.org/10.5195/ie.2020.134>
- Hwang, N., Graff, P., & Berends, M. (2023). Timing and Frequency Matter: Same Race/Ethnicity Teacher and Student Achievement by School Level and Classroom Organization. *Educational Policy*, 37(5), 1349–1379. <https://doi.org/10.1177/08959048221087212>
- Ialuna, F., Civitillo, S. & Jugert, P. (2024), Culturally Responsive Teaching, Teacher-Student Relationship and School Belongingness: A Multi-Informant Study in Ethnically Diverse Classrooms. *Learning, Culture and Social Interaction*, 47(3), 100839. <https://doi.org/10.1016/j.lcsi.2024.100839>
- Jahn, R. (2017). *Im Sog des Infantilen. Schulleitung als Beruf* [In the maelstrom of the infantile. School management as a profession]. Springer VS.
- Jayavant, S. (2016). Mapping the Complexities of Effective Leadership for Social Justice Praxis in Urban Auckland Primary Schools. *Education Sciences*, 6(1), 11. <https://doi.org/10.3390/educsci6010011>
- Jones, B. E., Bustamante, R. M., Gray, P., & Brooks, B. R. (2019). Exploring Cultural Responsiveness Among European American Principals in Rural Schools With High-Performing African American Readers. *Journal of School Leadership*, 29(2), 115–129. <https://doi.org/10.1177/1052684619832155>
- Kayser, A. A., Nash, A. M., & Kayser, B. (2020). Change-Makers: A Grassroots Approach to Culturally Responsive Leadership and Teaching. *Journal of Education Human Resources*, 38(1), 35–56. <https://doi.org/10.3138/jehr.2019-0009>
- Khalifa, M. (2018). *Culturally Responsive School Leadership*. Harvard Education Press.
- Khalifa, M. A., Gooden, M. A., & Davis, J. E. (2016). Culturally Responsive School Leadership: A Synthesis of the Literature. *Review of Educational Research*, 86(4), 1272–1311. <https://doi.org/10.3102/0034654316630383>
- Klein, E. D. (2018). *Erfolgreiches Schulleitungshandeln an Schulen in sozial deprivierter Lage: Eine Zusammenschau zentraler Grundlagen und Befunde aus der nationalen und internationalen Bildungsforschung. Expertise im Auftrag der Wübben Stiftung* [Successful school management at schools serving disadvantaged communities: A summary of central principles and findings from national and international educational research; expertise commissioned by the Wübben Foundation] (SHIP Working Paper Series, No. 02, Universität Duisburg-Essen. <https://doi.org/10.17185/duerpublico/45206>
- Klein, E. D., & Bronnert-Härle, H. (2022). Defizitorientierungen von Lehrkräften und ihr Zusammenhang mit der Führungspraxis an Schulen in sozialräumlich deprivierter Lage [Deficit-oriented perspectives of teachers and their relationship with leadership practices at schools serving disadvantaged communities]. *Zeitschrift für Pädagogik*, 68(5), 691–710. <https://doi.org/10.3262/ZP2205691>
- Klein, E. D., Bronnert-Härle, H., Boone, W. J., & Muslic, B. (2022). Constructs of Leadership and Diverging Institutional Environments. An Exploratory Comparative Study in the United States and Germany. *School Effectiveness and School Improvement*, 33(4), 564–587. <https://doi.org/10.1080/09243453.2022.2069129>

- Kollender, E. (2022). „Und das nenn’ ich Nehmermentalität.“ Rassistische Adressierungen von Eltern in der Schule im Kontext des sozialstaatlichen Aktivierungsdiskurses [“And that’s what I call a taker mentality.” Racist addressing of parents at school in the context of the welfare state activation discourse]. In Y. Akbaba, B. Bello, & K. Fereidooni (Eds.), *Pädagogische Professionalität und Migrationsdiskurse* (pp. 167–190). Springer Fachmedien Wiesbaden.
- Ladson-Billings, G. (1992). Reading between the lines and beyond the pages: A culturally relevant approach to literacy teaching. *Theory Into Practice*, 31(4), 312–320. <https://doi.org/10.1080/00405849209543558>
- Ladson-Billings, G. (1994). *The dreamkeepers: Successful teachers of African American children*. Jossey-Bass.
- Ladson-Billings, G. (2006). From the Achievement Gap to the Education Debt: Understanding Achievement in U.S. Schools. *Educational Researcher*, 35(7), 3–12. <https://doi.org/10.3102/0013189X035007003>
- Leithwood, K. (2021). A Review of Evidence about Equitable School Leadership. *Education Sciences*, 11(8), 377. <https://doi.org/10.3390/educsci11080377>
- Leithwood, K., Aitken, R., & Jantzi, D. (2006). *Making Schools Smarter. Leading With Evidence* (3rd ed.). Corwin Press.
- Leithwood, K., & Jantzi, D. (2006). Transformational school leadership for large-scale reform: effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement* 17(2), 201–227. <https://doi.org/10.1080/09243450600565829>
- Mastrogiovanni, D., & Freidus, A. (2023). “Deeper Than a Lesson”: New Jersey School Superintendents’ Approaches to Culturally Relevant Education. *Journal of Research on Leadership Education*, 19(4), 483–508. <https://doi.org/10.1177/19427751231215776>
- Maxwell, G. M., Locke, L. A., & Scheurich, J. J. (2013). Case study of three rural Texas superintendents as equity oriented change agents. *The Qualitative Report*, 18(11), 1–23. <https://doi.org/10.46743/2160-3715/2013.1546>
- Merten, R. (2003). Psychosoziale Folgen von Armut im Kindes- und Jugendalter [Psychosocial consequences of poverty in childhood and adolescence]. In C. Butterwegge & M. Klundt (Eds.), *Kinderarmut und Generationengerechtigkeit: Familien- und Sozialpolitik im demografischen Wandel* (pp. 137–151). VS Verlag für Sozialwissenschaften.
- Milner, H. R. (2003). Reflection, racial competence, and critical pedagogy: How do we prepare pre-service teachers to pose tough questions? *Race Ethnicity and Education*, 6(2), 193–208. <https://doi.org/10.1080/13613320308200>
- Moll, L., & Gonzalez, N. (1994). Lessons from research with language minority children. *Journal of Reading Behavior*, 26(4), 23–41. <https://doi.org/10.1080/10862969409547862>
- Narine, C. W., Jones, D., & Dawidowicz, P. (2022). Culturally Responsive Curriculum Implementation of Middle School Principals in Southern Texas. *Journal of Instructional Pedagogies*, 27, 1–20. <https://www.aabri.com/manuscripts/213438.pdf>
- Neugebauer, M. (2013). Wer entscheidet sich für ein Lehramtsstudium – und warum? Eine empirische Überprüfung der These von der Negativselektion in den Lehrerberuf [Who chooses to study education – and why? An empirical examination of the thesis of negative selection into the teaching profession]. *Zeitschrift für Erziehungswissenschaft*, 16(1), 157–184. <https://doi.org/10.1007/s11618-013-0343-y>
- Organization for Economic Co-operation and Development. (2019). *PISA 2018 Results (Volume II). Where All Students Can Succeed*. PISA, OECD Publishing. <https://doi.org/10.1787/b5fd1b8f-en>
- Oliver, G. P. (2023). English as a Second or Other Language-Focused Leadership to Support English Language Learners: A Case Study. *Journal of Education*, 203(3), 616–629. <https://doi.org/10.1177/00220574211054749>
- Palardy, G. J. (2008). Differential School Effects Among Low, Middle, and High Social Class Composition Schools. A Multiple Group, Multilevel Latent Growth Curve Analysis. *School Effectiveness and School Improvement*, 19(1), 21–49. <https://doi.org/10.1080/09243450801936845>

- Patterson, N., & Wang, A. H. (2024). Leading urban school teachers of Black and Brown students during social injustice and the pandemic. *School Leadership & Management*, 44(1), 76–95. <https://doi.org/10.1080/13632434.2023.2271515>
- Paris, D. (2012). Culturally Sustaining Pedagogy: A Needed Change in Stance, Terminology, and Practice. *Educational Researcher*, 41(3), 93–97. <https://doi.org/10.3102/0013189X12441244>
- Pettigrew, T. F. (1979). The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice. *Personality and Social Psychology Bulletin*, 5(4), 461–476. <https://doi.org/10.1177/014616727900500407>
- Potter, D., Reynolds, D., & Chapman, C. (2002). School Improvement for Schools Facing Challenging Circumstances: A review of research and practice. *School Leadership & Management*, 22(3), 243–256. <https://doi.org/10.1080/1363243022000020381>
- Pumptow, M., Nagengast, B., Hericks, N., & Neumann, M. (2024). SchuMaS-Ausgangserhebung: Rahmenbedingungen, Bedarfe und schulische Erfolgseinschätzungen zu Beginn der gemeinsamen Entwicklungsarbeit [SchuMaS initial survey: Conditions, needs and school success assessments at the start of the joint development work.]. In K. Maaz & A. Marx (Eds.), *SchuMaS – Schule macht stark. Sozialraumorientierte Schul- und Unterrichtsentwicklung an Schulen in schwierigen Lagen. Aufbau und erste Arbeitsergebnisse des Forschungsverbunds* (pp. 151–174). Waxmann.
- Racherbäumer, K. (2017). Rekonstruktionen zu Bedeutung und Funktionen der Lehrer-Schüler-Beziehung aus Sicht von Lehrerinnen und Lehrern an Schulen in sozial benachteiligter Lage [Reconstructions of the meaning and functions of the teacher-student relationship from the perspective of teachers at schools in socially disadvantaged circumstances]. In V. Manitijs & P. Döbelstein (Eds.), *Schulentwicklungsarbeit in herausfordernden Lagen* (pp. 123–139). wbv.
- Rodela, K. C., & Rodriguez-Mojica, C. (2020). Equity Leadership Informed by Community Cultural Wealth: Counterstories of Latinx School Administrators. *Educational Administration Quarterly*, 56(2), 289–320. <https://doi.org/10.1177/0013161X19847513>
- Salmen, C. (2019). *The evidence in evidence-based policy. The case of No Child Left Behind*. Dissertation. Universität Wien. <https://doi.org/10.25365/thesis.77494>
- Savvopoulos, D., Saiti, A., & Arar, K. (2024). The role of the school head in inclusion and culturally responsive leadership. *Journal of Educational Administration and History*, 56(2), 107–131. <https://doi.org/10.1080/00220620.2022.2122419>
- Schein, E. H. (2010). *Organizational Culture and Leadership* (3rd ed.). Jossey-Bass.
- Scribner, J. P., Weingand, D. H., & Sanzo, K. L. (2021). Fostering Cultural Responsiveness in an Urban High School: A Case Study. *NASSP Bulletin*, 105(3), 153–172. <https://doi.org/10.1177/01926365211036513>
- Sendzik, N., Edelstein, B., Hermstein, B., & Racherbäumer, K. (2023). Was kann die Einzelschule gegen Bildungsungleichheit ausrichten? Eine kritische Auseinandersetzung mit Prämissen, Ansätzen und Praktiken der Schulentwicklung(sforschung) [What can individual schools do to challenge educational inequality? A critical examination of premises, approaches and practices of school development (research)]. *Die Deutsche Schule*, 115(3), 177–188. <https://doi.org/10.31244/ddS.2023.03.01>
- Shannon-Baker, P., Porfilio, B. J., & Plough, B. (2020). Parents Want Their Voices to “Matter”: Perspectives on School Enrollment in a Shrinking Urban School District. *The Journal of Educational Foundations*, 33(1–4), 77–94. <https://img1.wsimg.com/blobby/go/7b7d8256-7ebe-43a8-9f8d-677e64c4940a/downloads/TEF2020.pdf?ver=1725638753774>

- Stubbe, T. C., Schaufelberger, R., Krieg, M., Kleinkorres, R., & Schlitter, T. (2023). Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen [School career preferences at the transition to secondary school and the association with performance-related and social characteristics]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. Stubbe (Eds.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 231–248). Waxmann.
- Sun, J., & Leithwood, K. (2015). Direction-setting school leadership practices: a meta-analytical review of evidence about their influence. *School Effectiveness and School Improvement*, 26(4), 499–523. <https://doi.org/10.1080/09243453.2015.1005106>
- Thompson, D. A., Virella, P., Goings, R. B., & Kelly, K. (2023). Black Women Principals and Expressions of Culturally Responsive School Leadership During Crisis: An Exploratory Study. *Journal of Urban Learning, Teaching, and Research*, 17(Special), 89–111. <https://julttr.online/article/87487-black-women-principals-and-expressions-of-culturally-responsive-school-leadership-during-crisis-an-exploratory-study>
- Torres Clark, J., & Chrispeels, J. H. (2022). Using multiple leadership frames to understand how two school principals are influencing teachers' practices and achievement of Hispanic English learners. *Journal of Educational Administration*, 60(3), 303–322. <https://doi.org/10.1108/JEA-03-2021-0054>
- Valencia, R. R. (2010). *Dismantling Contemporary Deficit Thinking: Educational Thought and Practice*. Routledge.
- Vassallo, B. (2022). Leading the flock: Examining the characteristics of multicultural school leaders in their quest for equitable schooling. *Improving Schools*, 25(1), 22–36. <https://doi.org/10.1177/1365480221999133>
- Wang, H., & Hall, N. C. (2018). A systematic review of teachers' causal attributions: Prevalence, correlates, and consequences. *Frontiers in Psychology*, 9, 2305, 1–22. <https://doi.org/10.3389/fpsyg.2018.02305>
- Warwas, J. (2012). *Berufliches Selbstverständnis, Beanspruchung und Bewältigung in der Schulleitung* [Professional self-image, stress and coping in school management]. VS Verlag für Sozialwissenschaften.
- Wellgraf, S. (2012). *Hauptschüler. Zur gesellschaftlichen Produktion von Verachtung* [Hauptschule-students. On the social production of contempt]. transcript Verlag. <https://doi.org/10.14361/transcript.9783839420539>
- Ylimaki, R., & Brunderman, L. (2019). School Development in Culturally Diverse U.S. Schools: Balancing Evidence-Based Policies and Education Values. *Education Sciences*, 9(2), 84. <https://doi.org/10.3390/educsci9020084>

Teil B:

Interventionskonzepte

Digitale Schreibförderung im Ganzttag an Grundschulen: Herausforderungen und Potenziale für Schulentwicklungsprozesse

Kathrin Gade & Vera Busse

Zusammenfassung

Das Kapitel stellt eine zielgruppenspezifische Intervention vor, die im Rahmen des BMBF-geförderten Projekts *KommSchreib!* im Ganzttag an Grundschulen mit schreibschwachen Lernenden in Klasse 3 und 4 durchgeführt wurde. Ziel der Intervention ist eine zum Regelunterricht komplementäre Förderung des prozessorientierten Schreibens und kooperativen Lernens. In Schreib-Arbeitsgemeinschaften (Schreib-AGen) wurde in kooperativen Lernsettings ein multimediales und mehrsprachiges Kochbuch unter Einsatz von Tablets erarbeitet. Der Beitrag erläutert die angestrebten Kompetenzentwicklungen und die Designprinzipien der Schreib-AGen. Anschließend werden Herausforderungen in der Implementierung herausgearbeitet. Diese zeigen, dass digitalisierungsbezogene Schulentwicklungsmaßnahmen im Bereich der Technologie- und Kooperationsentwicklung erforderlich sind, um die Potenziale zu realisieren, die die Einbindung des Ganztags in Bezug auf die Entwicklung fachlicher und überfachlicher Fähigkeiten bietet.

Schlagnvorte: Sprachförderung, Literalität, digitale Medien, Bildungsgleichheit, Ganzttag

Abstract

The chapter presents a target group-specific intervention conducted as part of the BMBF-funded project *KommSchreib!* in afternoon workshops at primary schools with struggling writers in Grades 3–4. The intervention aims to complement regular classroom instruction by supporting process-ori-

ented writing and cooperative learning. In the writing workshops, students created a multimedia and multilingual cookbook in cooperative learning settings using tablets. The chapter explains the targeted competence development and the design principles of the writing workshops. It then discusses implementation challenges highlighting that school development measures related to digitalization are necessary in the areas of technology and cooperation development to realize the potential that the integration of afternoon workshops offer for the development of both subject-specific and cross-curricular skills.

Keywords: Language support, literacy, digital media, educational inequality, full-day support

1. Einleitung

Die Förderung kompetenten Schreibens ist nicht nur für den erfolgreichen Übergang nach der Grundschule, sondern auch für den weiteren schulischen Erfolg von zentraler Bedeutung (Cutler & Graham, 2008) und vor dem Hintergrund sprachlicher und sozialer Heterogenität besonders relevant. Eine systematische Einbindung der Schreibförderung in den Ganztag an Grundschulen kann dabei fehlende Lerngelegenheiten im Bereich der sprachlichen Bildung kompensieren (Edele & Stanat, 2022), die interkulturelle Öffnung von Schule (Karakaşoğlu, 2011) stärken und somit (u. a. zugewanderungsbezogene) Bildungsnachteile mindern.

Der folgende Artikel stellt zunächst Ziele und Design des vom Bundesministerium für Bildung und Forschung geförderten Projekts *KommSchreib!*¹ dar, innerhalb dessen eine zielgruppenspezifische Intervention zur Förderung des Schreibens im Ganztag an Grundschulen durchgeführt wurde (Kapitel 2). Daran anschließend werden einzelne Kompetenzentwicklungsziele und Designprinzipien der zielgruppenspezifischen Intervention dargestellt (Kapitel 3). Anschließend skizzieren wir Herausforderungen, beschreiben kurzfristig gefundene Lösungen (Kapitel 4) und überlegen daran anknüpfend, welche Veränderungen nötig sind, um die Potenziale des Ganztags zu einer ergänzenden Schreibförderung optimal zu nutzen (Kapitel 5).

1 „KommSchreib! Kooperativ, kompetent, motiviert Schreiben“ ist ein Interventionsprojekt innerhalb der Förderrichtlinie des Bundesministeriums für Bildung und Forschung (BMBF) „Sprachliche Bildung in der Einwanderungsgesellschaft“, Laufzeit von 10/2022–09/2025 (Förderkennzeichen 01JM2210).

2. Überblick über das BMBF-Projekt *KommSchreib!*: Projektziele und Projektdesign

Übergeordnetes Ziel des BMBF-geförderten Projekts *KommSchreib!* ist es, Schreibkompetenz, Schreibmotivation und soziale Partizipation an Grundschulen zu fördern. Hierzu wurden Grundschullehrkräfte in prozessorientierten Methoden zur Schreibförderung fortgebildet, und gemeinsam mit ihnen wurden Lehrmaterialien angepasst. Die entsprechenden Methoden und Materialien wurden dann von den Lehrkräften im Regelunterricht in den Klassen 3 und 4 eingesetzt. Das Projekt ist als quasi-experimentelle Interventionsstudie mit Prä-, Post- und Follow-up-Testung angelegt. Die Stichprobe umfasst elf Grundschulen ($N = \sim 1.300$ Schüler*innen) in Nordrhein-Westfalen. Im Rahmen des Projekts wurde zusätzlich eine zielgruppenspezifische quasi-experimentelle Intervention in Form von Schreib-Arbeitsgemeinschaften (Schreib-AGen) im Ganztag in den Jahrgängen 3 und 4 durchgeführt, die Gegenstand des vorliegenden Beitrags ist. Übergeordnetes Ziel ist eine über den Regelunterricht hinausgehende zusätzliche Schreibförderung im Ganztag für schreibschwache Lernende an den jeweiligen Schulstandorten.

Die Lehrkräfte sowie das Ganztagspersonal identifizierten hierzu schreibschwache Lernende sowie Lernende mit geringen Deutschkompetenzen und schlugen diese für die Schreib-AGen vor; der Besuch war für die Kinder jedoch nicht verpflichtend. Bislang nahmen 53 Kindern an fünf Schulstandorten der Experimentalgruppen teil. Die jetzige Stichprobe erhielt also entsprechende Schreibfördermaßnahmen im Regelunterricht und nahm dazu an den wöchentlichen Schreib-AGen im Nachmittagsbereich des Ganztags teil. Die Schreib-AGen an den Schulen der Wartekontrollgruppe finden im Schuljahr 2024/25 statt.

Während die Lehrkräfte im Regelunterricht primär analog arbeiteten und prozessorientierte Schreibfördermaßnahmen insbesondere beim Erstellen narrativer und deskriptiver Texte zum Einsatz kamen, wurden in den Schreib-AGen prozessorientierte Schreibfördermaßnahmen zum Erstellen von instruktionalen Texten in einer digitalisierten Lernumgebung durchgeführt. Konkret erarbeiteten die Kinder in den Schreib-AGen in kooperativen Lernsettings mit Hilfe von Tablets ein multimediales und mehrsprachiges Kochbuch.

Die Schreib-AGen wurden von externen Honorarkräften (Lehramtsstudierende im Masterstudium) geleitet, die im Rahmen des Projekts entsprechend geschult und von der Erstautorin betreut wurden. Dem regulären Ganztagspersonal wurde angeboten, an den Schreib-AGen teilzunehmen, um das Konzept bei Bedarf in das Ganztagsangebot zu integrieren; die Teilnahme war

jedoch nicht verpflichtend. Die Lehrkräfte des Regelunterrichts waren über die Schreib-AGen informiert, aber nicht direkt eingebunden.

Zur Überprüfung der Wirksamkeit wurden Prä- und Posttestdaten sowie Fragebogendaten zu mehreren Messzeitpunkten erhoben und durch Unterrichtsbeobachtungen und Befragungen der Kinder sowie der Honorarkräfte ergänzt (Mixed-Methods-Ansatz).

Die Implementationstreue des entwickelten Konzepts wurde auf verschiedene Weise überprüft. Basierend auf dem prozessorientierten Schreibansatz und in Anlehnung an die Intervention im Regelunterricht wurde eine detaillierte Unterrichtsplanung für jede Sitzung erstellt. Die Erstautorin, eine erfahrene Lehrkraft, bereitete die Honorarkräfte auf die Durchführung der Schreib-AGen vor. Um den unterschiedlichen Bedürfnissen der Kinder gerecht zu werden, wurde ein Pool mit leistungsdifferenziertem und adaptierbarem Lernmaterial erstellt, auf den die Honorarkräfte Zugriff erhielten. Kernelemente der prozessorientierten Schreibförderung (vgl. Kapitel 3.2) sollten dabei in allen Sitzungen durchgeführt werden. Nach Ende jeder AG-Sitzung füllten die durchführenden Honorarkräfte einen kurzen Online-Fragebogen aus, in dem sie angaben, ob und in welchem Umfang die Sitzung gemäß der Unterrichtsplanung umgesetzt wurde. Bei Herausforderungen wurden in Absprache mit den Autorinnen entsprechende Maßnahmen ergriffen (vgl. auch Kapitel 4). Zudem wurde in jeder Sitzung eine Bewertung bzw. ein Kommentar zum bereitgestellten Lernmaterial abgegeben und es konnten Beobachtungen sowie mögliche Besonderheiten der Sitzungen (wie z. B. Krankheitsvertretungen) eingetragen werden. Alle Standorte wurden mehrmals von der Erstautorin besucht und der tatsächliche Ablauf der Schreib-AGen wurde mit der Unterrichtsplanung abgeglichen.

3. Angestrebte Kompetenzentwicklung und Design-Prinzipien

Die Interventionsmaßnahmen konzentrieren sich primär auf drei zentrale Bereiche der Kompetenzentwicklung: Inhaltskompetenz, Schreibkompetenz (zu der hier auch Schreibmotivation als Teilaspekt der Entwicklung von Schreibkompetenz im Sinne eines erweiterten Kompetenzbegriffs nach Weinert, 1994 zählt) und Medienkompetenz.

3.1. Inhaltskompetenz

Die angestrebte Entwicklung auf Ebene der Inhaltskompetenz zielt auf die kindgerechte Beschäftigung mit ausgewogener Ernährung ab. Für die ersten elf Sitzungen der Schreib-AGen wurden fünf konkrete Rezepte ausgewählt und didaktisch aufbereitet (Fruchtspieß, Salat mit Dressing, Frühstücksbrot, Dessert mit Obststückchen, Durstlöcher). Diese Auswahl knüpft an das in der Schuleingangsphase laut Lehrplan erworbene Wissen zu gesunder Ernährung an (vgl. MSB NRW, 2021) und erfolgte vor dem Hintergrund, dass Kinder und Jugendliche tendenziell zu wenig Obst, Gemüse oder Lebensmittel mit komplexen Kohlenhydraten zu sich nehmen (Mensink et al., 2021). Die vermittelten Inhalte (z. B. Vitamingehalt von Obst und Gemüse, Vollkornprodukte, Verzicht auf Industriezucker) berühren dabei die Vorgaben aus den Bildungsstandards bzw. Vorgaben der Bundesländer, sich in der Primarstufe mit den Themen Körper und Gesundheit sowie Natur und Umwelt (MSB NRW, 2021) zu befassen und Bezüge zwischen den Themen zu erkennen. Der (u. a. lebensmittelhygienisch und gesundheitlich begründete) Verzicht auf tierische Lebensmittel erlaubt außerdem, Ernährungsformen unter dem Aspekt Verbraucherbildung und Bildung für nachhaltige Entwicklung (MSB NRW, 2021) zu thematisieren. Der organisatorische Rahmen der Schreib-AGen ermöglicht in jeder zweiten Sitzung die Zubereitung und den anschließenden Verzehr kleiner Speisen auf der Grundlage der gemeinsam erarbeiteten Rezepte, was sich neben dem inhaltlichen Lernzuwachs positiv auf die Motivation der Lernenden (vgl. Kapitel 3.2) auswirken sollte.

3.2. Schreibkompetenz

Die angestrebte Entwicklung auf Ebene der Schreibkompetenz gliedert sich in drei Komponenten. Erste Komponente ist die Vermittlung *genrespezifischer Merkmale* von instruktionalen Texten wie Rezepten (Augst et al., 2007; Bachmann, 2002). Ziel einer Instruktion ist die Nachvollziehbarkeit einer Handlung für den Adressaten eines Textes (z. B. Weber, 1982). Auf der Grundlage der selbstverfassten Rezepte sollten also entsprechende Gerichte zubereitet werden können. Vermittelt wurden genrespezifische Merkmale instruktionaler Texte wie Textrahmung (z. B. Überschrift und Abschlussfloskel „Guten Appetit“), Textgliederungselemente (z. B. Zutatenliste und Arbeitsschritte), Angemessenheit des Schreibgegenstands (z. B. genaue und damit funktionale Beschreibung der Handlungsschritte) oder lexikalisch-semantisch treffender Wortschatz (z. B. „schälen“ statt „abmachen“).

Zweite Komponente ist die im Regelunterricht verfolgte *prozessorientierte Schreibstrategieförderung* (Graham et al., 2000; Graham et al., 2003). Hierzu wurden in den Schreib-AGen wiederkehrend Phasen des *Planens*, des *Verschriftens* und des *Revidierens* (vgl. auch Philipp, 2020) sowie (Peer-) Feedback (vgl. auch die Meta-Analyse von Scherer et al., 2024) in die einzelnen Sitzungen integriert.

Das *Planen* als erster Teilprozess des Schreibens begann zunächst mit der lexikalisch-semanticen und/oder diskursiven Entlastung des späteren eigenen Textes (zur Relevanz von Wortschatzförderung vgl. auch Busse & Hardy, 2023). So wurde jede der Sitzungen von den Kindern mit einem tabletbasierten Spiel in Einzel- oder Partnerarbeit begonnen, welches sich entweder auf den thematischen Wortschatz der Sitzung (z. B. Obst, Gemüse, Mengenangaben) oder auf Genrewissen zur Textsorte Rezept bezog (z. B. richtige Gliederung von Zutaten, Zubehör und Arbeitsschritte oder passende Satzanfänge). Vor der Erstellung von Wortschatzspielen mittels H5P (Schoblick 2021; Wörfel & Huesmann, 2020) für den Einstieg in die Sitzungen der Schreib-AGen wurde der Grundwortschatz für die Primarstufe in Nordrhein-Westfalen (MSB NRW, 2019) gesichtet und, sofern thematisch passend, in das jeweilige Spiel integriert. Auf der Grundlage des im Einstieg gefestigten thematischen Wortschatzes und der vermittelten genrespezifischen Textsortenkonventionen für Rezepte wurde dann gemeinsam mit den Kindern ein Rezept geplant. Anschließend modellierten die geschulten Honorarkräfte exemplarisch das Verfassen des Rezeptes, indem sie ihre mentalen Prozesse beim Schreiben kindgerecht verbalisierten, um diese für die Lernenden der Schreib-AGen nachvollziehbar zu machen (vgl. auch Philipp, 2020). Das Modellieren ist dabei dem SRSD-Ansatz (*self-regulated strategy development*) zur strategiegeleiteten Schreibförderung entnommen (Harris & Graham, 2000). Diese Phase wurde in den Schreib-AGen interaktiv gestaltet, d. h. die Kinder wurden durch Nachfragen in den von der Honorarkraft modellierten Planungsprozess ihres Rezeptes einbezogen (zum Modellieren und der didaktischen Umsetzung in den verschiedenen Phasen des Schreibprozesses vgl. auch Siekmann et al., 2022 und Sturm, 2022). Der beim Modellieren durch die Honorarkraft entstandene Text diente vor allem schreibschwächeren Kindern als Modelltext für die Erstellung eines eigenen Rezepts.

Das *Verschriften* als zweiter Teilprozess des Schreibens wurde von den Lernenden am Tablet ausgeführt. Hier konnten sie, je nach graphomotorischer Präferenz, wahlweise mittels interner oder externer Tastatur, einem Tablet-Stift oder via Touchscreen ihre Ideen zur Gestaltung eines eigenen Rezepts im E-Book verschriften. Aufgrund der kognitiven (anspruchsvollen) Anforderungen dieses Prozesses (vgl. auch Philipp, 2014; Sturm, 2022) konnten sich

schwächere Lernende die modellierten Planungs- und Verschriftungsprozesse eines Rezeptes auch noch einmal in einem Lernvideo am Tablet anschauen oder wahlweise mit einem Peer arbeiten.

Das *Revidieren* als dritter Teilprozess des Schreibens fand zu Beginn jeder zweiten Sitzung der Schreib-AGen statt. Dort analysierten die Kinder gemeinsam mit der Honorarkraft eigene und fremde Modelltexte hinsichtlich Verbesserungsmöglichkeiten (zum Einsatz von Modelltexten vgl. Siekmann et al., 2022). Anschließend gaben sich die Kinder mündliches Peer-Feedback zu ihren eigenen Texten anhand eines Kriterienrasters (zur Effektivität von Kriterienrastern vgl. Peltzer et al., 2024; zu Effektivität von Peer-Feedback vgl. Scherer et al., 2024). Bei der anschließenden gemeinschaftlichen Zubereitung der eigenen Rezepte sollten etwaige nicht funktionale Stellen in den Instruktionen oder unpassende Ausdrücke zunächst ausfindig gemacht und im Anschluss auch schriftlich revidiert werden. Zudem wurde für das gesamte Kochbuch in den letzten Sitzungen der Schreib-AGen Zeit für Revisionen eingeplant, die sich zum einen auf die Struktur des Kochbuchs an sich (z. B. Einfügen eines Inhaltsverzeichnisses) sowie auf Textoberflächenmerkmale bezogen. Übergeordnetes Designprinzip aller Schreibphasen war, das Schreiben in eine gemeinschaftliche Aktivität zu verwandeln und Interaktion über das Schreiben zwischen Lehrkraft und Lernenden und zwischen den Peers zu fördern (vgl. auch Rijlaarsdam et al., 2018).

Neben der Vermittlung *genrespezifischer Merkmale* und der *prozessorientierten Schreibstrategieförderung* stellt die Förderung von *Schreibmotivation* die dritte zentrale Komponente dar. Bezugnehmend auf die Selbstbestimmungstheorie (Ryan & Deci, 2000) kann allgemein angenommen werden, dass der Einsatz digitaler Medien Interesse am Schreiben weckt (vgl. Camacho et al., 2021) und dadurch intrinsische Motivation fördert. Dies gilt insbesondere für die spielerischen Elemente (z. B. Wortschatz-Memory), die sich als motivationsförderlich erweisen sollten, da sie dem Gamification-Ansatz folgen (vgl. Zeybek & Saygi, 2024). Zudem kann angenommen werden, dass authentische, gut situierte Schreibanlässe (Becker-Mrotzeck & Lemke, 2022), die den Nutzen des Schreibens deutlich machen, motivierend sind (vgl. auch Bruning & Horn, 2000). Insbesondere die darauffolgenden handlungsorientierten Phasen der Zubereitung kleiner Gerichte nach einer zuvor selbst verfassten Anleitung sollten sich positiv auf die intrinsische Motivation der Lernenden auswirken (vgl. Boscolo & Gelati, 2019). Darüber hinaus berücksichtigt die Lerneinheit die drei psychologischen Grundbedürfnisse, die in der Selbstbestimmungstheorie (Ryan & Deci, 2000) postuliert werden, wie folgt:

Erstens sollte die Selbstbestimmtheit bzw. Autonomie der Lernenden gefördert werden, indem die Kinder Rezepte adaptierten und um eigene

Zutaten und Schritte ergänzten oder kürzten. Sie waren frei in der Wahl ihres Schreibwerkzeugs (magnetische Tastatur, Touchpad-Tastatur, Stifteingabe oder Touchscreen), konnten zwischen Text- und Audioeingabe wählen oder beides kombinieren und die Gestaltung ihrer Medienprodukte frei bestimmen. Zudem konnten eigene Seiten im Kochbuch ergänzt werden wie Inhaltsverzeichnisse, Texte über sich selbst oder eigene Rezepte (insbesondere in Sitzungen 11–14; siehe Anhang). Zu Beginn und am Ende jeder Sitzung hatten die Kinder Wahlfreiheit bei der Auswahl und den Schwierigkeitsstufen der Wortschatzspiele.

Zweitens sollte das Kompetenzerleben der Lernenden in den Schreib-AGen durch das Erstellen eines sichtbaren Medienprodukts gestärkt werden. In der ersten Sitzung wurde das gemeinsame Schreibziel hergeleitet und den Kindern ein ausgedrucktes Muster-Kochbuch gezeigt. Am Ende jeder Sitzung wurde den Kindern der Fortschritt durch einen Abgleich zwischen Muster-Kochbuch und ihren eigenen E-Books visualisiert. In der abschließenden Sitzung erhielten alle Kinder Urkunden, ein ausgedrucktes Exemplar ihres Kochbuchs sowie die epub-Datei ihres Kochbuchs zur Weiterarbeit, um ihre Kompetenzentwicklung zu visualisieren.

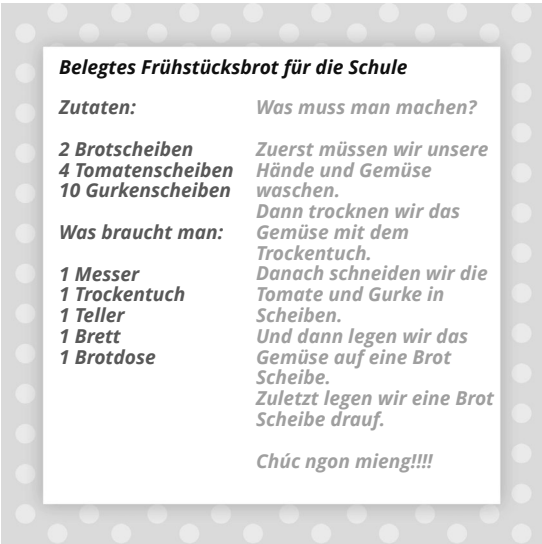
Drittens sollte das Gefühl sozialer Eingebundenheit durch kooperatives Lernen gestärkt werden (Huber, 2019; vgl. auch Schürer & van Ophuysen, 2022), wobei sich kooperative Schreibaktivitäten und Peer-Feedback nachweislich förderlich auf Schreibmotivation auswirken (Camacho et al., 2021; De Smedt et al., 2018). In jeder Sitzung konnten in der Planungsphase sowie zum Abschluss tabletbasierte Aktivitäten (vgl. Kapitel 2) in Partnerarbeit durchgeführt werden. Peer-Feedback zu den Rezepten während der Übungsphasen wurde in vier Sitzungen gegeben (Sitzungen 4, 6, 8, 10) und noch einmal zum Abschluss zum gesamten E-Book (Sitzung 12). Darüber hinaus sollte das gemeinsame Zubereiten von Rezepten, das gemeinsame Essen und anschließende Aufräumen das Gefühl sozialer Zugehörigkeit fördern.

Eine weitere Komponente zur Schaffung eines positiven emotionalen Schreibumfelds (vgl. Bruning & Horn, 2000) und zur Verbindung mit der persönlichen Lebenswelt der Kinder ist die Einbindung und Wertschätzung ihrer lebensweltlichen Sprachen. Unter dem Aspekt „Leben in Vielfalt“ (MSB NRW, 2021, S. 186) ist die Erwartung formuliert, dass Kinder auch mittels digitaler Tools Lebensgewohnheiten und Traditionen aus ihrem Alltag wie z. B. Essensgewohnheiten vorstellen. Mehrsprachig aufwachsende Kinder wurden ermutigt, ihr sprachliches Potenzial im digitalen Kochbuch darzustellen (z. B. mittels Text und/oder Audioaufnahme). Dies sollte den Lebensweltbezug unterstreichen und wertschätzende Einstellungen (Cummins, 2013),

Akzeptanz und Wohlbefinden (vgl. Busse et al., 2019) sowie das Erleben von sozialer Eingebundenheit fördern.

Die bislang entstandenen Medienprodukte zeigen, dass einige Kinder, wie im Musterkochbuch vorgeschlagen, auch von der Möglichkeit Gebrauch machten, nicht nur ein Inhaltsverzeichnis, sondern zusätzlich auch eigene Texte (z. B. „Das bin ich“) zu ergänzen. Darüber hinaus brachten einige Kinder auch ihre mehrsprachigen Ressourcen in ihr Medienprodukt ein. Dies reichte von Textrahmungselementen wie *Guten Appetit* (vgl. Abbildung 1) bis hin zu ganzen Rezepten in den Familiensprachen der Kinder.

Abbildung 1: Beispielseite aus einem E-Book (Verwendung der Familiensprache bei der Textrahmung)



3.3. Medienkompetenz

Die angestrebte Entwicklung auf Ebene der Medienkompetenz fokussierte innerhalb der Schreib-AGen insbesondere zwei Aspekte und zwar einerseits das *Bedienen und Anwenden* der Tablets sowie andererseits das *Produzieren und Präsentieren* von Medienprodukten (für das in den Projektschulen geltende Medienkompetenzmodell vgl. Medienberatung NRW, 2020).

Alle Kinder hatten Zugriff auf ein Tablet, das jeweils mit der App *BookCreator* zur Erstellung eines eigenen multimedialen E-Books ausgestattet war. An zwei Standorten wurden vom Schulträger konfigurierte Tablets

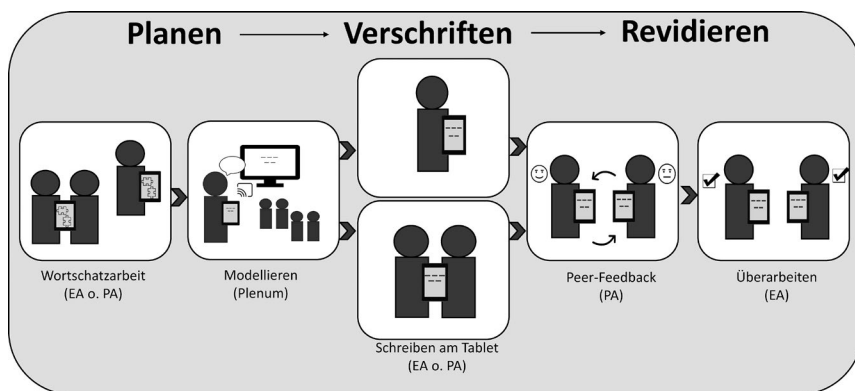
genutzt, während an drei weiteren Standorten vom BMBF-Projekt *Komm-Schreib!* finanzierte und konfigurierte Geräte zum Einsatz kamen.

Die Kinder übten neben dem *Bedienen* der Hardware zum Schreiben (z. B. Tastatur oder Tablet-Stift) auch das *Anwenden* verschiedener Apps ein (z. B. *BookCreator* oder Kamera). Um Daten sicher zu speichern, wiederzufinden und abzurufen (Medienberatung NRW, 2020), sicherten sie am Ende jeder Sitzung unter Anleitung der Honorarkraft selbständig ihre epub-Dateien.

Des Weiteren wurde das *Produzieren* und *Präsentieren* von Medienprodukten (Medienberatung NRW, 2020) eingeübt. Die Kinder nutzten neben den Gestaltungsmöglichkeiten in *BookCreator* (z. B. Schriftart, -größe, -farbe, Hintergrund) auch Bild-, Video- und Tonaufnahmen (teilweise in mehreren Sprachen) für ihre epub-Datei. Dabei wurde auch das Thema Urheber- und Nutzungsrechte behandelt, da einige Kinder Bildmaterial von sich oder aus dem Internet verwendeten.

Im Sinne einer möglichst vielfältigen Nutzung des Tablets (Haßler et al., 2016) und beziehungsweise auf die übergeordneten Projektziele, wurde das Tablet in den drei Schreibphasen der Sitzungen (Planen, Verschriften, Revidieren) in unterschiedlichen Arbeitskonstellationen (z. B. *one-to-one with and without collaboration*) eingesetzt (vgl. Abbildung 2).

Abbildung 2: Einsatz des Tablets in den verschiedenen Schreibphasen



4. Herausforderungen und Potenziale für Schulentwicklungsprozesse

Das Konzept der Schreib-AGen bietet Potenziale für die Entwicklung von Unterricht, da die Verzahnung von Regelunterricht mit dem Ganztags Synergieeffekte erzeugen und entsprechende Strategien des prozessorientierten

Schreibens am Tablet an verschiedenen Stellen in lebensweltlichen Kontexten eingeübt und gefestigt werden könnten. Diese Verzahnung von Schreib-AGen im Ganzttag mit Unterrichtskonzepten zum Schreiben im Regelunterricht am Vormittag setzen allerdings noch weitere Maßnahmen voraus, die innerhalb des zeitlich begrenzten *KommSchreib!*-Projekts nicht zu leisten waren. So müsste u. a. das Ganztagspersonal entsprechend qualifiziert werden, damit eine Verankerung im festen AG-Repertoire des Ganztags erfolgen kann. Die Umsetzung des Konzepts der Schreib-AGen, inklusive aller beschriebenen Maßnahmen zur Förderung von Inhalts-, Schreib- und Medienkompetenz, berührt auch weitere (digitalisierungsbezogene) Schulentwicklungsfragen. Im Modell von Eickelmann und Gerick (2018) gehören neben Unterrichts-, Personal- und Organisationsentwicklung auch die Technologie- und Kooperationsentwicklung zu den zentralen Bereichen. Zur Dimension der Technologieentwicklung zählt die IT-Ausstattung der Schule sowie die Unterstützung bei IT-Problemen (Labusch et al., 2020). Zur Dimension der Kooperationsentwicklung gehört bei der (außer-)unterrichtlichen Nutzung digitaler Medien zur Schreibförderung eine Kooperation verschiedener Akteure. Als gemeinsames Angebot von Schule sowie Trägern der Kinder- und Jugendhilfe kennzeichnet den Ganzttag genau dieses wichtige Zusammenspiel von verschiedenen Zuständigkeiten. Die folgenden Reflexionen konzentrieren sich insbesondere auf diese beiden Dimensionen, da sich Technologie- und Kooperationsentwicklung im Hinblick auf die Durchführbarkeit der Schreib-AGen auch für bereits geschulte Honorarkräfte als besonders herausfordernd erwiesen.

4.1. Technologieentwicklung

In Bezug auf die Technologieentwicklung an Grundschulen konstatieren Studien wie TIMSS 2019 (Selter et al., 2020) oder die KIM-Studie 2022 (Forschungsverbund Süd-West, 2022) insgesamt eine positive Ausstattungstendenz, die sich mit dem Ende des Abrufs aus Landes- oder Bundesförderprogrammen (DigitalPakt Schule, IT-Administration) bzw. deren Neuauflage (Digitalpakt 2.0) noch verstärken dürfte. Nach wie vor ist die Ausstattungslage an Schulen jedoch heterogen, wie die Gegebenheiten an den einzelnen Projektschulen zeigten. Insgesamt kann angenommen werden, dass die auch im internationalen Vergleich unterdurchschnittliche Bereitstellung von digitalen Medien für den Leseunterricht (Lorenz et al., 2023) auch für den Schreibunterricht an Grundschulen gilt. So hatten die Kinder an einigen Standorten kaum Erfahrung mit dem Schreiben am Tablet, und es konnte nicht von flüs-

sigem Schreiben (insbesondere mit der Tastatur) ausgegangen werden. Eine gezielte Förderung des Tastaturschreibens konnte innerhalb der Schreib-AGen nicht durchgeführt werden, daher wurde mehr Zeit für die Texteingaben eingeplant, und es konnte auch mit einem Tablet-Stift geschrieben werden. Generell wäre jedoch in den Schulen das Trainieren von Schreibflüssigkeit auch an der Tastatur wichtig, um hierarchiehöhere Schreibprozesse wie das Planen, Verschriften oder Überarbeiten zu entlasten (Graham & Harris, 2000).

Darüber hinaus stellten instabiles WLAN und fehlende Präsentationsmöglichkeiten via Apple TV oder Beamer Herausforderungen dar, daher wurden mehr zeitliche und personelle Ressourcen an diesen Standorten eingeplant, u. a. um mitgebrachte Geräte zu installieren und die von den Kindern erstellten Rezeptbücher ggf. lokal auf Geräten der Honorarkräfte zu speichern. Solche pragmatischen Herangehensweisen können jedoch keine langfristigen Lösungen für die skizzierten Probleme darstellen. Voraussetzung für den Transfer des vorgestellten Konzepts und dem damit verbundenen Potenzial für den Regelunterricht wäre ein Zugang zu einer funktionierenden technischen Infrastruktur der Schule und die Verfügbarkeit von IT-Support auch in den üblichen Randzeiten des Ganztags.

Insbesondere an Standorten in herausfordernden Lagen (vgl. Sozialindexstufen von Einzelschulen in Nordrhein-Westfalen) konnte festgestellt werden, dass die Kinder noch keine ausreichend ausgeprägte selbstregulative Fähigkeit im Umgang mit digitalen Werkzeugen (Medienberatung NRW, 2020) erworben hatten, die ihnen ermöglicht, Medien auch gezielt als Lernressource zu nutzen (KMK, 2016) statt vorrangig zur Unterhaltung (vgl. hierzu auch Paschke & Thomasius, 2024). Das Tablet und der damit verbundene Zugang zum Internet bot Ablenkungspotenzial (vgl. auch Schaumburg & Prasse, 2019; Stetina & Kryspin-Exner, 2009), das sich z. B. durch das übermäßige Gestalten von farbigen Hintergründen und Schriftarten zu Lasten des eigentlichen Schreibprozesses oder das Öffnen von *YouTube*-Clips ohne Themenbezug zu den Schreib-AGen (und unter Umgehen der vom Schulträger installierten Content-Filter) zeigte.

Um diesem Problem zu begegnen, wurden neben mehr personeller Ressource (i. d. R. durch Zweifachbesetzung an herausfordernden Standorten) alle Honorarkräfte verstärkt darin geschult, Regeln zum Umgang mit dem Tablet in den Schreib-AGen zu etablieren (zur Rolle von Regeln und Ritualen vgl. auch Helmke, 2022). Hier gehen technikbezogene Herausforderungen auch in die Medienkompetenzentwicklung über. Um Ablenkungen zu minimieren, sind neben Übungen zur Selbstregulation auch technisch-organisatorische Maßnahmen innerhalb der IT-Infrastruktur erforderlich. Über das Mobile-Device-Management des Schulträgers kann beispielsweise die

Überprüfung der Aktivitäten der Kinder ermöglicht werden. Eigene zeitnahe Anpassungen, wie z. B. am Content-Filter oder das schnelle Deaktivieren der WLAN-Verfügbarkeit durch die Honorarkraft waren nicht möglich. Dies ist jedoch für störungsfreies Arbeiten unabdingbar, um z. B. privates Surfen im Internet der Kinder zu unterbinden.

4.2. Kooperationsentwicklung zwischen Unterricht und Ganzttag

In Bezug auf eine gelingende Kooperationsentwicklung zwischen Unterricht und Ganzttag sind u. a. die Räumlichkeiten und die personellen Ressourcen relevant, für die in Nordrhein-Westfalen i. d. R. die Kreise und kreisfreien Städte zuständig sind. Insbesondere an Schulstandorten in herausfordernden Lagen fehlten oft angemessene räumliche Ressourcen für die Schreibphasen und die Zubereitung von Lebensmitteln. Zwar konnten durch die Doppelbesetzung an diesen Standorten provisorische Schreib- und Arbeitsmöglichkeiten eingerichtet werden, auf struktureller Ebene wäre jedoch eine Schulgebäudegestaltung wünschenswert, die vielseitig für den Regelunterricht und die Ganztagsbetreuung genutzt werden könnte. Die Bedeutung einer pädagogischen Architektur (Buddensiek, 2009; Stadler-Altmann, 2016) für zukunftsfähige Lehr- und Lernräumen ist diesbezüglich zu betonen.

Ein weiterer Aspekt der Kooperationsentwicklung betrifft die personellen Ressourcen im Ganzttag, für die ebenfalls die Kommune zuständig ist. Nur ein Schulstandort konnte eine durchgängige Begleitung durch das Ganztagspersonal bieten. Dies erwies sich in mehrfacher Hinsicht als nachteilig. Zum einen sollte das begleitende Ganztagspersonal im Bereich der Schreibförderung qualifiziert werden, um ähnliche Konzepte selbst durchführen zu können. Diese Qualifizierung war an Standorten mit Personalmangel nicht möglich. Zum anderen führte die häufige Abwesenheit des Ganztagspersonals dazu, dass Honorarkräfte z. B. nicht über Ausfälle des Ganztags und damit die Abwesenheit aller Kinder der Schreib-AGen informiert wurden oder der Ganzttag nicht für die Übergabe von Informationsdokumenten an die Erziehungsberechtigten erreichbar war. Dies führte zu zusätzlichem Organisationsaufwand und zu vermehrten Fehlzeiten der Kinder, die das Angebot als weniger verbindlich betrachteten.

5. Fazit und Ausblick

Die zuvor skizzierten Herausforderungen zeigen, dass die praktische Umsetzung von digitalen Schreibfördermaßnahmen im Ganzttag durchaus anspruchsvoll ist und verschiedene räumliche, technische und personelle Voraussetzungen erfordert. Neben basalen Inhalts- und Schreibkompetenzen (z. B. mittels der hier vorgestellten prozessorientierten Schreibfördermaßnahmen) müssen zudem auch grundlegende Medienkompetenzen angebahnt werden, etwa bezüglich digitaler Schreibwerkzeuge oder einer selbstregulierten Mediennutzung, um die Potenziale des vorgestellten Konzepts der Schreib-AGen voll auszuschöpfen.

Mit Blick auf das Ganztagsförderungsgesetz (GaFöG), das ab 2026 einen Rechtsanspruch auf Ganztagsbetreuung für Grundschulkinder garantiert, wird sich nach dem Ausbau von räumlichen und personellen Kapazitäten auch die inhaltliche Frage stellen, welche innovativen Konzepte allgemein geeignet sind, um Sprachbildung im Regelunterricht auch mit dem Nachmittagsbereich zu verzahnen. Um Konzepte zur Verbindung von fachlichem und überfachlichem Lernen im Ganzttag wie im Projekt *KommSchreib!* angedacht über die Einzelschulebene hinaus zu etablieren, ist daher eine enge Kooperation zwischen den verschiedenen schulischen und außerschulischen Akteuren nötig, die auch einer Steuerung der Gesamtzusammenhänge durch Bildungspolitik bzw. -administration im Sinne einer komplexen Schulentwicklung (Rolff, 2018) bedürfen. Die hier skizzierte multimodale Herangehensweise zur Schreibförderung könnte ein vielversprechender Ansatz sein, der schreibschwache Lernende im Ganzttag unterstützt, wenn ausreichend zeitliche, räumliche und personelle Ressourcen zur Verfügung gestellt werden. So müsste z. B. auch das Ganztagspersonal einen systematischen Zugang zum Fortbildungswesen für Lehrkräfte erhalten, um optimale Implementierungsergebnisse zu erzielen und das Potenzial des hier vorgestellten Konzepts im Sinne bestmöglicher Kompetenzentwicklung bei den Kindern zu entfalten.

Vorläufige Ergebnisse zeigen nicht nur eine Zunahme der Textlänge und Qualität in der Teilstichprobe, sondern auch ein hohes Wohlbefinden in den einzelnen Sitzungen, wobei die Schreibfreude am Tablet deutlich höher war als die allgemeine Schreibfreude (Gade et al., 2024). Die Auswertung der Wartekontrollgruppe sowie die Analyse der Textqualität und Schreibförderung im Regelunterricht stehen zwar noch aus, jedoch weisen diese ersten Ergebnisse damit trotz der skizzierten Herausforderungen in eine positive Richtung.

Literatur

- Augst, G., Disselhoff, K., Henrich, A., Pohl, T., & Völzing, P. L. (2007). *Text-Sorten-Kompetenz. Eine echte Longitudinalstudie zur Entwicklung der Textkompetenz im Grundschulalter* [Text variety literacy. A true longitudinal study on the development of text literacy at primary school age]. Lang.
- Bachmann, T. (2002). *Kohäsion und Kohärenz: Indikatoren für Schreibentwicklung. Zum Aufbau kohärenzstiftender Strukturen in instruktiven Texten von Kindern und Jugendlichen* [Cohesion and coherence: indicators for writing development. On the development of coherence-creating structures in instructive texts by children and young people]. Studien-Verlag.
- Becker-Mrotzeck, M., & Lemke, V. (2022). Gute Schreibaufgaben für alle Fächer [Good writing tasks for all subjects]. In V. Busse, N. Müller, & L. Siekmann (Eds.), *Schreiben fachübergreifend fördern. Grundlagen und Anregungen für Schule, Unterricht und Lehrkräftebildung* (pp. 73–95). Klett Kallmeyer.
- Boscolo, P., & Gelati, C. (2019). Motivating writers. In S. Graham, C. A. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed., pp. 51–78). Guilford Press.
- Bruning, R., & Horn, C. (2000). Developing motivation to write. *Educational Psychologist*, 35(1), 25–37. https://doi.org/10.1207/S15326985EP3501_4
- Buddensiek, W. (2009). Fraktale Schularchitektur [Fractal school architecture]. In J. Böhme (Ed.), *Schularchitektur im interdisziplinären Diskurs: Territorialisierungskrise und Gestaltungsperspektiven des schulischen Bildungsraums* (1st ed., pp. 315–329). VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-91868-6>
- Busse, V., Cenoz, J., Dalmann, N., & Rogge, F. (2019). Addressing linguistic diversity in the language classroom in a resource-oriented way: An intervention study with primary school children. *Language Learning*, 70(2), 382–419. <https://doi.org/10.1111/lang.12382>
- Busse, V., & Hardy, I. (2023). Literalität und Mehrsprachigkeit: Begriffsklärungen, Förderansätze und Forschungsbefunde [Literacy and multilingualism: definitions, approaches and research findings]. *Unterrichtswissenschaft*, 51(2), 149–168. <http://dx.doi.org/10.1007/s42010-023-00175-0>
- Camacho, A., Alves, R. A., & Boscolo, P. (2021). Writing motivation in school: A systematic review of empirical research in the early twenty-first century. *Educational Psychology Review*, 33(1), 213–247. <https://doi.org/10.1007/s10648-020-09530-4>
- Cummins, J. (2013). Immigrant students' academic achievement: Understanding the intersections between research, theory and policy. In I. Gogolin, I. Lange, U. Michel, & H. H. Reich (Eds.), *Herausforderung Bildungssprache – und wie man sie meistert* (pp. 19–41). Waxmann Verlag. <https://content.e-bookshelf.de/media/reading/L-829460-78cd5bfd4d.pdf>
- Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology*, 100 (4), 907. <https://doi.org/10.1037/a0012656>
- De Smedt, F., Graham, S., & Van Keer, H. (2018). The bright and dark side of writing motivation: Effects of explicit instruction and peer assistance. *The Journal of Educational Research*, 112(2), 152–167. <https://doi.org/10.1080/00220671.2018.1461598>
- Edele, A., & Stanat, P. (2022). Zuwanderung und soziale Ungleichheit [Immigration and social inequality]. In H. Reinders, D. Bergs-Winkels, A. Prochnow, & I. Post (Eds.), *Empirische Bildungsforschung. Eine elementare Einführung* (pp. 1105–1126). Springer Verlag. <https://doi.org/10.1007/978-3-658-27277-7>
- Eickelmann, B., & Gerick, J. (2018). Herausforderungen und Zielsetzungen im Kontext der Digitalisierung von Schule und Unterricht. Teil 2: Fünf Dimensionen der Schulentwicklung zur erfolgreichen Integration digitaler Medien [Challenges and objectives in the context of the digitalization of schools and teaching. Part 2: Five dimensions of school development for the successful integration of digital media]. *Schulverwaltung Hessen/Rheinland Pfalz. Zeitschrift für Schulentwicklung und Schulmanagement*, 23(6), 184–188.

- Gade, K., Graham, S., & Busse, V. (2024, 11th September). *Prozessorientierte Schreibförderung: Digitales und kooperatives Arbeiten in der Grundschule* [Process-oriented writing support: digital and collaborative work in elementary school]. AEPF 2024: Wissenstransfer zwischen Bildungsforschung, Bildungspraxis und Lehrer*innenbildung, Osnabrück, Deutschland.
- Gerick, J., Vennemann, M., Lorenz, R., & Eickelmann, B. (2014). Schulische Ausstattung mit digitalen Medien in der Grundschule [School equipment with digital media in elementary school]. In B. Eickelmann, R. Lorenz, M. Vennemann, J. Gerick, & W. Bos (Eds.), *Grundschule in der digitalen Gesellschaft. Befunde aus den Schulleistungstudien IGLU und TIMSS 2011* (pp. 19–34). Waxmann Verlag.
- Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and development. *Educational Psychologist* 35(1), 3–12. https://doi.org/10.1207/S15326985EP3501_2
- Graham, S., Harris, K. R., & Troia, G. A. (2000). Self-regulated strategy development revisited: Teaching writing strategies to struggling writers. *Topics in Language Disorders*, 20(4), 1–14. <https://doi.org/10.1097/00011363-200020040-00003>
- Graham, S., McKeown, D., Kiuvara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, 104(4), 879–896. <https://doi.org/10.1037/a0029939>
- Harris, K. R., Graham, S., & Mason, L. H. (2003). Self-regulated strategy development in the classroom: Part of a balanced approach to writing instruction for students with disabilities. *Focus on exceptional children*, 35(7), 1–16. <https://doi.org/10.17161/fec.v35i7.6799>
- Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, 32(2), 139–156. <https://doi.org/10.1111/jcal.12123>
- Helmke, A. (2022). *Unterrichtsqualität und Professionalisierung: Diagnostik von Lehr-Lern-Prozessen und evidenzbasierte Unterrichtsentwicklung* [Teaching quality and professionalization: diagnostics of teaching-learning processes and evidence-based teaching development]. Klett Kallmeyer.
- Hillmayr, D., Reinhold, F., Zierwald, L., & Reiss, K. (2016). *Digitale Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe. Einsatzmöglichkeiten, Umsetzung und Wirksamkeit* [Digital media in mathematics and science lessons at secondary level. Possible applications, implementation and effectiveness]. Waxmann Verlag. <https://doi.org/10.25656/01:15482>
- Huber, C. (2019). Ein integriertes Rahmenmodell zur Förderung sozialer Integration im inklusiven Unterricht – sozialpsychologische Grundlagen, empirische Befunde und schulpraktische Ableitungen [An integrated framework model for promoting social integration in inclusive education – socio-psychological foundations, empirical findings and practical implications for schools]. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 88(1), 27–43. <https://doi.org/10.2378/vhn2019.art06d>
- Karakaşoğlu, Y., Gruhn M., & Wojciechowicz, A. A. (2011). *Interkulturelle Schulentwicklung unter der Lupe. (Inter-)Nationale Impulse und Herausforderungen für Steuerungsstrategien am Beispiel Bremen* [Intercultural school development under the microscope. (Inter-)national impulses and challenges for steering strategies using the example of Bremen]. Waxmann Verlag.
- KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Ed.) (2016). *Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt.“* [Strategy of the Standing Conference of the Ministers of Education and Cultural Affairs “Education in the digital world.”]. Retrieved June, 11, 2024, from https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit_Weiterbildung.pdf

- Labusch, A., Eickelmann, B., & Conze, D. (2020). *ICILS 2018 #Transfer: Gestaltung digitaler Schulentwicklung in Deutschland* [ICILS 2018 #Transfer: Shaping digital school development in Germany] (1st ed.). Waxmann Verlag. <https://doi.org/10.25656/01:21351>
- Lorenz, R., Goldhammer, F., & Glondys, M. (2023). Digitalisierung in Grundschulen [Digitization in elementary school]. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. C. Stubbe (Eds.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (pp. 197–214). Waxmann Verlag. <https://doi.org/10.25656/01:28075>
- Medienpädagogischer Forschungsverbund Südwest (2022). *KIM-Studie 2022. Kindheit, Internet, Medien. Basisuntersuchung zum Medienumgang 6- bis 13- Jähriger* [KIM Study 2022: Childhood, Internet, Media. Basic study on media use by 6- to 13-year-olds]. Retrieved June, 11, 2024, from <https://www.mpf.de/fileadmin/files/Studien/KIM/2022/KIM-Studie2022websitefinal.pdf>
- Medienberatung NRW (Ed.) (2020). *Medienkompetenzrahmen NRW* [Media literacy framework NRW] (3. ed.). LVR Druck.
- Mensink, G. B. M., Haftenberger, M., Lage Barbosa, C., Brettenscheider, A.-K., Lehmann, F., Frank, M., Heide, K., Moosbrugger R., Patelakis, E. & Perlit, H. (2021). *EsKiMo II – Die Ernährungsstudie als KiGGS-Modul* [EsKiMo II – The nutrition study as a KiGGS module]. Robert Koch-Institut.
- MSB NRW (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen) (Ed.) (2019). *Hinweise und Materialien für einen systematischen Rechtschreibunterricht in der Primarstufe in NRW – Handreichung* [Information and materials for systematic spelling lessons at primary level in NRW – handout.]. Retrieved June, 11, 2024, from <https://www.schulentwicklung.nrw.de/cms/upload/grundwortschatz/Hinweise-und-Materialien-fuer-einen-systematischen-Rechtschreibunterricht-in-der-Primarstufe-in-NRW-Handreichung.pdf>
- MSB NRW (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen Sachunterricht) (Ed.) (2021). *Lehrplan für die Primarstufe in Nordrhein-Westfalen. Fach Sachunterricht* [Curriculum for the primary level in North Rhine-Westphalia. Subject: general studies. Retrieved June 11, 2024, from https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan/292/ps_lp_su_einzeldatei_2021_08_02.pdf
- Paschke, K., & Thomasius, R. (2024). Digitale Mediennutzung und psychische Gesundheit bei Adoleszenten – eine narrative Übersicht [Digital media use and mental health in adolescents – a narrative overview]. *Bundesgesundheitsblatt*, 67, 456–464. <https://doi.org/10.1007/s00103-024-03848-y>
- Peltzer, K., Lira Lorca, A., Krause, U.-M., & Busse, V. (2024). Effects of formative feedback on argumentative writing in English and cross-linguistic transfer to German. *Learning and Instruction*, 92, 101935. <https://doi.org/10.1016/j.learninstruc.2024.101935>
- Philipp, M. (2014). *Selbstreguliertes Schreiben: Schreibstrategien erfolgreich vermitteln* [Self-regulated writing: successfully teaching writing strategies]. Beltz Verlag.
- Philipp, M. (2020). *Grundlagen der effektiven Schreibdidaktik und der systematischen Schreibförderung* [Fundamentals of effective writing didactics and systematic writing support] (8th ed.). Schneider Verlag Hohengehren.
- Rijlaarsdam, G., Janssen, T., Rietdijk, S., & van Weijen, D. (2018). Reporting design principles for effective instruction of writing: Interventions as constructs. In R. Fidalgo Redondo, K. Harris, & M. Braaksma (Eds.), *Design Principles for Teaching Effective Writing* (pp. 280–313). Brill. https://doi.org/10.1163/9789004270480_013
- Rolf, H.-G. (2018). Grundlagen der Schulentwicklung [Basics of school development]. In C. G. Bühren & H.-G. Rolf, (Eds.), *Handbuch Schulentwicklung und Schulentwicklungsberatung* (2nd ed.) (pp. 12–39). Beltz.

- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Schaumburg, H., & Prasse, D. (2019). *Medien und Schule* [Media and school]. Bertelsmann Verlag.
- Scherer, S., Graham, S., & Busse, V. (2024). How effective is feedback for L1, L2, and FL learners' writing? A meta-analysis. *Learning and Instruction*, 92. <https://doi.org/10.1016/j.learninstruc.2024.101961>
- Schoblick, R. (2021). *Multimedial lehren und lernen: Digitale Lerninhalte erstellen mit H5P* [Multimedia teaching and learning: creating digital learning content with H5P]. Carl Hanser Verlag.
- Schürer, S., & van Ophuysen, S. (2022). Relationship between group cohesion and social participation of pupils with learning and behavioural difficulties. *European Journal of Special Needs Education*, 37(5), 866–881. <http://dx.doi.org/10.1080/08856257.2021.1963150>
- Selter, C., McElvany, N., Wendt, H., Steffensky, M., Kasper, D., Köller, O., & Schwippert, K. (2020). TIMSS 2019. *Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* [TIMSS 2019: An international comparison of the mathematical and scientific skills of primary school children in Germany]. Waxmann Verlag. <https://doi.org/10.31244/9783830993193>
- Siekmann, L., Müller, N., & Busse, V. (2022). Wie können Lehrkräfte Feedback zur Schreibförderung in größeren Lerngruppen einsetzen? Fortbildungsinhalte aus dem Projekt WeLiKe Feedback [How can teachers use feedback to promote writing in larger learning groups? Training content from the WeLiKe Feedback project]. In V. Busse, N. Müller, & L. Siekmann (Eds.), *Schreiben fachübergreifend fördern. Grundlagen und Anregungen für Schule, Unterricht und Lehrkräftebildung* (pp. 268–287). Klett Kallmeyer.
- Stadler-Altman, U. (2016). Lernumgebungen. Erziehungswissenschaftliche und architekturkritische Perspektiven auf Schulgebäude und Klassenzimmer [Learning environments. Educational science and architecture-critical perspectives on school buildings and classrooms]. In U. Stadler-Altman (Ed.), *Lernumgebungen. Erziehungswissenschaftliche Perspektiven auf Schulgebäude und Klassenzimmer*. (pp. 7–16). Verlag Barbara Budrich. <https://doi.org/10.25656/01:18185>
- Stetina, B. U., & Kryspin-Exner, I. (2009). Gesundheitsrelevante Aspekte neuer Medien [Health-related aspects of new media]. In B. U. Stetina & I. Kryspin-Exner (Eds.), *Gesundheit und Neue Medien: Psychologische Aspekte der Interaktion mit Informations- und Kommunikationstechnologien* (pp. 1–36). Springer Verlag.
- Sturm, A. (2022). Prozess- und produktorientierte Schreibförderung in Kombination [Process and product-oriented writing support in combination]. In V. Busse, N. Müller, & L. Siekmann (Eds.), *Schreiben fachübergreifend fördern. Grundlagen und Praxisanregungen für Schule, Unterricht und Lehrerinnen- und Lehrerbildung*. Klett Kallmeyer.
- Weber, U. (1982). *Reihe Germanistische Linguistik, 41. Instruktionsverhalten und Sprechhandlungsfähigkeit: Eine empirische Untersuchung zur Sprachentwicklung* [Instructional behavior and speech negotiation skills: An empirical study of language development]. De Gruyter. <https://doi.org/10.1515/9783111371092>
- Weinert, F. E. (2014). Vergleichende Leistungsmessungen in Schulen – eine umstrittene Selbstverständlichkeit [Comparative performance measurements in schools – a controversial matter of course]. In F. E. Weinert (Ed.), *Leistungsmessungen in Schule* (3rd ed.) (pp. 17–32). Beltz.

- Woerfel, T., & Huesmann, I. (2020). Unterricht mit digitalen Medien organisieren. Sprachsensibel unterrichten. Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache [Organizing lessons with digital media. Language-sensitive teaching. Mercator Institute for Language Support and German as a Second Language]. Retrieved June 11, 2024, from <https://www.bimm.at/themenplattform/wp-content/uploads/2024/03/woerfelhuesmann-sprachsensibelunterrichten.pdf>
- Zeybek, N., & Saygi, E. (2024). Gamification in education: Why, where, when, and how? – A systematic review. *Games and Culture*, 19(2), 237–264. <https://doi.org/10.1177/15554120231158625>

Promoting Vocabulary and Grammar in German First-Grade Classrooms: Design and Structure of a Digital Language Intervention

Leonie Dargiewicz*, Svenja Wehrhöfer*, Annika Ohle-Peters, Fani Lauermann & Nele McElvany

* Shared first authorship

Abstract

Language skills are essential for educational success and social participation. Vocabulary and grammar are key dimensions of language skills. Differences in language skills between children in these dimensions are already apparent at the start of formal schooling, which is why systematic early support is necessary. The project “Shaping linguistic education together with elementary schools, families, and educational professionals: What works?” (SPEAK) is developing a digital language intervention in vocabulary and grammar for classroom instruction in first grade by providing successive exercises, developed based on theoretical approaches and available evidence. SPEAK focuses on training school-relevant vocabulary and the gender and case grammar system. This article aims to present the conception of the intervention. Vocabulary and grammar are trained twice a week during regular lessons over 15 weeks. The effectiveness will be tested in an experimental pre-post-test design.

Keywords: Vocabulary, grammar, intervention, digital media, Grade 1, elementary school

Zusammenfassung

Sprachkompetenzen sind essentiell für den Bildungserfolg und die gesellschaftliche Teilhabe. Wortschatz und Grammatik bilden zentrale Dimen-

sionen von Sprachkompetenzen. Sprachliche Disparitäten in diesen Bereichen zwischen jungen Kindern zeigen sich bereits zu Beginn der formalen Beschulung, weswegen eine systematische frühe Förderung notwendig ist. In dem Projekt „Sprachliche Bildung gemeinsam mit Grundschulen, Familien und pädagogischen Fachkräften gestalten: Was wirkt?“ (SPEAK) wird eine digitale Sprachförderung in den Bereichen Wortschatz und Grammatik für den Unterricht in der ersten Klassenstufe entwickelt, die auf theoretischen Ansätzen und Erkenntnissen aus der Wissenschaft basiert. Es wird schulisch relevanter Wortschatz sowie das Genus- und Kasusystem trainiert. Im vorliegenden Beitrag wird die Konzeption der Intervention vorgestellt. Das Wortschatz- und Grammatiktraining wird über 15 Wochen zweimal wöchentlich im Schulalltag durchgeführt.

Schlagworte: Wortschatz, Grammatik, Intervention, digitale Medien, 1. Klasse, Grundschule

1. Introduction

Language skills are central prerequisites for academic success and social participation (Dale et al., 2023). Both vocabulary and grammatical skills are critical dimensions of language competencies and lay the foundation for other literacy skills, such as reading comprehension (Hjetland et al., 2020; Zheng et al., 2023). As reading skills determine academic and occupational success, students need sufficient preparatory language skills, such as vocabulary and grammatical skills, to support their reading comprehension (e.g., McElvany & Artelt, 2009). However, students from families with low socioeconomic status (SES) or second language learners (SLLs) who speak a different language at home than in school may encounter particular challenges (e.g., Levine et al., 2020). These differences also apply to Germany, where 25% of preschoolers are SLLs, and one in five minors lives in families with extremely low SES (Funcke & Menne, 2023; Maaz et al., 2024). There are significant differences between students with high and low SES as well as first language learners (FLLs) and SLLs in vocabulary and grammatical skills at the beginning of formal schooling in Germany (e.g., Limbird et al., 2014; Passaretta et al., 2022; Schulz & Grimm, 2018) and this discrepancy continues into later elementary school (e.g., Heppt & Stanat, 2020). Furthermore, Wanzek (2014) noted that teachers rarely explicitly promote students' vocabulary in their regular lessons. Thus, early interventions in both vocabulary and grammatical skills are needed.

Accordingly, the research project “Shaping linguistic education together with elementary schools, families, and educational professionals: What

works?” (SPEAK) at the Center for Research on Education and School Development (IFS) at TU Dortmund University developed a digital-based language intervention addressing first-grade classrooms in linguistically and socially heterogeneous neighborhoods.

Various theoretical and empirically evaluated approaches use promising implicit and explicit methods, such as storybook reading or providing syntactic cues (full overview see chapter 3), to support children’s vocabulary and grammatical skills (e.g., Bebout & Belke, 2017; Grøver et al., 2020). Against the background of children’s diversity, especially in elementary school, digital media can take this into account and amplify these methods’ effectiveness and wide-spread use (e.g., Hammerschmidt-Snidarich et al., 2019). Besides adaptive methods, the increasing number of all-day elementary schools in Germany (Maaz et al., 2024) provides new opportunities for extensive and systematic language instruction for first graders, such as the one implemented in SPEAK.

This book chapter provides an overview of the specific vocabulary and grammatical skills identified to be relevant at the beginning of formal schooling and the didactic approaches and methods used within the intervention to foster vocabulary and grammatical skills. Subsequently, we provide an outline of the structure of the intervention accompanied by task examples.

2. Identification of Relevant Vocabulary and Grammatical Skills

In terms of vocabulary, the SPEAK-intervention aims to train school-relevant vocabulary. On the one hand, this contains academic vocabulary, which is a prerequisite for academic success and further knowledge acquisition (Anstrom et al., 2010). Academic vocabulary includes vocabulary for specific disciplines (such as science) and domain-general academic vocabulary used across different disciplines (Baumann & Graves, 2010). On the other hand, we focused on vocabulary that children require to comprehend the content in school and to understand conversations or explanations in the lessons. This kind of vocabulary is important because most school interactions are contextually reduced, potentially hampering children’s understanding of class content and ability to participate in class activities (Bailey & Butler, 2003). For the purposes of the intervention, it was also important to train vocabulary that was likely unknown to first graders, even though it is needed for their academic success.

To identify relevant school-based vocabulary, we analyzed dictionaries with basic as well as more advanced vocabulary for German FLLs and SLLs

because they contain vocabulary relevant to understanding class discourse and task operators. Likewise, we considered vocabulary tests for children this age as they test relevant vocabulary children this age should know. As a third category of target words, “likely unknown” words were identified using the frequency distribution from the *childLex* corpus, a German lexical database that lists word frequencies based on their occurrence in children’s literature (Schroeder et al., 2015). After identifying a corpus of 653 potential target words, the next step was to reduce the corpus to relevant vocabulary for first graders. The “relevance” of school-based vocabulary was determined by two research assistants who rated the words independently based on the conceptualization of school-relevant vocabulary mentioned above. At least one of the following criteria must apply: (i) the word is either relevant for the curriculum and, therefore, for a specific discipline, or (ii) the word is relevant across different disciplines, or (iii) the word is relevant for school-based communication. Words with a 100% agreement rate were considered. Taking into account the word type, we identified 88 target words that also fit semantically in order to be trained context-based.

In terms of grammar, we aimed for training those grammatical skills that cause the greatest difficulties for German SLLs, but the selected learning objectives should also attend to the language needs of FLLs as the intervention is conducted in regular classrooms. Accordingly, we decided to focus on gender and case classes as various studies found great difficulties, especially for SLLs in this area (Ronniger et al., 2019; Schulz & Grimm, 2018). Anyhow, evidence suggests that also a large proportion of FLLs have not acquired specific case classes, such as the dative case when entering elementary school (Ulrich, 2017). We therefore assumed that a focus on gender and case classes would meet the language needs of the majority of students in German first-grade classrooms.

To decide on a sequence order of grammatical targets, we aimed to follow the natural progression of language acquisition. Studies of German gender and case acquisition in both FLLs and SLLs provide indications about the phases in which acquisition might occur (e.g., Binanzer, 2017; Clahsen, 1984; Ulrich, 2017; Ulrich et al., 2021). As these findings need to be understood in the context of the German gender/case system, a brief overview of it is provided below.

The German language is not gender-neutral and has three gender classes (masculine, feminine, and neuter), which are usually not applied to the noun itself but to the article, the attributive adjective (internal nominal group), and pronouns (external nominal group) for nouns in the singular (Wöllstein, 2022). Thus, gender establishes reference points within a sentence (Köpcke & Zubin, 2009). There are four classes of cases (nominative, genitive, dative, and

accusative), which are also generally applied to articles, attributive adjectives, and pronouns (Wöllstein, 2022). Gender and case classifications in German operate in a complex system, as shown in Table 1 for the singular definite article. Case inflection, therefore, requires knowledge of the gender of the noun. Notably, there is no one-to-one correspondence between the form and function of the article, which can aggravate students’ language acquisition (Kauschke, 2012).

Table 1: The Singular Declination of the Definite Article in German (in English ‘the’)

Case	Gender category		
	Masculine	Feminine	Neuter
Nominative	der	die	das
Genitive	des	der	des
Dative	dem	der	dem
Accusative	den	die	das

Based on her research on how FLLs and SLLs in German elementary schools acquire grammatical gender in German, Binanzer (2017) proposed a model of the acquisition phases. According to the model, language learners first establish semantic gender congruence (in the order: personal pronoun > definite article > attributive adjective > relative pronoun) and later move on to formal grammatical congruence (in the order: definite article > attributive adjective > relative pronoun > personal pronoun), with more valid patterns being acquired before less valid patterns. In this context, validity means the frequency with which a noun with a particular formal characteristic is assigned to a gender class expected based on that pattern. Hence, the central function of gender, namely the establishment of reference, must be emphasized within a didactic concept, and the acquisition sequence must be taken into account (Binanzer & Wecker, 2020). Concerning case acquisition, studies indicate a pattern in the order of acquisition for both FLLs and SLLs (e.g., Clahsen, 1984; Ulrich et al., 2021). The accusative case (in the order: nominal phrase > prepositional phrase) is acquired before the dative case (in the order: prepositional phrase > nominal phrase). We used evidence from previous studies to develop a sequence order of the grammatical learning objectives trained within the intervention (see Table 2), even though empirical study results are limited and results are ambiguous (e.g., Ulrich, 2017).

Table 2: Sequence of Grammatical Targets Within the Intervention

Order	Grammatical target	Further description
1	Xe → feminine	Nouns ending on the [a] vowel → feminine
2	Xel, Xer, Xen → masculine	Nouns ending on <el>, <er>, and <en> → masculine
3	X _{monosyllabic} → neuter	Monosyllabic nouns → neuter
4	Accusative nominal phrase	
5	Accusative prepositional phrase	
6	Dative prepositional phrase	
7	Dative nominal phrase	

3. Approaches and Methods to Foster Vocabulary and Grammatical Skills

Language interventions are typically categorized into explicit and implicit approaches and methods (Dickinson et al., 2019; Goo et al., 2015; Norris & Ortega, 2000). In vocabulary instruction, implicit methods are used to infer the word’s meaning from contextual cues (McElvany et al., 2017; Sternberg & Powell, 1983). In contrast, in explicit methods, definitions and explanations are provided (Marulis & Neuman, 2010; National Institute of Child Health and Human Development, 2000).

In the context of grammar instruction, implicit instruction means the absence of rule presentation and any activities that require the learner to attend to specific grammatical forms. Explicit instruction, on the other hand, involves either rule presentation (deductive) or activities that require the learner to attend to specific grammatical forms to make generalizations on their own (inductive; Norris & Ortega, 2000).

Meta-analyses suggest that combining implicit and explicit methods is most effective for vocabulary learning (Marulis & Neuman, 2010, 2013). For grammar, the results of meta-analyses of general second language instructions are mixed, with a tendency in favor of explicit training approaches (Goo et al., 2015; Kang et al., 2019; Norris & Ortega, 2000; Spada & Tomita, 2010). However, these meta-analyses included only a few primary studies with preschool and early elementary school children, so no clear statement can be made about this age group. On the contrary, primary studies suggest that a combination of implicit and explicit approaches may be effective in promoting the grammatical skills of elementary school SLLs (e.g., Stanat et al., 2012; Whittle & Lyster, 2016).

Important to note is that methods and approaches to foster vocabulary and grammar are not generally comparable. In terms of vocabulary, research differentiates between implicit and explicit methods that are implemented with a variety of tasks. For grammar, on the other hand, relatively little is

known about the concrete tasks themselves, rather than approaches that can be incorporated into specific tasks. For this reason, the following descriptions differ in their level of specificity. The vocabulary section describes the specific methods used in the vocabulary tasks, but the grammar section depicts more general approaches woven into the grammar tasks.

3.1. Overview of Methods for Vocabulary Training

A widely used implicit method for training vocabulary is storybook reading (Grøver et al., 2020; Lowman et al., 2018). Unknown target words appear in the story, which can be read aloud by the teacher or experimenter (e.g., Lenhart et al., 2018) or read by the children (e.g., Kelley & Kinney, 2017). Contextual cues facilitate the development of the word meanings without requiring an explicit explanation (Dujardin et al., 2021). Different studies indicated that storybook reading effectively supports vocabulary acquisition (Lenhart et al., 2018; Neuman et al., 2021).

Likewise, studies use a variety of explicit methods to foster vocabulary development. Usually, the focus is on child-friendly definitions that can be supported by pictures (Dixon et al., 2020). Evidence suggests that contextual knowledge about words is crucial for students to use them effectively (e.g., Vaahtoranta et al., 2018). Accordingly, studies that use explicit definitions usually combine them with different activities to deepen word comprehension. Thus, words are presented in sentences (e.g., Chung et al., 2022), or questions are asked (Dixon et al., 2020). Available research supports the effectiveness of both implicit and explicit methods (Language and Reading Research Consortium et al., 2017).

Because a combination of implicit and explicit methods seems to be most effective for word learning (Marulis & Neuman, 2010, 2013), most interventions integrate implicit as well as explicit methods (Grøver et al., 2020; Hadley et al., 2016; Zucker et al., 2021). Vadasy et al. (2015) define the target words before the reading and try to activate children's background knowledge by asking questions about the story. In different interventions, the target words are explained while reading, and images of the words are presented (Hadley et al., 2016; Neuman et al., 2021). Various activities can be implemented after reading, in which the children explain the words (Hadley et al., 2016) or answer questions about them (Neuman & Kaefer, 2018). Different mapping games can also be played (Heller et al., 2019; Lowman et al., 2018), or children fill out cloze tasks (Dixon et al., 2020; Zucker et al., 2021).

3.2. Overview of Approaches for Grammar Training

Grammar instructions can be distinguished by the type of focus that is paid to a target grammatical form. A distinction can be made between (i) *Focus-on-Meaning*, (ii) *Focus-on-Forms*, and (iii) *Focus-on-Form*. While *Focus-on-Meaning* requires the learner to focus solely on the meaning that is central within instruction, *Focus-on-Forms* directs the learner's attention only to the target grammatical form (e.g., past -ed in English) without integrating it into a communicative context. *Focus-on-Form*, however, integrates attention to both form and meaning by implementing form focus in meaningful communicative contexts (Ellis, 2001; Long, 1991). Thus, the *Focus-on-Form* approach is best suited for linking implicit and explicit instruction while at the same time taking into account usage-based acquisition models, which consider the communicative use of language as a basic prerequisite for language learning (Tomasello, 2003).

Planned *Focus-on-Form* instruction can be implemented using various implicit and explicit methods. Typically, implicit training is provided through a targeted modification of the input that learners receive during instruction. One way to modify the input to promote learning of a target form is to consider the distribution of representatives of the new target form in the input. According to the principle of *skewed input*, learners should first be exposed to a low-variance input containing a few types of a specific construction combined with a high token frequency in order to facilitate their abstraction of the form¹ (Boyd & Goldberg, 2009). Corpus analyses have shown that such a distribution resembles the input in natural speech (Goldberg et al., 2004; Wulff et al., 2009), leading to the argument that a skewed input might facilitate form learning processes. The results of studies on skewed input in language learning are mixed (e.g., Casenhiser & Goldberg, 2005; Pulido, 2024; Révész et al., 2014) and reveal a need for more classroom-based studies with morpho-syntactic target forms, focussing on young learners (for a review see Zhang & Mai, 2023). Nevertheless, the theoretical assumptions of skewed input seem promising and justify its implementation in a language intervention.

A further input modification approach is *context-optimization*, originally developed for students with developmental language disorders (DLDs) in the German-speaking area. It provides instructions and methods for making the target form more salient and thus drawing the learner's attention to it (Motsch, 2017). Hence, it is largely compatible with the *Focus-on-Form* approach because it also integrates form focus in meaningful communica-

1 While a type means the abstract representation of a structure or a word, a token refers to the individual occurrence of these in the written or oral language (Barber, 2009).

tive contexts. Implicit components of context-optimization include the use of contrasts, emphasizing the target morpheme marker in speech, and using symbols, vocal gestures, or writing to make the target form more salient. In addition, using the shortest possible target structure helps to eliminate linguistic distractors. Explicit elements include, for example, brief meta-linguistic reflections on the target structure embedded in meaningful activities and communication (Motsch, 2017). Intervention studies with preschool and early elementary school students with DLDs have demonstrated the superiority of context-optimization over conventional instruction in promoting inflectional targets (e.g., Motsch & Riehemann, 2008).

Another input modification that likely plays a particularly relevant role in learning the German gender/case system is the targeted design of the morphosyntactic environment. Since, with very few exceptions, there are no fully valid gender patterns in the German language (Wegener, 1995), and there is no one-to-one correspondence between the form and function of gender/case indicators (Kauschke, 2012), it may also be beneficial to specifically modify the linguistic environment of the nouns. In this way, the learner is provided with syntactic cues that, when combined, can indicate the gender of a noun. In their study with university students using artificial gender-case-like markers, Bebout and Belke (2017) found that a blocked presentation of nouns was significantly superior to a random presentation. Another study with preschool children in Germany suggests that specially designed children's books that present gender and case categories in such a structured way can support the acquisition of gender and case in young SLLs (Belke et al., 2020).

4. Intervention

As described above, there are a variety of methods and approaches for developing young learners' vocabulary and grammatical skills, as well as a variety of intervention studies that support their effectiveness. Before describing the SPEAK-intervention, we would like to highlight the research gaps that necessitate the development of a new intervention.

Given persistent achievement gaps in vocabulary and grammatical skills linked to students' SES and SLL status (e.g., Levine et al., 2020), there is an urgent need for theory-driven and evidence-based interventions that foster young learners' language skills. Many intervention studies focus on training new vocabulary (e.g., Kelley & Kinney, 2017) or vocabulary that appears in reading assignments (e.g., Dickinson et al., 2019) regardless of its relevance for students' learning and school life. Some studies (e.g., Neuman & Kaefer,

2018; Spencer et al., 2020) follow the three-tier framework of Beck et al. (2013), which categorizes words according to their frequency. However, no such list exists for the German language.

With regard to grammar interventions, there are generally few school-based interventions for young FLLs and SLLs that include purposeful and structured grammar training and that have been evaluated rigorously (Wehrhöfer et al., 2024). Furthermore, most grammar interventions are conducted in small groups (e.g., Heller et al., 2019; Phillips et al., 2021), which is not always feasible in schools due to staff shortages (for Germany see Kultusministerkonferenz, 2023).

In addition, to our knowledge, there is no empirically evaluated digital training program in the German-speaking area for either vocabulary or grammar that supports these language components in the early elementary school grades. This is a notable gap in the literature because the use of digital intervention formats offers many opportunities, such as interactive and multimodal features as well as individual adaptive learning progressions (Hammerschmidt-Snidarich et al., 2019; Lee & Aspiranti, 2023). Our intervention was designed to address these gaps in applied and basic research on language interventions by developing a language-focused intervention for students in the early elementary school years that trains school-focused vocabulary and grammar skills in a digital format.

4.1. Structure of the Intervention

The duration of the intervention is 15 weeks, which was deemed necessary for covering a range of vocabulary and grammatical skills relevant to German first graders. Vocabulary and grammar are divided into three modules of five weeks each and are taught twice weekly for 45 minutes. The first weekly session focuses on vocabulary, and the second on grammar, with the grammar sessions incorporating the target vocabulary whenever possible.

Target words change weekly, and the grammatical targets are trained for one to two weeks, depending on their complexity. The number of target words increases over time, starting with five words per week in the first module, seven words in the second, and ten in the third module. In this way, a total of 88 target words are trained. Nouns ($n = 44$, e.g., *Lamm* – ‘lamb’, *Strategie* – ‘strategy’), verbs ($n = 26$, e.g., *unterstreichen* – ‘underline’, *vergleichen* – ‘compare’), and adjectives ($n = 18$, e.g., *digital* – ‘digital’, *rechteckig* – ‘rectangular’) are practiced weekly. Approximately 50% of the target words are nouns. As recommended in research, the target words of each week match semantically,

so they can be introduced within a coherent story (e.g., Language and Reading Research Consortium et al., 2017). The last week of every module focuses on the repetition of words and grammar. In these units, the module's target words and grammatical targets are combined in different tasks, and no new content is introduced.

The intervention is implemented within the regular curriculum and conducted by the classroom teacher for the subject of German. Aiming at a standardized implementation, the teacher introduces each session supported by a manual. In the manual teachers find information about the theoretical background of the intervention, an overview of the target words and grammatical target forms, and a detailed description of the tasks. Students then complete the tasks independently, being equipped with a tablet and a headset. As German first graders do not receive any formal reading and writing instruction before entering elementary school and thus, are usually not yet able to read or write, all tasks are accompanied with audio that was recorded by a female adult.

The digital implementation allows for individual learning as students can work at their own pace and get the possibility to listen to every audio as often as they need. Contrary to an approach addressing the whole classroom, digital tasks ensure that every child gets a chance to answer each task on their own and to receive immediate feedback. Another advantage of digitalization is the implementation of adaptivity, which is integrated in vocabulary and grammar tasks. The adaptive structure allows for students who have a low accuracy rate at certain tasks to receive an additional in-depth task, addressing the mistakes made, before proceeding. Hence, students with a high accuracy rate, spend less time on the tasks as they tend to require less training.

Another advantage of the digital implementation is that students can complete the digital tasks individually at their own pace. Both the vocabulary and the grammar sessions end with two recurring questions asking students about their motivation and effort in completing the tasks. They can rate the questions on a seven-point Likert scale, indicated by seven colored thumbs that vary in color (from red to green) and orientation (from down to up), depending on the level of agreement. For example, a thumb indicating complete disagreement is red and points downward. A thumb indicating complete agreement is green and points upward.

4.2. Type of Tasks Within the Intervention

Sessions With Focus on Vocabulary

The first of the two weekly sessions focuses on vocabulary. Every session begins with a short introduction by the teacher (5–10 minutes) followed by individual learning on the tablet (35–40 minutes). As every vocabulary session follows a different theme (e.g., nature), the teacher activates the children’s prior knowledge by asking questions about their experiences with the theme (following Vadasy et al., 2015). The questions are specified in the manual. After that, children spend the rest of the time working on the digital tasks using their tablets and headsets. They work on different implicit, explicit, and combined tasks, which appear alternately. A brief overview of examples of these task formats is provided in Table 3. Here, we focus on tasks that occur frequently.

Table 3: Overview of Sample Digital Vocabulary Task Formats

Method	Task format
(i) Implicit	e.g., digital storybook reading
(ii) Explicit	e.g., audio-image-mapping games, listening tasks, vocabulary questions
(iii) Combination implicit/explicit	e.g., digital storybook reading with definitions of target words or cloze tasks

Every vocabulary session starts with (i) the implicit digital storybook reading. Children listen individually to the story on their tablet with headsets. The story contains the current week’s target words and grammatical forms. Typical length is 600 to 800 words. It is accompanied by different images that illustrate the narrative. As already mentioned, the combination of the implicit storybook reading with explicit activities is most effective. Accordingly, the story is enriched with different explicit elements described below.

In terms of (ii) explicit methods, we are using different audio-image-mapping games in which children match audio of the target word (following Heller et al., 2019) or word definition (Lowman et al., 2018) with the corresponding image from a pool of images of different target words. One example is a memory game where a pair consists of an image and the corresponding audio. Since contextual knowledge of words is important for a deep understanding of the word’s meaning (e.g., Vaahtoranta et al., 2018), students work on different tasks that foster their vocabulary depth. One task format is a listening task. Students see a picture of a target word and listen to a contextual example (e.g., Hadley et al., 2016). In another task format, students listen to questions presenting the target words in various contexts. While listening to the question, the target word is presented as a picture. They can choose the correct answer

by clicking on an image or audio (e.g., Dixon et al., 2020; Grøver et al., 2020). In this way, students learn more about the word's usage in different contexts as well as about e.g., their function or look.

As described above, we (iii) combine the implicit method of storybook reading with different explicit activities. We use definitions of the target words to introduce word meaning. Children listen to the story on their tablets with headsets. After the first appearance of a target word, the story stops automatically. Then, an image of the word is shown along with an explanation (Hadley et al., 2016). Afterward, the story continues. The combination of implicit storybook reading and explicit definitions of words enables students to learn both the meaning of words and their usage simultaneously. Following different studies, we also combined storybook reading with oral cloze tasks. When the story interrupts, students have to answer questions that involve the target words (e.g., Kelley & Kinney, 2017). The question appears orally and children can choose between two possible answers that are supported both visually and audibly.

Sessions With Focus on Grammar

The second session of each week is dedicated to grammar training. As with the vocabulary sessions, the digital grammar tasks are also preceded by a short introduction given by the teacher (10–15 minutes) following the teachers' intervention manual. The introduction aims to introduce the grammatical target by conducting a joint classroom activity such as a rhyme containing the target grammatical form or a blackboard game eliciting the grammatical target form. By integrating the grammatical targets into meaningful input and providing the students with corrective feedback (Ranta & Lyster, 2017), the functionality of the grammatical targets is emphasized. All classroom activities are accompanied by analog picture cards that teachers receive for that purpose. Most of these classroom activities are picked up in the digital tasks so that learners have already established some prior knowledge.

After the joint classroom introduction, the students individually work on various digital grammar tasks using their tablets and headsets (30–35 minutes). Analogously to the vocabulary tasks, the grammar part of the intervention also uses different digital task formats that incorporate the approaches described in section 3.2 (see Table 4 for an overview). All tasks have in common that they are integrated into meaningful contexts. Students' responses generally rely on the content. The focus on the grammatical form, however, is achieved by the modification of the input. Additionally, there are few tasks requiring the learners to explicitly select the correct grammatical form within

a phrase or a sentence that is meaningful to the learners. This procedure ensures a combination of meaning-focused and form-focused instruction (Focus-on-Form) including implicit and explicit tasks (Ellis, 2001; Long, 1991).

Table 4: Overview of Sample Digital Grammar Task Formats

Approach	Task format
(i) Skewed input	e.g., memory tasks, digital storybook reading
(ii) Modified morphosyntactic environment	e.g., digital storybook reading, cloze tasks
(iii) Contrasts	e.g., arrangement games, rhymes
(iv) Shortest target structure	e.g., elicitation tasks, digital storybook reading with questions
(v) Increasing the salience of target forms	e.g., true/false tasks, building fun nonsense sentences
(vi) Metacognition	e.g., reflection impetuses

An example of the implementation of (i) skewed input (Boyd & Goldberg, 2009) is the instruction on the gender pattern *Xe* in week 1. Here, the noun *Blume* (‘flower’) serves as an anchor and occurs very frequently in both the story and the following grammar tasks. Other nouns with the same gender pattern are gradually included in the tasks. To emphasize the reference function of gender (Binanzer & Wecker, 2020), the nouns are first presented with an inflected definite article and an attributive adjective (see 1), and in the further course of the session, with an inflected relative pronoun (see 2) and a personal pronoun (see 3).

- (1)

die

ART.DEF.SG.F.

NOM²

‘the blue flower’

blau-e

blue-SG.F.

NOM

Blume

flower
- (2)

die

ART.DEF.SG.F.

NOM

‘the blue flower that is beautiful’

blau-e

blue-SG.F.

NOM

Blume,

flower

die

REL.SG.F.

NOM³

schön

beautiful

ist

is
- (3)

die

ART.DEF.SG.F.

NOM

‘the blue flower, it is beautiful’

blau-e

blue-SG.F.

NOM

Blume,

flower

sie

PERS.SG.F.

NOM⁴

ist

is

schön

beautiful

Many tasks also provide a (ii) modified linguistic environment of the noun in form of syntactic cues (Bebout & Belke, 2017). For example, one task exposes the learners to different masculine and neuter nouns corresponding to the

2 ART = article, DEF = definite, SG = singular, F = feminine, NOM = nominative.

3 REL = relative pronoun.

4 PERS = personal pronoun.

trained gender patterns. The students listen to two cloze sentences dealing with the same noun. While one sentence uses the indefinite article, the other sentences requires a definite article. The learners have to choose the right attributive adjective in terms of content by selecting the corresponding audio. Then the two sentences are presented again directly one after the other (see 4). Here, the indefinite article alone does not give reliable information regarding the gender of the noun as *ein* can apply to both masculine and neuter nouns. However, together with the definite article (*der* for masculine nouns and *das* for neuter nouns) and the strongly inflected adjective (-er suffix for masculine nouns and -es suffix for neuter nouns), the gender of the noun becomes unambiguous.

- (4) *Ich bin ein schwarz-er Kater*
 I am ART.IND.SG.M. black-SG.M.NOM cat
 NOM⁵

‘I am a black cat.’

- sagt der schwarz-e Kater*
 says ART.DEF.SG.M. black-SG.M.NOM cat
 NOM

‘says the black cat’

Apart from that, most of the tasks training the accusative and the dative case contain (iii) contrasting structures to make the target case more salient (Motsch, 2017). For example, in one task, students listen to a prompt asking them to arrange objects on the screen in a specific way (see 5). The nominative and the accusative case are presented here in direct contrast.

- (5) *Das ist der Koffer*
 This is ART.DEF.SG.M. suitcase
 NOM

‘This is the suitcase.’

- Lege den Teddybären in den Koffer*
 Put the teddy bear in ART.DEF. suitcase
 SG.M.ACC⁶

‘Put the teddy bear in the suitcase.’

5 IND = indefinite, M = masculine.

6 ACC = accusative.

The principle of using (iv) the shortest target structure possible (Motsch, 2017) is also followed in many tasks. When the learners are supposed to answer a question eliciting an answer in a specific grammar form, not the whole answer sentence is presented but only the nominal phrase. Another feature making the grammatical target (v) more salient is emphasizing the morpheme marker of the inflected article in the speech (Motsch, 2017), either by the teacher or the digital program. In the context of the German case marking, the phonological difference between *den* and *dem* is particularly relevant. From the second module of the intervention, when first graders have acquired initial decoding skills, the emphasis on the morpheme marker is increased by an additional presentation of the written form of the articles with the <m> in *dem* highlighted in bold. Lastly, in order to address (vi) the learners' metacognition (Motsch, 2017), most grammar sessions include at least one impetus for reflection pointing explicitly to the target grammatical form and how it may differ from other forms. In terms of training the gender assignment, the students are provided with a learning strategy that helps them to identify nouns with a specific gender. These reflections are embedded in the exercises and are not detached from context.

5. Conclusion and Outlook

Given the great importance of language skills for educational success and social participation, early promotion of vocabulary and grammar is essential. Yet, there is little research on fostering young learners' vocabulary and grammatical skills simultaneously and systematically.

Several studies have focused on vocabulary interventions by using effective implicit and explicit activities (e.g., Zucker et al., 2021). Despite the multitude of intervention studies, no empirically evaluated digital program for the German context is available. Regarding grammar, approaches to foster grammatical skills of (second language) learning beginners exist (e.g., Bebout & Belke, 2017), although they are in large part not yet well researched. As vocabulary and grammatical skills are important prerequisites for later reading comprehension (Hjetland et al., 2020) and discrepancies in these skills are evident at the beginning of formal schooling (Levine et al., 2020), a systematic and empirically evaluated intervention is necessary.

The SPEAK-intervention combines empirical findings and theoretical assumptions for a digital-based intervention that can be used in regular classroom lessons and is suitable for children without reading and writing skills. We combine different implicit and explicit strategies for maximum promo-

tional impact. Digital administration enables individual and adaptive working as well as the integration of feedback.

The effectiveness of the presented intervention will be evaluated in an experimental pre-post-test design with two intervention groups and a control group. One intervention group receives training in their regular German lessons while the other intervention group receives additional training in all-day-schools in the afternoon. Accordingly, the interventions' effects on students' vocabulary and grammar skills as well as the potential of extended learning time in all-day schools and at home will be investigated.

At the end of the project SPEAK, a theory-based and empirically evaluated intervention will be available that focuses on promoting the prerequisites of important competencies that determine academic success. Against the background of educational inequalities already manifesting at the beginning of formal schooling, this intervention makes an important contribution to reducing those inequalities. In the future, the tasks can be used not only in the classroom context but also for individual students with difficulties. Furthermore, the intervention can be combined with other exercises that focus on further language skills for holistic support.

References

- Anstrom, K., DiCerbo, P., Butler, F., Katz, A., Millet, J., & Rivera, C. (2010). *A review of the literature on academic English: Implications for K-12 English language learners*. The George Washington University Center for Equity and Excellence in Education. <https://nysrti.org/files/resources/ells/litreviewacademicenglish.pdf>
- Bailey, A. L., & Butler, F. A. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document*. <https://cresst.org/wp-content/uploads/r611.pdf>
- Barber, A. (2009). Type/token. In S. Chapman & C. Routledge (Eds.), *Key ideas in linguistics and the philosophy of language* (pp. 246–247). Edinburgh University Press.
- Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy*, 54(1), 4–12. <https://doi.org/10.1598/jaal.54.1.1>
- Bebout, J., & Belke, E. (2017). Language play facilitates language learning: Optimizing the input for gender-like category induction. *Cognitive Research: Principles and Implications*, 2(1), 11. <https://doi.org/10.1186/s41235-016-0038-z>
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). The Guilford Press. <http://site.ebrary.com/lib/alltitles/Doc?id=10649559>
- Belke, E., Lehmden, F. von, & Müller-Brauers, C. (2020, September 3–5). *Specially designed children's books and songs for training the acquisition of German gender and case* [Poster Presentation]. Architectures and Mechanisms for Language Processing, Potsdam, Germany.
- Binanzer, A. (2017). *DaZ-Forschung: Vol. 17. Genus – Kongruenz und Klassifikation: Evidenzen aus dem Zweitspracherwerb des Deutschen*. [DaZ-Research: Vol. 17. Gender – congruence and classification: Evidence from the second language acquisition of German]. De Gruyter. <https://doi.org/10.1515/9783110548549>

- Binanzer, A., & Wecker, V. (2020). Lernergrammatik und Grammatikunterricht – Überlegungen zu natürlichen Erwerbssequenzen und didaktischer Progression [Learner grammar and grammar instruction – considerations on natural acquisition processes and didactic progression]. In M. Langlotz (Ed.), *Grammatikdidaktik: Theoretische und empirische Zugänge zu sprachlicher Heterogenität: Vol. 33. Thema Sprache – Wissenschaft für den Unterricht* (pp. 61–86). Schneider Verlag Hohengrehren GmbH.
- Boyd, J. K., & Goldberg, A. E. (2009). Input effects within a constructionist framework. *The Modern Language Journal*, 93(3), 418–429. <http://www.jstor.org/stable/40264096>
- Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, 8(6), 500–508. <https://doi.org/10.1111/j.1467-7687.2005.00441.x>
- Chung, K. K. H., Li, X., Lam, C. Y., Lam, C. B., Fung, W. K., & Lai, P. Y. (2022). The effects of bilingual reading program on Chinese children from low socioeconomic status families. *Early Education and Development*, 33(2), 204–218. <https://doi.org/10.1080/10409289.2021.1891803>
- Clahsen, H. (1984). Der Erwerb von Kasusmarkierungen in der deutschen Kindersprache [The acquisition of case markers in German children's language]. *Linguistische Berichte*, 89, 1–31.
- Dale, P. S., Paul, A., Rosholm, M., & Bleses, D. (2023). Prediction from early childhood vocabulary to academic achievement at the end of compulsory schooling in Denmark. *International Journal of Behavioral Development*, 47(2), 123–134. <https://doi.org/10.1177/01650254221116878>
- Dickinson, D. K., Nesbitt, K. T., Collins, M. F., Hadley, E. B., Newman, K., Rivera, B. L., Ilgez, H., Nicolopoulou, A., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Teaching for breadth and depth of vocabulary knowledge: Learning from explicit and implicit instruction and the storybook texts. *Early Childhood Research Quarterly*, 47, 341–356. <https://doi.org/10.1016/j.ecresq.2018.07.012>
- Dixon, C., Thomson, J., & Fricke, S. (2020). Evaluation of an explicit vocabulary teaching intervention for children learning English as an additional language in primary school. *Child Language Teaching and Therapy*, 36(2), 91–108. <https://doi.org/10.1177/0265659020925875>
- Dujardin, E., Auphan, P., Bailloud, N., Ecalle, J., & Magnan, A. (2021). Tools and teaching strategies for vocabulary assessment and instruction: A review. *Social Education Research*, 3(1), 34–66. <https://doi.org/10.37256/ser.3120221044>
- Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning*, 51(s1), 1–46. <https://doi.org/10.1111/j.1467-1770.2001.tb00013.x>
- Funcke, A., & Menne, S. (2023). *Kinderarmut in Deutschland: Factsheet* [Child poverty in Germany: Factsheet]. <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/factsheet-kinder-und-jugendarmut-in-deutschland>
- Goldberg, A. E., Casenhiser, D., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, 15(3), 289–316. <https://doi.org/10.1515/cogl.2004.011>
- Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning. In P. Rebuschat (Ed.), *Implicit and Explicit Learning of Languages: Vol. 48. Studies in Bilingualism* (pp. 443–482). John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.48.18goo>
- Grøver, V., Rydland, V., Gustafsson, J.-E., & Snow, C. E. (2020). Shared book reading in pre-school supports bilingual children's second-language learning: A cluster-randomized trial. *Child Development*, 91(6), 2192–2210. <https://doi.org/10.1111/cdev.13348>
- Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, 51(2), 181–198. <https://doi.org/10.1002/rq.130>
- Hammerschmidt-Snidarich, S. M., Edwards, L. M., Christ, T. J., & Thayer, A. J. (2019). Lev-eraging technology: A multi-component personalized system of instruction to teach sight words. *Journal of School Psychology*, 72, 150–171. <https://doi.org/10.1016/j.jsp.2018.12.005>

- Heller, M. C., Lervåg, A., & Grøver, V. (2019). Oral language intervention in Norwegian schools serving young language-minority learners: A randomized trial. *Reading Research Quarterly*, 54(4), 531–552. <https://doi.org/10.1002/rrq.248>
- Heppt, B., & Stanat, P. (2020). Development of academic language comprehension of German monolinguals and dual language learners. *Contemporary Educational Psychology*, 62, Article 101868. <https://doi.org/10.1016/j.cedpsych.2020.101868>
- Hjetland, H. N., Brinchmann, E. I., Scherer, R., Hulme, C., & Melby-Lervåg, M. (2020). Pre-school pathways to reading comprehension: A systematic meta-analytic review. *Educational Research Review*, 30, Article 100323. <https://doi.org/10.1016/j.edurev.2020.100323>
- Kang, E. Y., Sok, S., & Han, Z. (2019). Thirty-five years of ISLA on form-focused instruction: A meta-analysis. *Language Teaching Research*, 23(4), 428–453. <https://doi.org/10.1177/1362168818776671>
- Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen* [Child language acquisition in German]. De Gruyter. <https://doi.org/10.1515/9783110283891>
- Kelley, E. S., & Kinney, K. (2017). Word learning and story comprehension from digital storybooks. *Journal of Educational Computing Research*, 55(3), 410–428. <https://doi.org/10.1177/0735633116669811>
- Köpcke, K.-M., & Zubin, D. A. (2009). Genus [Gender]. In E. Hentschel & P. M. Vogel (Eds.), *Deutsche Morphologie* (pp. 132–154). Walter de Gruyter.
- Kultusministerkonferenz. (2023). *Lehrkräfteeinstellungsbedarf und -angebot in der Bundesrepublik Deutschland 2023–2035 – Zusammengefasste Modellrechnungen der Länder* [Teacher recruitment needs and supply in the Federal Republic of Germany 2023–2035 – Summarized model calculations of the federal states]. https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_238_Bericht_LEB_LEA_2023.pdf
- Language and Reading Research Consortium, Jiang, H., & Davis, D. (2017). Let's Know! Proximal impacts on prekindergarten through grade 3 students' comprehension-related skills. *The Elementary School Journal*, 118(2), 177–206. <https://doi.org/10.1086/694220>
- Lee, S. H., & Aspiranti, K. B. (2023). Using multimodal educational apps to increase the vocabulary of children with and without reading difficulties. *International Journal of Child-Computer Interaction*, 36, Article 100579. <https://doi.org/10.1016/j.ijcci.2023.100579>
- Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2018). Incidental vocabulary acquisition from listening to stories: A comparison between read-aloud and free storytelling approaches. *Educational Psychology*, 38(5), 596–616. <https://doi.org/10.1080/01443410.2017.1363377>
- Levine, D., Pace, A., Luo, R., Hirsh-Pasek, K., Golinkoff, R. M., Villiers, J. de, Iglesias, A., & Wilson, M. S. (2020). Evaluating socioeconomic gaps in preschoolers' vocabulary, syntax and language process skills with the Quick Interactive Language Screener (QUILS). *Early Childhood Research Quarterly*, 50(1), 114–128. <https://doi.org/10.1016/j.ecresq.2018.11.006>
- Limbird, C. K., Maluch, J. T., Rjosk, C., Stanat, P., & Merckens, H. (2014). Differential growth patterns in emerging reading skills of Turkish-German bilingual and German monolingual primary school students. *Reading and Writing*, 27(5), 945–968. <https://doi.org/10.1007/s11145-013-9477-9>
- Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign Language Research in Cross-Cultural Perspectives* (pp. 39–52). John Benjamins.
- Lowman, J., Stone Laura T, & Guo, J. (2018). Effects of interactive book reading for increasing children's knowledge of instructional verbs. *Communication Disorders Quarterly*, 39(4), 477–489. <https://doi.org/10.1177/152574011774563>

- Maaz, K., Artelt, C., Buchholz, S., Kuger, S., Kühne, S., Leerhoff, H., Schrader, J., Seeber, S., Arnoldt, B., Bach, M., Herrmann-Feichtenbeiner, V., Hoffmann, M., Kerst, C., Kohl, J., Kopp, K., Krug von Nidda, S., Lühe, J., Lochner, S., Malchin, A., ... Ziesmann, T. (2024). *Bildung in Deutschland 2024: Ein indikatorengestützter Bericht mit einer Analyse zu beruflicher Bildung*. <https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2024/pdf-dateien-2024/bildungsbericht-2024.pdf>, <https://doi.org/10.3278/6001820iw>
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300–335. <https://doi.org/10.3102/0034654310377087>
- Marulis, L. M., & Neuman, S. B. (2013). How vocabulary interventions affect young children at risk: A meta-analytic review. *Journal of Research on Educational Effectiveness*, 6(3), 223–262. <https://doi.org/10.1080/19345747.2012.755591>
- McElvany, N., & Artelt, C. (2009). Systematic reading training in the family: Development, implementation, and initial evaluation of the Berlin Parent-Child Reading Program. *Learning and Instruction*, 19(1), 79–95. <https://doi.org/10.1016/j.learninstruc.2008.02.002>
- McElvany, N., Ohle, A., El-Khechen, W., Hardy, I., & Cinar, M. (2017). Förderung sprachlicher Kompetenzen – Das Potential der Familiensprache für den Wortschatzerwerb aus Texten [Fostering language skills – the potential of the family language for vocabulary acquisition from texts]. *Zeitschrift Für Pädagogische Psychologie*, 31(1), 13–25. <https://doi.org/10.1024/1010-0652/a000189>
- Motsch, H.-J. (2017). *Kontextoptimierung: Evidenzbasierte Intervention bei grammatischen Störungen in Therapie und Unterricht* [Context-Optimization: Evidence-based intervention for grammatical disorders in therapy and teaching] (4th ed.). Ernst Reinhardt Verlag.
- Motsch, H.-J., & Riehemann, S. (2008). Effects of 'Context-Optimization' on the acquisition of grammatical case in children with specific language impairment: An experimental evaluation in the classroom. *International Journal of Language & Communication Disorders*, 43(6), 683–698. <https://doi.org/10.1080/13682820701794728>
- National Institute of Child Health and Human Development. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. National Reading Panel. <https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf>
- Neuman, S. B., & Kaefer, T. (2018). Developing low-income children's vocabulary and content knowledge through a shared book reading program. *Contemporary Educational Psychology*, 52, 15–24. <https://doi.org/10.1016/j.cedpsych.2017.12.001>
- Neuman, S. B., Samudra, P., & Danielson, K. (2021). Effectiveness of scaling up a vocabulary intervention for low-income children, pre-k through first grade. *The Elementary School Journal*, 121(3), 385–408. <https://doi.org/10.1086/712492>
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528. <https://doi.org/10.1111/0023-8333.00136>
- Passareta, G., Skopek, J., & van Huizen, T. (2022). Is social inequality in school-age achievement generated before or during schooling? A European perspective. *European Sociological Review*, 38(6), 849–865. <https://doi.org/10.1093/esr/jcac005>
- Phillips, B. M., Kim, Y.-S. G., Lonigan, C. J., Connor, C. M., Clancy, J., & Al Otaiba, S. (2021). Supporting language and literacy development with intensive small-group interventions: An early childhood efficacy study. *Early Childhood Research Quarterly*, 57, 75–88. <https://doi.org/10.1016/j.ecresq.2021.05.004>
- Pulido, M. F. (2024). Optimizing the input for learning of L2-specific constructions: The roles of Zipfian and balanced input, explicit rules and working memory. *Studies in Second Language Acquisition*, 46(2), 379–403. <https://doi.org/10.1017/S0272263124000081>
- Ranta, L., & Lyster, R. (2017). Form-focused instruction. In P. Garrett & J. M. Cots (Eds.), *The Routledge Handbook of Language Awareness* (pp. 40–56). Routledge.

- Révész, A., Sachs, R., & Hama, M. (2014). The effects of task complexity and input frequency on the acquisition of the past counterfactual construction through recasts. *Language Learning*, 64(3), 615–650. <https://doi.org/10.1111/lang.12061>
- Ronniger, P., Rißling, J.-K., Petermann, F., & Melzer, J. (2019). Entwicklung von sprachlichen Fähigkeiten und Basiskompetenzen des Spracherwerbs bei mehrsprachig im Vergleich zu einsprachig aufwachsenden Vorschulkindern [Development of language skills and basic language acquisition skills in multilingual compared to monolingual preschool children]. *Zeitschrift für Erziehungswissenschaft*, 22(1), 203–224. <https://doi.org/10.1007/s11618-018-0837-8>
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). ChildLex: A lexical database of German read by children. *Behavior Research Methods*, 47(4), 1085–1094. <https://doi.org/10.3758/s13428-014-0528-1>
- Schulz, P., & Grimm, A. (2018). The age factor revisited: Timing in acquisition interacts with age of onset in bilingual acquisition. *Frontiers in Psychology*, 9, Article 2732. <https://doi.org/10.3389/fpsyg.2018.02732>
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308. <https://doi.org/10.1111/j.1467-9922.2010.00562.x>
- Spencer, T. D., Moran, M., Thompson, M. S., Petersen, D. B., & Restrepo, M. A. (2020). Early efficacy of multitiered dual-language instruction: Promoting preschoolers' Spanish and English oral language. *AERA Open*, 6(1). <https://doi.org/10.1177/2332858419897886>
- Stanat, P., Becker, M., Baumert, J., Lüdtke, O., & Eckhardt, A. G. (2012). Improving second language skills of immigrant students: A field trial study evaluating the effects of a summer learning program. *Learning and Instruction*, 22(3), 159–170. <https://doi.org/10.1016/j.learninstruc.2011.10.002>
- Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, 38(8), 878–893. <https://doi.org/10.1037//0003-066x.38.8.878>
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard Univ. Press. <https://doi.org/38840>
- Ulrich, T. (2017). *Grammatikerwerb und grammatische Störungen im Kindesalter. Ergebnisse des Forschungsprojekts GED 4–9 und ihre Implikationen für sprachdiagnostische und -therapeutische Methoden* [Grammar acquisition and grammatical disorders in childhood. Results of the research project GED 4–9 and their implications for language diagnostic and therapeutic methods]. Universität zu Köln. https://kups.ub.uni-koeln.de/9011/1/Habilita-tion_Ulrich_OnlinePub.pdf
- Ulrich, T., Thater, S., & Mennicken, S. (2021). Kasusfähigkeiten mehrsprachiger Achtjähriger: Eine explorative Pilotuntersuchung in Regelgrundschulen [Case marking skills of multilanguage eight-years-olds: An exploartive pilot study in elementary schools]. *Logos*, 29(2), 84–95.
- Vaahantoranta, E., Suggate, S., Jachmann, C., Lenhart, J., & Lenhard, W. (2018). Can explaining less be more? Enhancing vocabulary through explicit versus elaborative storytelling. *First Language*, 38(2), 198–217. <https://doi.org/10.1177/0142723717737452>
- Vadasy, P. F., Sanders, E. A., & Nelson, J. R. (2015). Effectiveness of supplemental kindergarten vocabulary instruction for English learners: A randomized study of immediate and longer-term effects of two approaches. *Journal of Research on Educational Effectiveness*, 8(4), 490–525. <https://doi.org/10.1080/19345747.2015.1033120>
- Wanzek, J. (2014). Building word knowledge: Opportunities for direct vocabulary instruction in general education for students with reading difficulties. *Reading & Writing Quarterly*, 30(2), 139–164. <https://doi.org/10.1080/10573569.2013.789786>
- Wegener, H. (1995). *Die Nominalflexion des Deutschen – verstanden als Lerngegenstand* [The nominal inflection of German – understood as a learning objective]. De Gruyter. <https://doi.org/10.1515/9783110952230>

- Wehrhöfer, S., Dargiewicz, L., Ohle-Peters, A., Lauermann, F., & McElvany, N. (2024). *Target groups, characteristics, and efficacy of early school-based grammar interventions: A systematic literature review* [Manuscript in preparation]. Center for Research on Education and School Development, TU Dortmund University.
- Whittle, A., & Lyster, R. (2016). Focus on Italian verbal morphology in multilingual classes. *Language Learning*, 66(1), 31–59. <https://doi.org/10.1111/lang.12131>
- Wöllstein, A. (Ed.). (2022). *Die Grammatik. Struktur und Verwendung der deutschen Sprache. Sätze – Wortgruppe – Wort* [Grammar. Structure and use of the German language. Sentences – word groups – word]: Vol. 4. Der Duden in zwölf Bänden (10th ed.). Dudenverlag.
- Wulff, S., Ellis, N. C., Römer, U., Bardovi-Harlig, K., & Leblanc, C. J. (2009). The acquisition of tense-aspect: Converging evidence from corpora and telicity ratings. *The Modern Language Journal*, 93(3), 354–369. <https://doi.org/10.1111/j.1540-4781.2009.00895.x>
- Zhang, X., & Mai, C. (2023). Impact of constructional complexity and intralingual influence on the effectiveness of skewed input. *Language Teaching Research*, 27(5), 1217–1245. <https://doi.org/10.1177/1362168820981395>
- Zheng, H., Miao, X., Dong, Y., & Yuan, D.-C. (2023). The relationship between grammatical knowledge and reading comprehension: A meta-analysis. *Frontiers in Psychology*, 14, Article 1098568. <https://doi.org/10.3389/fpsyg.2023.1098568>
- Zucker, T. A., Carlo, M. S., Montroy, J. J., & Landry, S. H. (2021). Pilot test of the Hablemos Juntos Tier 2 academic language curriculum for Spanish-speaking preschoolers. *Early Childhood Research Quarterly*, 55, 179–192. <https://doi.org/10.1016/j.ecresq.2020.11.009>

Autor*innen

Dr. Christine Becks, wissenschaftliche Mitarbeiterin der AG Bildungsforschung und Co-Leitung des Forschungsbereichs „Kontextsensible Erforschung und Entwicklung des Schulsystems“. Forschungs- und Arbeitsschwerpunkte: Formen kontext- und kulturresponsiver Formen von Schulentwicklung; Angewandte Phänomenologie; International und vergleichende Forschung zu Bildungsreformen.

Dr. phil. Christine Becks
Fakultät für Bildungswissenschaften
Universität Duisburg-Essen
Universitätsstr. 2
45276 Essen
christine.becks@uni-due.de

Prof. Dr. Vera Busse, Professorin am Institut für Erziehungswissenschaft (IfE), Universität Münster, Arbeitsgruppe für Mehrsprachigkeit und Bildung. Forschungs- und Arbeitsschwerpunkte: sprachliche Bildung und Mehrsprachigkeit, Diversität im Unterricht, Literalitätsförderung, Lehrkräftebildung.

Prof. Dr. Vera Busse
Institut für Erziehungswissenschaft
Universität Münster
Georgskommende 33
48143 Münster
vbusse@uni-muenster.de

Dr. Susanne J. Czaja, wissenschaftliche Mitarbeiterin am Institut für Allgemeine Didaktik und Schulpädagogik an der Technischen Universität Dortmund und Koordinatorin des Inhaltsclusters „Schulentwicklung und Führung“ im Forschungsverbund „Schule macht stark – SchuMaS“. Forschungs- und Arbeitsschwerpunkte: Schulentwicklung an sozialräumlich benachteiligten Standorten; schulisches Führungshandeln; designbasierte sowie systemische Problemlösestrategien im schulischen Kontext.

Dr. phil. Susanne J. Czaja
Fakultät 12 – Erziehungswissenschaft, Psychologie und Bildungsforschung
Technische Universität Dortmund
Martin-Schmeißer-Weg 4–8
44227 Dortmund
susanne.czaja@tu-dortmund.de

Leonie Dargiewicz, M.A., wissenschaftliche Mitarbeiterin, Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Interventionsstudien, Sprachförderung und Sprachentwicklung, Einsatz digitaler Medien zur Kompetenzstärkung.

Leonie Dargiewicz, M.A.

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

leonie.dargiewicz@tu-dortmund.de

Dr. phil. Paul Fabian, bis Ende Februar 2024 wissenschaftlicher Mitarbeiter, Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund. Seit März 2024 akademischer Rat a. Z. am Lehrstuhl für Allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung, Ludwig-Maximilians-Universität München. Forschungs- und Arbeitsschwerpunkte: irreguläre Bildungsverläufe (insb. Klassenwiederholungen), herkunftsbedingte Ungleichheiten im Bildungswesen, längsschnittliche Bildungsforschung, Forschungs- und Analysemethoden.

Dr. phil. Paul Fabian

Ludwig-Maximilians-Universität München

Leopoldstr. 13

80802 München

paul.fabian@lmu.de

Dr. Katrin Gabler, wissenschaftliche Mitarbeiterin im Arbeitsbereich Sonderpädagogik – Sprachentwicklung im Fachbereich Erziehungswissenschaft und Psychologie, Freie Universität Berlin. Forschungs- und Arbeitsschwerpunkte: Diagnostik und Förderung sprachlicher Kompetenzen im Schulkontext, Lehrkräftebildung im Bereich Sprachförderung, Grammatikentwicklung im (Vor-)Schulalter.

Dr. Katrin Gabler

Arbeitsbereich Sonderpädagogik/Sprachentwicklung

FU Berlin

Fabeckstraße 35

14195 Berlin

katrin.gabler@fu-berlin.de

Kathrin Gade, M.A., abgeordnete Lehrkraft, Institut für Erziehungswissenschaft (IfE), Universität Münster, Arbeitsgruppe für Mehrsprachigkeit und Bildung. Forschungs- und Arbeitsschwerpunkte: Schreibförderung, multimediale Lernsettings, (digitalisierungsbezogene) Schulentwicklung, Lehrkräftebildung.

Kathrin Gade, M.A.

Institut für Erziehungswissenschaft

Universität Münster

Bispinghof 2

48143 Münster

gade@uni-muenster.de

Prof. Dr. Hanna Gaspard, bis Ende März 2025 Universitätsprofessorin für Educational Data Science, Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund. Seit April 2025 Universitätsprofessorin für Lehr-Lernforschung mit Schwerpunkt Diagnostik und Intervention an der Universität Konstanz. Forschungs- und Arbeitsschwerpunkte: Heterogenität in der Entwicklung von Lerneinstellungen und deren Auswirkungen auf Bildungs- und Karriereverläufe, Pädagogisch-psychologische Interventionen zur Förderung von Lerneinstellungen und Bildungserfolgen, Evaluation von Interventionen im Bildungskontext, Qualitätsvoller Unterricht und dessen Zusammenhänge mit Merkmalen der Lehrenden und Lernenden.

Prof. Dr. Hanna Gaspard
Geisteswissenschaftliche Sektion
Empirische Bildungsforschung
Universitätsstraße 10
Fach 45
D-78464 Konstanz
hanna.gaspard@uni-konstanz.de

Dr. Alyssa Laureen Grecu, Akademische Rätin (a. Z.), Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Soziale Schichtung und Bildungsungleichheiten, Lernen in formalen und informellen Bildungsumfeldern, Schulentfremdung, Bildungsverläufe.

Dr. Alyssa Laureen Grecu
Institut für Schulentwicklungsforschung (IFS)
TU Dortmund
Vogelpothsweg 78
44227 Dortmund
alyssa.grecu@tu-dortmund.de

Prof. Dr. Ilonca Hardy, Professorin für Erziehungswissenschaft mit Schwerpunkt Grundschulpädagogik/Empirische Bildungsforschung, Fachbereich Erziehungswissenschaften, Goethe-Universität Frankfurt am Main. Forschungs- und Arbeitsschwerpunkte: Adaptivität in Schule und Unterricht, sprachliche Bildung und Mehrsprachigkeit im Elementar- und Primarbereich, Entwicklung und Förderung naturwissenschaftlicher Kompetenzen.

Prof. Dr. Ilonca Hardy
Fachbereich Erziehungswissenschaften
Goethe-Universität Frankfurt am Main
Theodor-W.-Adorno-Platz 6
60323 Frankfurt am Main
hardy@em.uni-frankfurt.de

Dr. Sofie Henschel, stellvertretende wissenschaftliche Leiterin am Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin. Forschungs- und Arbeitsschwerpunkte: Entwicklung, Implementation und Evaluation von Sprach- und Lesefördermaßnahmen, datengestützte Unterrichtsentwicklung, Determinanten und Effekte lehr- und lernbegleitender Emotionen.

Dr. Sofie Henschel

Humboldt-Universität zu Berlin

Institut zur Qualitätsentwicklung im Bildungswesen

Unter den Linden 6

10099 Berlin

sofie.henschel@iqb.hu-berlin.de

Prof. Dr. Birgit Heppt, Professorin für Empirische Bildungsforschung im schulischen Kontext, Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Diagnostik und Förderung sprachlicher Kompetenzen, Professionalisierung von Lehr- und Fachkräften, Erfassung und Wirksamkeit von Unterrichtsprozessen.

Prof. Dr. Birgit Heppt

Institut für Schulentwicklungsforschung (IFS)

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

birgit.heppt@tu-dortmund.de

Prof. Dr. Esther Dominique Klein, Professorin für Allgemeine Didaktik und Schulpädagogik an der Technischen Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Schulentwicklungsforschung, insbesondere Führung und Schulkultur, Schulentwicklung an sozialräumlich benachteiligten Standorten und Rolle der Unterstützungssysteme; Educational Governance; International Vergleichende Bildungsforschung.

Prof. Dr. Esther Dominique Klein

Fakultät 12 – Erziehungswissenschaft, Psychologie und Bildungsforschung

Technische Universität Dortmund

Martin-Schmeißer-Weg 4–8

44227 Dortmund

dominique.klein@tu-dortmund.de

Dr. Ruben Kleinkorres, wissenschaftlicher Mitarbeiter, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Wohlbefinden im schulischen Kontext, kompetenzbedingende Faktoren.

Dr. Ruben Kleinkorres

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

ruben.kleinkorres@tu-dortmund.de

Anne Kosubek, M.A., wissenschaftliche Mitarbeiterin, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund, Forschungs- und Arbeitsschwerpunkte: die Transmission motivationaler und emotionaler Aspekte von Lehrkräften auf die Motivation von Lernenden, Unterrichtsqualität und die Motivation von Lehrkräften und Lernenden.

Anne Kosubek, M.A.

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

anne.kosubek@tu-dortmund.de

Prof. Fani Lauermann, Ph.D., Universitätsprofessorin für Empirische Bildungsforschung und Pädagogische Psychologie, Bonner Zentrum für Lehrerbildung (BZL), Universität Bonn, Forschungs- und Arbeitsschwerpunkte: Akademische Motivation und Leistung, Bildungsverläufe, Bildungs- und Berufswahl, Lehrkraftmotivation und Unterrichtsqualität.

Prof. Fani Lauermann, Ph.D.

Bonner Zentrum für Lehrerbildung

Universität Bonn

Poppelsdorfer Allee 15

53115 Bonn

fani.lauermann@uni-bonn.de

Denise Löfflad, wissenschaftliche Mitarbeiterin in der Arbeitsgruppe „Sprache und KI in der Bildung“, Leibniz-Institut für Wissensmedien (IWM) Tübingen. Forschungs- und Arbeitsschwerpunkte: Entwicklung und Anwendung computerlinguistischer Methoden zur Analyse sprachlicher Komplexität im Erst- und Zweitspracherwerb, Digitalisierung im schulischen Kontext.

Denise Löfflad

Arbeitsgruppe „Sprache und KI in der Bildung“

Leibniz-Institut für Wissensmedien (IWM)

Schleichstr. 6

72076 Tübingen

d.loefflad@iwm-tuebingen.de

PD Dr. Ramona Lorenz, seit 2009 wissenschaftliche Mitarbeiterin, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Empirische Schul- und Unterrichtsforschung, Digitale Medien in Schule und Unterricht, Leseleistungen von Schüler*innen.

PD Dr. Ramona Lorenz

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

ramona.lorenz@tu-dortmund.de

Dr. Ulrich Ludewig, wissenschaftlicher Mitarbeiter, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkt: Lesekompetenz, Large-Scale-Assessment, Testentwicklung.

Dr. Ulrich Ludewig
Institut für Schulentwicklungsforschung
TU Dortmund
Vogelpothsweg 78
44227 Dortmund
ulrich.ludewig@tu-dortmund.de

Prof. Dr. Nele McElvany, Geschäftsführende Direktorin des Instituts für Schulentwicklungsforschung (IFS) und Prorektorin Forschung der Technischen Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Lehr-/Lernforschung im schulischen Kontext; Kompetenzen von Lehrkräften und Unterrichtsqualität; Entwicklung und Förderung von Schriftsprachkompetenzen; Bildung und Migration; Pädagogisch-psychologische Diagnostik und Large-Scale Assessments.

Prof. Dr. Nele McElvany
Institut für Schulentwicklungsforschung
TU Dortmund
Vogelpothsweg 78
44227 Dortmund
nele.mcelvany@tu-dortmund.de

Prof. Dr. Detmar Meurers, Leiter der Arbeitsgruppe „Sprache und KI in der Bildung“, Leibniz-Institut für Wissensmedien (IWM) Tübingen. Forschungs- und Arbeitsschwerpunkte: Erforschung und Entwicklung KI-gestützter Methoden für den Bildungsbereich, insbesondere von Intelligenten Tutorsystemen (ITS), sprachliche Modellierung des Zweitspracherwerbs, Entwicklung effektiver Lehr-Lernmethoden durch die Verbindung von Computerlinguistik und empirischer Bildungsforschung.

Prof. Dr. Detmar Meurers
Arbeitsgruppe „Sprache und KI in der Bildung“
Leibniz-Institut für Wissensmedien (IWM)
Schleichstr. 6
72076 Tübingen
d.meurers@iwm-tuebingen.de

Katharina Molitor, M.A., wissenschaftliche Mitarbeiterin, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Reduktion von Bildungsungleichheit, Mediation sozialer Herkunftseffekte, Einfluss motivationaler Faktoren auf den Bildungserfolg.

Katharina Molitor, M.A.
Institut für Schulentwicklungsforschung
TU Dortmund
Vogelpothsweg 78
44227 Dortmund
katharina.molitor@tu-dortmund.de

Sira Neuhaus, M.A., wissenschaftliche Mitarbeiterin, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Ethnische Ungleichheit, insbesondere die Untersuchung der Rolle von Schule und individuellen Faktoren in der wahrgenommenen Diskriminierung von Schüler*innen mit Migrationshintergrund und den daraus resultierenden Konsequenzen.

Sira Neuhaus, M.A.

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44225 Dortmund

sira.neuhaus@tu-dortmund.de

PD Dr. phil. nat. Annika Ohle-Peters, Akademische Rätin am Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Unterrichtsqualität; professionelle Kompetenzen von Lehrenden; Interventionsstudien; multikriteriale Zielerreichung in der Grundschule.

PD Dr. Annika Ohle-Peters

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

annika.ohle-peters@tu-dortmund.de

Prof. Dr. Hannes Schröter, Universitätsprofessor für Kognition und Lernen Erwachsener, FernUniversität in Hagen und Leiter der Abteilung „Lehren, Lernen, Beraten“, Deutsches Institut für Erwachsenenbildung Leibniz-Zentrum für Lebenslanges Lernen e.V. (DIE), Bonn, Forschungs- und Arbeitsschwerpunkte: Professionelle Kompetenzen von Lehrkräften, Lernvoraussetzungen und Lernentwicklungen Erwachsener.

Prof. Dr. Hannes Schröter

Deutsches Institut für Erwachsenenbildung

Heinemannstraße 12–14

53175 Bonn

schroeter@die-bonn.de

Dr. Justine Stang-Rabrig, wissenschaftliche Mitarbeiterin, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Individuelle, soziale und institutionelle Bedingungsfaktoren des multikriterialen Schulerfolgs (Kognitive Kompetenzen, Motivation, Wohlbefinden).

Dr. Justine Stang-Rabrig

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

justine.stang-rabrig@tu-dortmund.de

Sebastian N. T. Vogel, M.Sc., wissenschaftlicher Mitarbeiter, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Heterogenität im Bildungswesen; Individuelle, soziale und institutionelle Bedingungsfaktoren des Schulerfolgs; Übergänge im Schulsystem.

Sebastian N. T. Vogel

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 7

44227 Dortmund

sebastian2.vogel@tu-dortmund.de

Svenja Wehrhöfer, M.Ed., wissenschaftliche Mitarbeiterin, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Interventionsstudien, Sprachentwicklung und Sprachförderung, Soziale Ungleichheit.

Svenja Wehrhöfer, M.Ed.

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

svenja.wehrhoefer@tu-dortmund.de

Dr. Sittipan Yotyodying, wissenschaftlicher Mitarbeiter, Institut für Schulentwicklungsforschung (IFS), Technische Universität Dortmund. Forschungs- und Arbeitsschwerpunkte: Schulisches Elternengagement, partnerschaftliche Zusammenarbeit von Schule und Eltern, schulische Medienbildungsforschung und digitale Interventionsforschung.

Dr. Sittipan Yotyodying

Institut für Schulentwicklungsforschung

TU Dortmund

Vogelpothsweg 78

44227 Dortmund

sittipan.yotyodying@tu-dortmund.de