

Socioeconomic Disparities and *De-Tracking*
Reforms: Students' Educational Aspirations
as Well as Socioeconomic Composition Effects
on Student Achievement

Dissertation
zur Erlangung des Doktorgrades
der Philosophischen Fakultät
der Christian-Albrechts-Universität zu Kiel
vorgelegt von

Marlen Holtmann

Kiel, Januar 2024

Erstgutachter: Prof. Dr. Oliver Lüdtkke

Zweitgutachter: Prof. Dr. Olaf Köller

Tag der mündlichen Prüfung: 03.05.2024

Durch den zweiten Prodekan, Prof. Dr. Jörg Kilian

zum Druck genehmigt am: 28.06.2024

Contents

1	Introduction	1
2	Theoretical, Empirical, and Methodological Background	4
2.1	Tracking Practices and Socioeconomic Disparities	4
2.1.1	Introduction to Tracking Practices With a Focus on the German Context	4
2.1.2	Between-School Tracking Practices and Student Achievement .	6
2.1.3	Between-School Tracking Practices and Students' Educational Aspirations	9
2.1.4	Between-School Tracking Practices and Socioeconomic Disparities in Student Achievement and in Students' Educational Aspirations	10
2.1.5	Between-School Tracking Practices and the Socioeconomic Composition of the Student Body	15
2.1.6	The Relationship Between Socioeconomic Composition and Student Achievement	16
2.2	Socioeconomic Disparities and <i>De-Tracking</i> Reforms in Some German States	18
2.2.1	Overview on Between-School Tracking Practices in Germany in Recent Years	18
2.2.2	The <i>De-Tracking</i> Reforms of Some German States	20
2.2.3	Socioeconomic Composition in the Course of the <i>De-Tracking</i> Reforms	23

CONTENTS

2.2.4	Socioeconomic Composition Effects on Student Achievement in the Course of the <i>De-Tracking</i> Reforms	25
2.2.5	Social Disparities in Students' Educational Aspirations in the Course of the <i>De-Tracking</i> Reforms	29
2.3	Estimating the Effects of the <i>De-Tracking</i> Reforms	32
3	Objectives of the Dissertation	39
4	Study 1: Socioeconomic Composition Effects and <i>De-Tracking</i> Reforms	42
4.1	Introduction	43
4.2	Theoretical Background	45
4.2.1	Tracking, SEC and <i>De-Tracking</i> Reforms	45
4.2.2	Tracking and SEC Effects on Student Achievement	47
4.2.3	<i>De-Tracking</i> Reforms and SEC Effects on Student Achievement .	48
4.3	Materials and Methods	52
4.3.1	Participants	52
4.3.2	Measures	53
4.3.3	Data Analysis	55
4.4	Results	57
4.4.1	RQ1: SEC in Different School Types	57
4.4.2	RQ2: SEC Effects in Different School Types	58
4.4.3	RQ3: Mediation of SEC Effects by the Level of Educational Expectations	60
4.5	Discussion	62
4.5.1	Limitations and Future Research	64
4.5.2	Conclusions	65
5	Study 2: <i>De-Tracking</i> Reforms, Educational Aspirations, and SES	67
5.1	Introduction	68
5.2	Theoretical Background	70
5.2.1	Students' Educational Aspirations and Their Socioeconomic Backgrounds	70
5.2.2	Students' Educational Aspirations in Tracked School Systems .	72

CONTENTS

5.2.3	<i>De-tracking</i> Reforms in Germany	74
5.2.4	Students' Educational Aspirations and <i>De-tracking</i> Reforms in Germany	76
5.2.5	Present Study	77
5.3	Materials and Methods	78
5.3.1	Data	78
5.3.2	Measures	81
5.3.3	Analyses	82
5.4	Results	85
5.5	Discussion	93
5.5.1	Limitations and Future Research	96
5.6	Conclusion	97
6	Conclusion	99
6.1	Integration of the Key Findings Into the State of Research	99
6.2	Limitations and Methodological Evaluation	106
6.2.1	Data Basis	106
6.2.2	Grouping States and Estimating Effects of <i>De-tracking</i> Reforms .	110
6.3	Implications	114
6.3.1	For Educational Policy	114
6.3.2	For Future Research	117
	References	121
	Appendix A Socioeconomic Composition Effects and <i>De-Tracking</i> Reforms	153
A.1	Tables	154
A.2	Authors' contributions	155
	Appendix B <i>De-Tracking</i> Reforms, Aspirations, and SES	156
B.1	Tables	157
B.2	Authors' contributions	158

List of Figures

2.1	School Structure in Germany Before and After <i>De-Tracking</i> Reforms . . .	21
2.2	SEC Effects on Student Achievement (Mediation Model)	27
2.3	Identification of the Difference-in-Differences Estimator	36
4.1	Mediation Model of SEC Effects on Student Achievement	49
5.1	Study Design: Tracking Practices in States With <i>De-tracking</i> Reforms and States Without Reforms	75
5.2	Trends in Students' Educational Aspirations	86
5.3	Trends in Students' Educational Aspirations According to Socioeconomic Groups.	88

List of Tables

4.1	Descriptive Statistics of Demographic Variables in the States With <i>De-Tracking</i> Reforms.	53
4.2	RQ1: SEC in Different School Types in the States With <i>De-Tracking</i> Reforms.	57
4.3	RQ2: SEC Effects in Different School Types in the States With <i>De-Tracking</i> Reforms.	59
4.4	RQ3: Mediation of SEC Effects by the Level of Educational Expectations – Comprehensive Schools	61
5.1	Descriptive Statistics of Dependent and Independent Variables and Sample Characteristics in States With <i>De-Tracking</i> Reforms and States Without <i>De-Tracking</i> Reforms.	80
5.2	Difference-in-Differences Analyses: Effects of <i>De-Tracking</i> Reforms on Students’ Educational Aspirations	87
5.3	Difference-in-Differences Analyses: Effects of <i>De-Tracking</i> Reforms According to Socioeconomic Groups.	89
5.4	Regression Analyses: Students’ Educational Aspirations on Their Socioeconomic Backgrounds.	92
A.1	RQ4: Mediation of SEC Effects by the Level of Educational Expectations in the States with Reforms – Comprehensive Schools (2).	154
B.1	Mathematical Achievement Between 2012 and 2018.	157

Acknowledgements

The dissertation accompanied me during an intensive and fascinating phase of my life. I mainly worked on it during my time as a research fellow at the Institute for Educational Quality Improvement (IQB) in Berlin, but also at the European Atlantic coast, in Kiel, California, and Hamburg. Our children were born during this time.

First, I would like to deeply thank my supervisor Oliver Lüdtke. His thoughtful feedback improved my work. My thanks also go to Oliver Lüdtke and Olaf Köller for reviewing this dissertation. Furthermore, I am very grateful to Petra Stanat whose passion for high quality educational research is an inspiration. I learned a lot from the exchange with her and her feedback. Petra Stanat and Camilla Rjosk initialized this dissertation project.

It is Camilla to whom I would like to express my very special thanks. She fueled my interest in educational research early during my studies and supervised me with dedication until the end of my time as a PhD student. Her dedication is evident in her never-ending curiosity, focused conversations and really helpful feedback, which have contributed significantly to my work and my passion as an educational researcher.

In addition, I would also like to thank the co-authors of the journal articles of this dissertation: Camilla Rjosk, Oliver Lüdtke, Petra Stanat, and Malte Jansen for the valuable and stimulating discussions and feedback.

I really enjoyed my time at the IQB Berlin. For their tireless willingness to share their methodological knowledge, but also for their support in handling the data, I would like to thank Aleksander Kocaj, Benjamin Becker, Christin Rüdiger, Karoline

Acknowledgements

Sachse, Sebastian Weirich, and Stefan Schipolowski; Sarah Lenz and Jenny Kölm for the enriching exchange. My colleagues Benjamin Becker, Jenny Kölm, and Lena Külker made this special time even more fun – thanks a lot!

During my stay at UC Berkeley, I experienced an extraordinary kindness and openness, especially from Allison Bradford, Libby Gerard, Marcia Linn, and Weiying Li, for which I am very grateful. I enjoyed their curiosity in new thoughts, ideas, and methods, and I am very happy to be part of their team.

Since last year, I am part of the Research and Analysis team at IEA Hamburg and I would like to thank my new colleagues Alec Kennedy, Andrés Strello, Aysel Gojayeva, Rolf Strietholt, and Sabine Meinck for their feedback; and especially Aysel for making me feel good from day one.

I would also like to thank Joschka Wanner for encouraging my efforts in econometrics.

I am deeply grateful for the extensive support of my family, for whom education is of the highest value. Above all, I would like to thank Hendrik, my co-conspirator and inspiration, who goes on adventures with me. He never questions my potential or gets tired of listening to my ideas or sharing his – finding the balance to challenge and encourage me to make the most out of this dissertation, life, and me.

Zusammenfassung

Die Sekundarschulsysteme von Ländern unterscheiden sich unter anderem darin, wie sie mit Leistungsunterschieden von Schüler:innen umgehen (M. Becker et al., 2017; Chmielewski, 2014; Terrin & Triventi, 2023). Deutschland ist durch eine vorgeblich leistungsbezogene Zuweisung zu hierarchisch gegliederten Schularten geprägt, auch als *between-school tracking practices* bezeichnet (vgl. Baumert et al., 2006a; M. Becker et al., 2017; Chmielewski, 2014; Maaz et al., 2008). Traditionell umfasste das hierarchisch gegliederte Sekundarschulsystem in Deutschland drei Schularten: Hauptschulen, Realschulen und Gymnasien. Diese Schularten richteten sich an Schüler:innen mit unterschiedlichem Leistungsniveau und führten zu verschiedenen Schulabschlüssen (z.B. Buchmann & Dalton, 2002; Buchmann & Park, 2009; Maaz et al., 2008; Parker et al., 2016). Die Schulart, die ein:e Schüler:in besucht, prägt wiederum entscheidend den Bildungsweg und die Chancen auf dem Arbeitsmarkt (Buchmann & Dalton, 2002; Parker et al., 2016).

Seit Jahrzehnten werden Unterschiede in den sozioökonomischen Disparitäten zwischen Ländern im Hinblick auf verschiedene *tracking practices*, also die Art der leistungsbezogenen Gruppierung von Schüler:innen, diskutiert (Baumert et al., 2006a; Chmielewski, 2014; Dämmrich & Triventi, 2018; Duru-Bellat & Suchaut, 2005; Hanushek & Wößmann, 2006; Maaz et al., 2008; Skopek & Passaretta, 2021; Skopek et al., 2019; Strello et al., 2021; Terrin & Triventi, 2023; Van de Werfhorst & Mijs, 2010; Van de Werfhorst, 2019). Auch aufgrund der im internationalen Vergleich überdurchschnittlichen sozioökonomischen Disparitäten wird die Debatte um die *tracking practices* in Deutschland auf Basis von internationalen und nationalen

großangelegten vergleichenden Schulleistungsstudien wie der PISA-Studie oder den IQB-Ländervergleichsstudien und IQB-Bildungstrends intensiv geführt (siehe z.B. Kuhl et al., 2016; Mahler & Kölm, 2018; Müller & Ehmke, 2016; Niemietz et al., 2023; OECD, 2019). Sozioökonomische Disparitäten werden in Systemen mit *between-school tracking practices* unter anderem durch primäre, sekundäre und tertiäre Herkunftseffekte erklärt, die sich maßgeblich im Übergang von der Grundschule zur weiterführenden Schule manifestieren (Baumert et al., 2006a; Boudon, 1974; Esser, 2016; Gresch, 2012; Maaz et al., 2008). *Between-school tracking practices* führen zudem zu schulartspezifischer sozioökonomischer Segregation, die wiederum zu sozioökonomischen Disparitäten beitragen könnte (Baumert et al., 2006a; Maaz et al., 2014; Strello et al., 2022; Van de Werfhorst & Mijs, 2010).

In einigen deutschen Bundesländern wurden in den vergangenen fünfzehn Jahren *de-tracking* Reformen implementiert, die das mehrgliedrige Sekundarschulsystem in ein Zwei-Wege-Modell überführten (vgl. Hurrelmann, 2013). Durch Umwandlung oder Zusammenlegung von Haupt-, Real- und Gesamtschulen reduzierten diese Bundesländer im Wesentlichen die nicht-gymnasiale Schullandschaft von mehreren Schularten auf eine Gesamtschule. Die reformierte Gesamtschule soll den Schüler:innen die Möglichkeit bieten, alle Schulabschlüsse zu erreichen. Zu den Zielen dieser *de-tracking* Reformen zählte die Reduktion von sozioökonomischen Disparitäten, unter anderem durch die leistungsbezogene Durchmischung der Schülerschaft in der Gesamtschule (Baumert et al., 2013, 2017).

Bisherige Forschung zu den Zusammenhängen zwischen *tracking practices* und sozioökonomischen Disparitäten basieren vornehmlich auf Vergleichen zwischen unterschiedlichen Ländern. Dabei werden in der Regel entweder Unterschiede in den sozioökonomischen Disparitäten zwischen Ländern mit verschiedenen *tracking practices* beschrieben (vgl. Dämmrich & Triventi, 2018; Duru-Bellat & Suchaut, 2005; Schütz et al., 2008; Skopek & Passaretta, 2021; Skopek et al., 2019; Van de Werfhorst & Mijs, 2010) oder Veränderungen in den sozioökonomischen Disparitäten von der Grundschule zur weiterführenden Schule zwischen Ländern mit verschiedenen *tracking practices* verglichen (z.B. Ammermüller, 2005; Hanushek & Wößmann, 2006;

Strello et al., 2021). Ein weiterer Ansatz ist der Vergleich von Ländern mit *de-tracking* Reformen und Ländern ohne entsprechende Reformen (z.B. Van de Werfhorst, 2018, 2019). Ländervergleichende Ansätze gehen mit einem gewissen Maß an unbeobachteten Unterschieden einher, die die Schätzung der Auswirkungen von verschiedenen *tracking practices* beeinträchtigen könnten (Piopiunik, 2021; Woessmann, 2016). Trotz des Bildungsföderalismus in Deutschland sorgt die Kultusministerkonferenz (KMK) für eine gewisse Harmonisierung der Bildungspolitik zwischen den Bundesländern, um vergleichbare Bildungsstandards und -abschlüsse zu gewährleisten (KMK, 2020). Vergleiche zwischen deutschen Bundesländern sollten daher mit einem geringeren Ausmaß an unbeobachteten Unterschieden verbunden sein als Vergleiche zwischen Ländern und damit eine präzisere Schätzung der Zusammenhänge zwischen *tracking practices* und sozioökonomischen Disparitäten ermöglichen (siehe auch Piopiunik, 2021). Darauf aufbauend untersucht diese Dissertation die Einführung von *de-tracking* Reformen in einigen der deutschen Bundesländer und vergleicht die Entwicklung von sozioökonomischen Disparitäten zwischen diesen Bundesländern und denen ohne *de-tracking* Reformen vor und nach der Einführung dieser Reformen (vgl. Piopiunik, 2021). Dadurch können die mit der vorliegenden Dissertation erarbeiteten Ergebnisse den Forschungsstand zu den Zusammenhängen zwischen *tracking practices* und sozioökonomischen Disparitäten erweitern.

Konkret untersucht diese Dissertation zum einen die bisher unbeantwortete Frage, ob sich die sozioökonomische Zusammensetzung der Schülerschaft und ihre Zusammenhänge mit der Leistung von Schüler:innen durch die *de-tracking* Reformen tatsächlich verändert haben. Zahlreiche Studien konnten Effekte der sozioökonomischen Komposition auf die Leistung von Schüler:innen aufzeigen, also dass Schüler:innen in Klassen oder Schulen mit einem niedrigeren durchschnittlichen sozioökonomischen Status schlechtere Leistungen erbrachten als vergleichbare Schüler:innen in einer Klasse oder Schule mit einem höheren durchschnittlichen sozioökonomischen Status (z.B. Opdenakker & Van Damme, 2001; Palardy, 2013; Sciffer et al., 2022; Tan et al., 2023; Van Ewijk & Slegers, 2010). Aufgrund des Zusammenhangs zwischen der Leistung und den sozioökonomischen Hintergründen

der Schüler:innen kann eine durch die *de-tracking* Reformen induzierte leistungsbezogene Mischung der Schülerschaft auch eine Veränderung der sozioökonomischen Zusammensetzung bedeuten. Zu der Erklärung von sozioökonomischen Kompositionseffekten scheinen Interaktionen zwischen Klassenkamerad:innen beizutragen (vgl. Baumert et al., 2006a; Harker & Tymms, 2004; Palardy, 2013; Van Ewijk & Sleegers, 2010), zum Beispiel über das Teilen von Normen, Werten und Erwartungen, die das Lernen erleichtern (z.B. Tenenbaum et al., 2020). Aufgrund der veränderten Schulabschlussaussichten an den Gesamtschulen im Vergleich zu den Haupt- und Realschulen vor den Reformen wurde daher auch untersucht, ob der Anteil der Schüler:innen, von denen erwartet wird, die allgemeine Hochschulreife zu erreichen – als Indikator für geteilte Normen zwischen Schüler:innen – zur beabsichtigten Reduktion der sozioökonomischen Kompositionseffekte beigetragen hat. Die Analysen des repräsentativen IQB-Ländervergleichs 2009 (Köller et al., 2011) vor den *de-tracking* Reformen und des IQB-Bildungstrends 2015 (Stanat et al., 2018) nach den *de-tracking* Reformen zeigen, dass die sozioökonomischen Kompositionseffekte auf die Leistung der Schüler:innen an Gesamtschulen nach den *de-tracking* Reformen auf einem vergleichbaren Niveau lagen, wie sie an Realschulen vor den Reformen beobachtet wurden. Ferner deuten die Ergebnisse darauf hin, dass die sozioökonomischen Kompositionseffekte nicht durch das Niveau der Abschlusserwartungen der Schüler:innen in einer Klasse mediiert wurden.

In Systemen mit *between-school tracking practices* werden neben der Schülerleistung die Abschlussaspirationen der Schüler:innen maßgeblich von dem Besuch der Schulart geprägt (z.B. Buchmann & Dalton, 2002; Buchmann & Park, 2009; Maaz et al., 2008; Zimmermann, 2020). Die Abschlussaspirationen der Schüler:innen sind – in Zusammenspiel mit der Leistung – eine wichtige Determinante für den Erwerb von Bildungsabschlüssen und den späteren Bildungserfolg (Beal & Crockett, 2010; Bozick et al., 2010; Sewell et al., 1970; Sewell et al., 1969). Daher wurde zum anderen im Rahmen dieser Dissertation untersucht, ob die *de-tracking* Reformen mit dem intendierten Anstieg der Abschlussaspirationen der Schüler:innen einhergingen und ob diese den Zusammenhang zwischen den Abschlussaspirationen der Schüler:innen

und ihren sozioökonomischen Hintergründen veränderten. Mit einem Differenz-in-Differenzen-Ansatz wurden repräsentative Daten aus Bundesländern mit und ohne *de-tracking* Reformen vor und nach der Implementation der Reformen analysiert (Pant et al., 2015; Stanat et al., 2018; Stanat et al., 2022). Darüber hinaus wurden die Zusammenhänge zwischen den Abschlussaspirationen der Schüler:innen und ihren sozioökonomischen Hintergründen nach den *de-tracking* Reformen mit denen vor der Implementation der Reformen sowie zwischen den Bundesländern mit und ohne *de-tracking* Reformen verglichen.

Zwischen 2012 und 2015, als die Neuntklässler:innen die reformierte Gesamtschule besuchten, ist der Anteil der Schüler:innen, die die allgemeine Hochschulreife anstrebten, in den Bundesländern mit *de-tracking* Reformen gestiegen, während er sich in den Bundesländern ohne Reformen nicht verändert hat. Der Anstieg des Anteils der Schüler:innen, die anstrebten, die allgemeine Hochschulreife zu erwerben, war in der Gruppe der Schüler:innen aus sozioökonomisch begünstigten Hintergründen höher als in der Gruppe der Schüler:innen aus sozioökonomisch benachteiligten Hintergründen. Unter Kontrolle der Leistung blieb der Zusammenhang zwischen den sozioökonomischen Hintergründen der Schüler:innen und ihren Abschlussaspirationen bestehen und wurde durch die *de-tracking* Reformen nicht reduziert. Die Ergebnisse deuten insgesamt darauf hin, dass die *de-tracking* Reformen die sozioökonomischen Disparitäten in den Abschlussaspirationen der Schüler:innen nicht abgeschwächt haben - im Gegenteil, sie schienen sie eher zu verstärken. Dies könnte damit zusammenhängen, Schüler:innen aus sozioökonomisch begünstigten Hintergründen (und ihre Eltern) dem höheren Schulabschluss einen höheren Wert beimessen als Schüler:innen aus sozioökonomisch benachteiligten Hintergründen - trotz des gleichen Leistungsniveaus (vgl. Bittmann & Schindler, 2021; Triventi et al., 2020).

Insgesamt weisen die Ergebnisse darauf hin, dass die *de-tracking* Reformen in der Tat zu Veränderungen geführt haben, z.B. in der sozioökonomischen Komposition in den Gesamtschulen oder beim allgemeinen Anstieg der Abschlussaspirationen der Schüler:innen. Trotz dieser Veränderungen wurde keine Verringerung der

sozioökonomischen Disparitäten beobachtet, weder in den sozioökonomischen Kompositionseffekten noch in den Abschlussaspirationen der Schüler:innen.

Abstract

This dissertation expands the body of research on the relationship between tracking practices and socioeconomic disparities in student achievement and students' educational aspirations. The topic is particularly relevant in Germany due to above-average socioeconomic disparities. Achievement-based differentiation such as allocation to hierarchically structured school types, also referred to as *between-school tracking practices*, determines to a large extent how the student body is distributed across different school types. The school type a student attends in turn has a decisive influence on their educational path and their chances on the labor market.

This dissertation contributes to the discussion of tracking practices and socioeconomic disparities by focusing on whether the recently implemented *de-tracking* reforms in some German states have indeed contributed to reducing socioeconomic disparities. These *de-tracking* reforms reduced the school systems of multiple types of lower secondary schools to a two-path school structure in which there are only academic track and comprehensive schools. After the *de-tracking* reforms, comprehensive schools offer all school-leaving qualifications including the general university entrance qualification that previously could only be obtained at the academic track.

The analyses are based on representative data from the IQB National Assessment Studies 2009 and 2012 before *de-tracking* reforms and the IQB Trends in Student Achievement Studies 2015 and 2018 after the *de-tracking* reforms. The results showed that the socioeconomic composition effects on student achievement at comprehensive schools after the *de-tracking* reforms were at a comparable level to the SEC effects observed at the intermediate track before the reforms. Moreover, the results revealed

that students' educational aspirations increased as a result of the *de-tracking* reforms, but that socioeconomic disparities in students' educational aspirations increased rather than decreased.

Overall, the results indicate that the *de-tracking* reforms have indeed led to changes, e.g., in the socioeconomic composition of comprehensive schools or in the general increase in students' educational aspirations. Despite these changes, no reduction in socioeconomic disparities was observed, neither in terms of socioeconomic composition effects nor in terms of students' educational aspirations.

1

Introduction

Disparities in educational outcomes between students from socioeconomically advantaged and students from socioeconomically disadvantaged backgrounds regularly receive public attention as they are consistently discussed in the context of international and national large-scale assessment studies such as the “Programme for International Student Assessment” (PISA) study or the IQB Trends in Student Achievement Studies (see, e.g., Kuhl et al., 2016; Mahler & Kölm, 2018; Müller & Ehmke, 2016; Niemietz et al., 2023; OECD, 2019). In Germany, educational disparities are considered relevant, partly because socioeconomic disparities are significantly larger than the average of the participating countries in the PISA study (e.g., Müller & Ehmke, 2016; OECD, 2019). In this context, countries are evaluated on the basis of whether they succeed in providing all students with the skills they need to participate fully in society. To achieve this goal, it is considered desirable that student achievement is as little related to their socioeconomic backgrounds as possible (cf. Dumont & Ready, 2023; OECD, 2019; Roemer & Trannoy, 2016).

Tracking practices are being discussed as being associated with socioeconomic disparities (e.g., Baumert et al., 2006b; Bittmann & Schindler, 2021; Buchmann & Dalton, 2002; Buchmann & Park, 2009; Hanushek & Wößmann, 2006; Skopek & Passaretta, 2021; Skopek et al., 2019; Strello et al., 2021; Terrin & Triventi, 2023; Triventi et al., 2020; Woessmann, 2016). For a long time, Germany was dominated by

a hierarchically structured secondary school system, i.e., a system with between-school tracking practices, comprising three school types: vocational, intermediate, and academic track schools (cf. M. Becker et al., 2017). These three school types were designed for students of different ability levels and lead to different school-leaving qualifications. While higher-achieving students usually attend an academic track school, lower-achieving students mostly attend a (less demanding) vocational track school in the traditional three-tier school system. Due to socioeconomic disparities in student achievement and the relationship of transition decisions after elementary school and students' socioeconomic backgrounds, between-school tracking practices are associated with socioeconomic school segregation (Baumert et al., 2006b; Boudon, 1974; Maaz et al., 2008). The resulting socioeconomic composition (SEC) of the student body in the different tracks may contribute to the reinforcement of socioeconomic disparities in students' educational outcomes (see also, among others, Ditton, 2013; Hanushek & Wößmann, 2006; Knigge & Hannover, 2011; Müller-Benedict, 2007; Strello et al., 2021; Terrin & Triventi, 2023; Trautwein et al., 2007; Woessmann, 2016).

With the aim of reducing socioeconomic disparities in educational outcomes, a two-path school structure was established in several German states over the past fifteen years (cf. Baumert et al., 2013, 2017). The states with *de-tracking* reforms reduced their tracking practices with multiple school types to a two-path school structure consisting of academic track and comprehensive schools mostly by converting or merging vocational track, intermediate track, and comprehensive schools. The reformed comprehensive schools offer students the opportunity to reach all school-leaving qualifications.

Existing research exploring the link between tracking practices and socioeconomic disparities predominantly relies on cross-national comparisons. These studies typically either describe the variance in socioeconomic disparities among countries with different tracking practices (cf. Dämmrich & Triventi, 2018; Duru-Bellat & Suchaut, 2005; Schütz et al., 2008; Skopek & Passaretta, 2021; Skopek et al., 2019; Van de Werfhorst & Mijs, 2010) or compare changes in socioeconomic disparities

from the elementary to the secondary school level between countries with different tracking practices (Ammermüller, 2005; Hanushek & Wößmann, 2006; Strello et al., 2021). Additionally, other studies compare countries implementing *de-tracking* reforms with countries without such reforms (e.g., Van de Werfhorst, 2018, 2019). Such cross-national comparisons are accompanied by a certain degree of unobserved differences, whereas comparisons between federal states are assumed to have a smaller degree of unobserved differences (e.g., Woessmann, 2016). This dissertation takes advantage of the introduction of *de-tracking* reforms in only a few federal states and compares states with and without *de-tracking* reforms before and after their introduction (see also Piopiunik, 2021). Its findings can therefore extend the state of research on the relationship between tracking practices and socioeconomic disparities.

Specifically, this dissertation examines, on the one hand, the development of the SEC and its relationship with student achievement, and, on the other hand, the development of students' educational aspirations and their association with students' socioeconomic backgrounds in the course of *de-tracking* reforms. The structure is as follows: Chapter 2 presents the theoretical, empirical, and methodological framework of the dissertation. Chapter 2.1 describes tracking practices and their relationship with socioeconomic disparities. Chapter 2.2 explains the *de-tracking* reforms implemented in some German states and how these reforms might affect socioeconomic disparities in terms of SEC and its relationship to student achievement as well as in terms of students' educational aspirations. Chapter 2.3 addresses the methodological challenges of estimating the effects of the *de-tracking* reforms. Chapter 3 summarizes the objectives of this dissertation. Following the presentation of the studies, Chapter 6 discusses their results (see 6.1), their limitations and methodological evaluation (see 6.2) as well as their implications for research and educational policy (see 6.3).

2

Theoretical, Empirical, and Methodological Background

2.1 Tracking Practices and Socioeconomic Disparities

2.1.1 Introduction to Tracking Practices With a Focus on the German Context

As opportunities for low-skilled workers continue to decline, the demand for education systems to provide students with the essential skills required for success in modern labor markets is increasing (OECD, 2010, 2021). National school systems vary significantly in their ability to meet this challenge (e.g., OECD, 2019; Woessmann, 2016). One contentious matter in this context is the country's tracking practices, which involves grouping students based on their abilities (e.g., M. Becker et al., 2017; Chmielewski, 2014; Terrin & Triventi, 2023). Chmielewski (2014) differentiated among three types of tracking practices: course-by-course tracking, within-school tracking, and between-school tracking. While some countries, such as Finland, do not practice between-school tracking and rely solely on comprehensive compulsory schooling in lower secondary level, others, such as Germany, practice between-school

tracking and separate students into one of several hierarchically ordered school types. Course-by-course tracking streams students within the school with respect to certain subjects and within-school tracking streams students across subjects into different tracks within schools. To better understand German between-school tracking practices, a brief historical overview is first provided before addressing the current national and international debate around tracking practices.

For a long time, the discussion of tracking practices was driven by theoretical considerations and ideologies. In Germany, despite early calls at the beginning of the 19th century by educational reformers such as Wilhelm von Humboldt for basic schooling regardless of family background (Günther, 1988), it took until the end of the 19th century for lower elementary schools for the ordinary (rural) population to be expanded – also in response to the growing demands of industrialization on the education of the working class (cf. Höffken, 2006). The sophisticated schools for the educated elite existed in parallel and stayed untouched. At that time, three hierarchically ordered school types began to coexist in a three-tier school system: the lower elementary schools, the newly introduced “middle schools”, and the academic track schools for the children of the educational elite. Even after World War I, the establishment of a comprehensive school of at least eight years could not be pushed through against conservative forces, but a four-year compulsory comprehensive elementary school was introduced. After World War II, a unified school system was introduced in the Soviet occupation zone, which continued for ten years in the polytechnic secondary school with a common schooling in the General Democratic Republic until reunification (cf. Benner & Sladek, 1994). In West Germany, the three-tier school structure of the Weimar Republic was resumed in 1955 in the framework of the Duesseldorf Agreement, in contrast to East Germany and despite demands, e.g., from the Americans, for a comprehensive school system that transcended social classes (cf. Drewek, 1994). A debate of varying intensity began here, primarily influenced by party politics and ideology (Terhart, 2001). In the 1960s and 1970s, educational scientists, as new actors, engaged in one of the key questions – still without empirical support – about the extent to which school structure could contribute both to increasing student achievement and to decreasing

socioeconomic disparities (cf. Dahrendorf, 1965; Glaser, 2017; Herrlitz et al., 2009; Picht, 1964). In reaction of this debate, comprehensive school experiments were introduced (Fend, 1982; Köller, 2005). The core findings of these comprehensive school experiments – including, a reduction of the association between students’ socioeconomic backgrounds and their educational attainment, but also achievement gaps compared to academic track students – were primarily addressed in party politics (cf. Fend, 1982; Köller, 2005). Reunification brought about various reform efforts in West and East Germany, but the so-called “PISA shock” in 2001 in particular triggered a new high point in the debate on tracking practices.

From now on, the debate was underpinned with empirical results of the international and national large-scale assessment studies such as the PISA study, Trends in International Mathematics and Science Study (TIMSS), and the IQB National Assessment Studies (Breen & Jonsson, 2005). The results of the first PISA and TIMSS studies revealed for Germany below-average achievement with above-average socioeconomic disparities, which triggered another peak in the discussion about between-school tracking practices and school structural reforms (cf. Barz, 2012; Baumert et al., 2001; Baumert et al., 2002; Baumert et al., 1999; Baumert & Lehmann, 1997; Tillmann, 2012, 2016). An intense debate ensued, internationally and also among German scholars, about the relationship between tracking practices and student achievement, as well as socioeconomic disparities, which continues to this day. The next section starts with a brief overview on tracking practices and student achievement.

2.1.2 Between-School Tracking Practices and Student Achievement

There is significant variation among countries regarding the extent and age at which students are tracked (Betts, 2011; Chmielewski, 2014; Skopek et al., 2019). Since the first large-scale assessment studies, a number of studies have used these data to assess the relationship between student achievement and tracking practices, for

example, by comparing countries with different tracking practices (for an overview, see, e.g., Terrin & Triventi, 2023). Findings and arguments in favor of between-school tracking practices include that tracking students into different school types based on their achievement in elementary schools leads to more efficient classrooms, allowing for better adaptation of curricula, and instruction speed in lower secondary school level, which should benefit all students' learning (Duflo et al., 2011; Ireson & Hallam, 2001; Pischke & Manning, 2006; Van de Werfhorst, 2019).

In contrast, there are critical objections that between-school tracking practices only benefit high-achieving students (cf. Epple et al., 2002; Pischke & Manning, 2006) or may not even provide any achievement benefits at all (cf. Matthewes, 2021; Terrin & Triventi, 2023). There are several mechanisms that could cause the effects of between-school tracking practices to be unequal (cf. Terrin & Triventi, 2023). First, if high-achieving peers are beneficial to learning (or low-achieving ones are harmful) (cf. Tenenbaum et al., 2020), between-school tracking practices could have an unequal effect, as it deprives lower-track students of more capable peers (Pischke & Manning, 2006; Sacerdote, 2011). Second, separating students by achievement could have motivational consequences. Lower-track students, who know they are considered less able, may feel discouraged and reduce their learning efforts (Knigge & Hannover, 2011). Third, if (financial) resources are unevenly distributed among tracks, students from certain tracks may be at a disadvantage (Betts, 2011). For example, because of the selection of teachers across school types, teachers with better grades from teacher education are more likely to teach at academic track schools and to provide better teaching quality, as Roloff et al. (2020) showed for Germany. Even if between-school tracking practices in theory are equally beneficial for all achievement levels, practical implementation is likely to be error-prone, as achievement is not directly observable and proxies like teacher assessments and tests are imprecise and may be biased by, for example, the students' socioeconomic backgrounds (see, e.g., Brunello & Checchi, 2007; Stubbe et al., 2016).

This issue of whether between-school tracking practices increase student achievement has also been intensely discussed in Germany in recent years. Esser and Seuring (2020)

examined whether different tracking practices in lower secondary education enhanced average student achievement. Based on data from the National Educational Panel Study in Germany, they used variations in the strictness of between-school tracking practices (e.g., whether teacher recommendations are mandatory) in 13 of the 16 German states. Their results suggested that stricter between-school tracking practices after grade 4 may lead to higher achievement in grade 7 and that this may be due to greater homogeneity of classes in states with stricter between-school tracking practices. Heisig and Matthewes (2022) and Lorenz et al. (2023) disagreed with their findings, criticizing that measures of class composition were based on too small numbers of observed students and that the results could be due to unobserved differences between the states. Extending the analyses by Esser and Seuring, both groups of authors found no evidence of a positive relationship between class homogeneity and student achievement, which Esser and Seuring (2020) assumed to mediate the positive effect of strict tracking practices. Moreover, they argued that the relationship between student achievement and the strictness of between-school tracking practices is largely due to the differences in average level of cognitive abilities between the tracks (Lorenz et al., 2023) or achievement differences prior to the start of the tracking practices (Heisig & Matthewes, 2022).

A recently published meta-analysis by Terrin and Triventi (2023) on tracking practices and student achievement suggests that student achievement does not differ systematically between countries with different tracking practices, although individual studies showed mixed results (lower achievement in systems with an early start of between-school tracking practices, e.g., Hanushek & Wößmann, 2006; Lavrijsen & Nicaise, 2016; higher achievement in systems with an early start of between-school tracking practices, e.g., Piopiunik, 2014; Van de Werfhorst, 2019). Some studies suggest that students who showed lower achievement benefited from learning with higher-achieving classmates, while higher-achieving students experienced either minimal or no negative effects (Guyon et al., 2012; Matthewes, 2021). Other studies on course-by-course tracking have found that students selected for high-achievement classrooms benefited without negative effects for the non-selected students (Card & Giuliano, 2016; Duflo et al., 2011; Figlio & Page,

2002).

Beyond student achievement, students' educational aspirations are a crucial determinant of their educational career, which is significantly influenced by between-school tracking practices and is therefore briefly presented in the following section.

2.1.3 Between-School Tracking Practices and Students' Educational Aspirations

Students' educational aspirations are pivotal as they are considered to influence students' daily behavior in school, thus contributing to shaping their academic careers (e.g., Beal & Crockett, 2010; Bozick et al., 2010; Sewell et al., 1970). They reflect students' assessment of their probable or possible educational achievements and success (Haller, 1968). According to the theories of reasoned action and planned behavior, these aspirations critically forecast choices throughout an educational journey (Ajzen, 1985; Madden et al., 1992), making them a substantial element of educational success. An important distinction is that between realistic and idealistic aspirations (Haller, 1968; Stocké et al., 2011). When students set realistic educational aspirations, they're essentially outlining their accessible educational goals or plans. Their idealistic educational aspirations represent their desired, less normative expectations for their educational careers. Factors like institutional attributes and barriers are commonly included in students' realistic aspirations (Finger, 2016; Kerckhoff, 1977). They thus describe predictions for educational trajectories that take into account factors that facilitate or impede educational success such as tracking practices (Finger, 2016; Stocké et al., 2011), which is why the focus of this dissertation is on students' realistic rather than idealistic aspirations.

In countries with between-school tracking practices such as Germany, students' opportunities to follow certain educational pathways may be limited due to the selection of a school type after elementary school (Buchmann & Dalton, 2002; Parker et al., 2016). Once attending one of the hierarchically ordered school types, the

school type strongly determines the educational aspirations (and the educational pathway) (Buchmann & Dalton, 2002; Buchmann & Park, 2009; Maaz et al., 2008). Several studies have examined the relationship between tracking practices and socioeconomic disparities in students' educational aspirations 2.2.5 (e.g., Bittmann & Schindler, 2021; Buchmann & Dalton, 2002; Buchmann & Park, 2009; Kerckhoff, 1977; Parker et al., 2016), which is explained in more detail in the following section and in section. The next section first discusses the relationship between tracking practices and socioeconomic disparities in terms of student achievement and students' educational aspirations.

2.1.4 Between-School Tracking Practices and Socioeconomic Disparities in Student Achievement and in Students' Educational Aspirations

Beyond the debate about how tracking practices might most efficiently promote learning, tracking practices are further discussed in the context of socioeconomic disparities, most prominently in student achievement, but also in students' educational aspirations. Between-school tracking practices may contribute to magnifying or perpetuating socioeconomic disparities from early education into later life (cf. Brunello & Checchi, 2007; Dämmrich & Triventi, 2018; Schindler et al., 2023; Triventi et al., 2020). The evidence suggests that there are considerable achievement gaps between students in different tracks (e.g., Dustmann, 2004; Müller & Ehmke, 2016; OECD, 2019; Rauch et al., 2016), and countries practicing between-school tracking typically have greater socioeconomic disparities in student achievement and students' educational aspirations (Hanushek & Wößmann, 2006; Parker et al., 2016; Strello et al., 2021; Terrin & Triventi, 2023; Van de Werfhorst & Mijs, 2010).

The correlations between student achievement and family background characteristics are, on the one hand, attributed to direct influences of socioeconomic family backgrounds on learning, so-called primary effects of origin, independent of tracking practices (cf. Boudon, 1974). The higher the socioeconomic and cultural resources of

the family, the more intensive and effective the use of school learning opportunities and, above all, extracurricular stimulation, support and encouragement (cf. Boudon, 1974; Maaz, 2006). As a result, children from socioeconomically disadvantaged families often show lower achievement than children from socioeconomically advantaged families at the end of the common elementary school years (e.g, Haag et al., 2017; Stubbe, Krieg, et al., 2020). In the German system with between-school tracking practices, teachers make a recommendation for the lower secondary school type a student will attend in the future based on the performance in elementary school. Assuming that teachers assess children's potentials in a purely objective and accurate manner, and that parents strictly follow their transition recommendations, the lower secondary school should be selected according to requirements of the school and the student's performance.

However, the selection process is additionally influenced by factors related to students' socioeconomic backgrounds independent of their achievement (Jackson, 2013). Achievement-independent influences of students' socioeconomic backgrounds on educational decisions are referred to as secondary effects of origin (cf. Boudon, 1974). Parents' calculations of the benefits, costs, and likelihood of success of a decision regarding their child's education may vary depending on their socioeconomic status (SES) and education level (cf. Breen & Goldthorpe, 1997). In between-school tracking systems, one result of these different decision-making calculations is that parents with higher SES are more likely to choose more challenging school types than parents with lower SES, regardless of their children's achievement in elementary school (e.g., Ditton & Krüsken, 2006; Dumont, Neumann, Becker, et al., 2013; Stocké, 2013; Stubbe, Kasper, et al., 2020). Accordingly, for example, Bittmann (2022) found that students' (and parents') educational aspirations mediated the association between students' socioeconomic backgrounds and the choice of school type in the transition from elementary to lower secondary school in Germany. The school type, in turn, has a decisive influence on students' educational aspirations (Buchmann & Dalton, 2002; Buchmann & Park, 2009; Parker et al., 2016).

Research on secondary effects of origin in the development of students' educational

aspirations often refers to the Wisconsin Model of Status Attainment and the rational choice theory (see, e.g., Zimmermann, 2020). The Wisconsin Model of Status Attainment assumes that the parents – and also peers as the reference groups – as their immediate environment contribute to shaping students' educational aspirations, in addition to student achievement (Sewell et al., 1970; Sewell et al., 1969; Sewell et al., 2003). It suggests that students adjust their educational aspirations in response to their parents' academic expectations, which are informed by their own education and occupation. Thus, students from socioeconomically disadvantaged backgrounds tend to develop lower educational aspirations due to lower academic expectations set by their parents. Students' educational aspirations are considered to act as mediator between students' socioeconomic backgrounds and their behavior in school, explaining motivational and achievement differences among students from different socioeconomic backgrounds (Stocké et al., 2011). In parallel, rational choice theory assumes that students' educational aspirations arise from a calculated balance of educational benefits, risks or costs, and prospective success of a decision regarding the educational career (Breen & Goldthorpe, 1997). This concept aligns with the expectancy-value model of achievement motivation, positing that a student's motivation for an educational goal rests on the goal's value coupled with anticipated success (see, e.g., Wigfield & Eccles, 2000). The effort to maintain family SES through academic success is seen as a central factor in understanding socioeconomic disparities in students' educational aspirations (Breen & Goldthorpe, 1997; Van de Werfhorst & Hofstede, 2007).

In addition to primary and secondary effects of origin, tertiary effects of origin describe the observations that children from socioeconomically disadvantaged backgrounds tend to be graded lower by teachers for the same achievement and are less likely to receive a recommendation to attend academic track schools (e.g., Esser, 2016; Helbig & Nikolai, 2017; Neuenschwander & Niederbacher, 2021; Stubbe et al., 2016).

Studies analyzing differences in socioeconomic disparities in student achievement between countries with different tracking systems suggest that between-school tracking practices are associated with larger socioeconomic disparities in student

achievement (cf. Dämmrich & Triventi, 2018; Duru-Bellat & Suchaut, 2005; Schütz et al., 2008; Skopek & Passaretta, 2021; Skopek et al., 2019; Van de Werfhorst & Mijs, 2010). Other studies estimate Difference-in-Differences models and compare the changes in socioeconomic disparities in student achievement from elementary to lower secondary school level between countries with different tracking practices (e.g., Hanushek & Wößmann, 2006; Strello et al., 2021). The results tend to show that between-school tracking practices are related to greater socioeconomic disparities (Ammermueller, 2013; Hanushek & Wößmann, 2006; Strello et al., 2021). In addition, Van de Werfhorst (2018, 2019) compared countries with reforms that postponed the onset of between-school tracking practices with countries without an early start of between-school tracking practices. His results suggest lower socioeconomic disparities in countries that introduced these reforms.

Furthermore, there are few studies that examined the changes in student achievement, taking into account their socioeconomic backgrounds, when reforms have been implemented in individual countries to delay the start of between-school tracking practices. For example, Meghir and Palme (2005) observed that student achievement, particularly students from socioeconomically disadvantaged backgrounds, increased following the introduction of a reform in Sweden in the 1950s that expanded comprehensive schooling and introduced a national curriculum. During the 1970s, Finland transitioned from a two-track school to a comprehensive school system up to the age of 15. In the course of these reforms, Kerr et al. (2013) found an increase in student achievement among students from socioeconomically disadvantaged backgrounds, but not among students from socioeconomically advantaged backgrounds. Piopiunik (2014) analyzed changes during a reform in Bavaria in 2000, in which the start of between-school tracking practices in the two lower tracks was moved from grade 6 to grade 4, and argued that the observed decline in student achievement could contribute to an increase in socioeconomic disparities. The study by Piopiunik (2021) used a Difference-in-Differences to examine *de-tracking* reforms introduced in Germany between 2009 and 2012. It categorized eight federal states as reform states that reduced the intensity of between-school tracking practices (either by merging two existing school types, by

offering more school-leaving qualifications in one particular school type or by offering an additional comprehensive school). In this definition of *de-tracking* reforms, the newly established school type differs in whether it offers students the opportunity to obtain the general university entrance qualification. The findings indicated an increase in reading achievement associated with the reforms, but no change in mathematical or science achievement. The increase for reading and mathematical achievement appeared to be larger among students from socioeconomically disadvantaged students. In summary, these studies tend to show that reforms which postponed the start of between-school tracking practices or reduced the intensity of between-school tracking practices might contribute to a reduction in the association of student achievement and their socioeconomic backgrounds.

Regarding socioeconomic disparities in students' educational aspirations, numerous studies demonstrate a positive relationship between students' socioeconomic backgrounds and their educational aspirations (e.g., Bittmann & Schindler, 2021; Bodovski, 2014; Sewell et al., 1969). Students' socioeconomic backgrounds and their educational aspirations are intertwined with student achievement (Parker et al., 2016; Van de Werfhorst & Hofstede, 2007). Socioeconomically advantaged students, on average, show higher achievement, justifying their higher educational aspirations. However, in the case of achievement parity, students from socioeconomically disadvantaged backgrounds are more likely to attend lower track schools in systems with between-school tracking practices and develop lower educational aspirations due to the secondary and tertiary effects of origin (Buchmann & Dalton, 2002; Buchmann & Park, 2009; Lee, 2014; Parker et al., 2016; Van de Werfhorst & Hofstede, 2007).

To summarize, between-school tracking practices tend to perpetuate or even reinforce socioeconomic disparities in student achievement and students' educational aspirations at an early stage, before students complete their education, due to the effects of origin. Related to the effects of origin, the described transition decisions after elementary school also lead to different socioeconomic compositions of the student body at the different lower secondary school types, which is explained in more detail in the next section.

2.1.5 Between-School Tracking Practices and the Socioeconomic Composition of the Student Body

The socioeconomic composition (SEC) of a school often reflects the demographic composition of its neighborhood, since most schools enroll predominantly students from the surrounding area. This means that residential segregation can affect the SEC of a school. Additionally, the SEC can be influenced by horizontal (or within-group) and vertical (or between-group) segregation (Hindriks et al., 2010; OECD, 2019; Willms, 2010). When students with certain socioeconomic backgrounds are more concentrated at individual schools than at other schools of the same school type, i.e., when segregation occurs between individual schools of the same school type, it is referred to as horizontal segregation. In systems with between-school tracking practices, vertical segregation describes the segregation between different hierarchically ordered school types. Students are segregated between different school types due to the achievement-based tracking practices or admission policies. In this way, differences in the achievement composition in various schools or school types are intentionally created. Because of the relationship between student achievement and their socioeconomic backgrounds and school selection processes (cf. 2.1.4; Boudon, 1974), the SEC may vary accordingly. The primary effects are considered to give rise to vertical socioeconomic segregation between the hierarchically differentiated school types. Secondary and tertiary effects of origin are then discussed to have a reinforcing effect on vertical socioeconomic segregation between the different school types, independently of student achievement (cf. Maaz et al., 2008). Thus, vertical socioeconomic segregation is an undesirable but hardly avoidable side effect of between-school tracking practices.

These considerations are supported by numerous empirical analyses. Using data from large-scale assessment studies from 2000 to 2018, for example, several studies indicate that socioeconomic segregation is substantially more pronounced in countries with between-school tracking practices than in countries without between-school tracking practices (Alegre & Ferrer, 2010; Jenkins et al., 2008; Strello et al., 2021).

Jenkins et al. (2008), for example, found that in the countries with between-school tracking practices like Germany, Austria, and Hungary, more than half of overall segregation was due to vertical segregation between different school types. Strello et al. (2021) used a Difference-in-Differences approach with 16 assessment cycles and in 75 different countries or regions and found evidence for the hypothesis that between-school tracking practices increase socioeconomic segregation across schools (not differentiating between horizontal and vertical segregation). Results on socioeconomic segregation in Germany and possible changes since the implementation of *de-tracking* reforms in some German states will be presented in section 2.2.3. The next section addresses the relationship of SEC and student achievement.

2.1.6 The Relationship Between Socioeconomic Composition and Student Achievement

The Coleman report represents a landmark study showing, among other things, that SEC is a school characteristic significantly associated with student achievement (Coleman et al., 1966). That means, student achievement is not solely determined by individual characteristics, but is significantly influenced by the SEC. The relationship between SEC and students' educational outcomes, referred to as socioeconomic composition (SEC) effects, generally characterizes the association of the aggregated group-level characteristic, e.g., average SES, controlling for the individual student characteristic, e.g., student SES, in multilevel analyses (Raudenbush & Bryk, 2002). Composition effects aim to explain how differences in achievement between similar students can result from learning environments with different compositions. In systems with between-school tracking practices, composition effects are observed in addition to or in interaction with characteristics of the school type, which are referred to as institutional effects (Baumert & Schümer, 2001; Köller & Baumert, 2001; Scharenberg et al., 2014).

Research concerning composition effects in general and SEC effects in particular has faced critique on both theoretical and methodological aspects (Armor et al., 2018;

M. Becker et al., 2022; Zhou & Ma, 2023). Criticisms include, for example, the operationalization of SEC measures (e.g., when using free/reduced-price lunch/schooling as an indicator; Hill & Jenkins, 2001; Marks, 2015; Sirin, 2005; Van Ewijk & Sleegers, 2010), and issues with model misspecification resulting from omitted variables (Lauen & Gaddis, 2013; Marks, 2015; Mickelson et al., 2013; Snijders & Bosker, 2012). Potential unreliability or high measurement error of individual characteristics can create statistically significant composition effects, also termed as the “phantom effect” or “contextual fallacy” (Harker & Tymms, 2004; Hauser, 1970; Lüdtke et al., 2011; Lüdtke et al., 2008; Lüdtke et al., 2002). Aggregation bias can occur when these individual attributes, prone to measurement errors, are aggregated to form a (class or school) characteristic (e.g., Baumert et al., 2006b; Harker & Tymms, 2004; Lüdtke et al., 2011; Marsh et al., 2012). A small sample size of student information within a class or school represents another possible source of error (sampling error; Lüdtke et al., 2011; Lüdtke et al., 2008). Especially in cross-sectional studies, given that students are not just influenced by their class environment, but they also actively shape it, a simultaneity or reciprocity issue may arise (cf. Baumert et al., 2006b; G. J. Duncan et al., 2004; G. J. Duncan & Raudenbush, 1999).

Given these methodological challenges, several studies on the association of SEC and student achievement found that students in classes or schools with lower average SES showed lower achievement compared to students with the same background characteristics in a class or school with higher average SES (e.g., Opdenakker & Van Damme, 2001; Palardy, 2013; Sciffer et al., 2022; Tan et al., 2023; Van Ewijk & Sleegers, 2010). A meta-analysis conducted by Van Ewijk and Sleegers (2010), investigated the impact of peer SES at varying levels, including the classroom and the school level, and found that the sizes of SEC effects depend on its measure (e.g., free lunch eligibility or a composite of different SES measures). A recently published meta-analysis by Tan et al. (2023) synthesizes findings of 97 studies on SEC effects. Most of these studies employ the proportion of free/reduced-price lunch or schooling as a singular indicator to gauge SEC, focusing on SEC effects on student achievement (primarily in languages and mathematics), rather than student attainment (such as

reaching key educational milestones like college admissions or graduation), and are predominantly US-based. Their findings demonstrate an average association of SEC and student educational outcomes of $r = .58$, which were observed across both private and public elementary and secondary schools.

In systems with between-school tracking practices, selection processes in the transition from elementary to lower secondary school lead to different SEC at the beginning of tracking. Based on theoretical assumptions, researchers have aimed to understand how and why SEC is related to student achievement, distinguishing between interrelated paths (see, e.g., Baumert et al., 2006b; Harker & Tymms, 2004; Rumberger & Palardy, 2005; Tan et al., 2023). How many paths are differentiated varies (three paths: Baumert et al., 2006b; Harker & Tymms, 2004; Thrupp et al., 2002; Van Ewijk & Slegers, 2010; six paths: Tan et al., 2023). In the case of three paths, instructional quality or teaching, school resources, as well as interactions and mutual influences between classmates are discussed (Harker & Tymms, 2004; Thrupp et al., 2002; Van Ewijk & Slegers, 2010). These assumed mediation processes are described in section 2.2.3, where they are presented with reference to the *de-tracking* reforms. Before that, the next section starts with presenting the between-school tracking practices in Germany and the *de-tracking* reforms of some states.

2.2 Socioeconomic Disparities and *De-Tracking* Reforms in Some German States

2.2.1 Overview on Between-School Tracking Practices in Germany in Recent Years

In Germany, sovereignty in education policy lies with the state governments due to federalism. However, in order to ensure comparable educational standards and qualifications, the Standing Conference of the Ministers of Education and Cultural Affairs of the States, known as the Kultusministerkonferenz (KMK), harmonizes

educational policies between the states to a certain extent, e.g., with regard to basic structure, teaching methods, and curricula (KMK, 2020). Nonetheless, some differences exist within certain areas of schooling policy, particularly in tracking practices. Compared to cross-national approaches, however, less unobserved differences between federal states can be assumed (Waldinger, 2007; Woessmann, 2016), so that changes in tracking practices in certain German states present an opportunity to investigate the relationship between tracking practices and socioeconomic disparities (cf. 2.3).

In Germany, mandatory education begins at the age of six through comprehensive elementary school, which spans the first four grades (exception: Berlin and Brandenburg the first sixth grades). Variations between states arise after the completion of elementary school. In the traditional (West) German secondary school system, according to its between-school tracking practices, students are placed into one of three vertically structured school types, after completing elementary school: vocational, intermediate, and academic tracks (cf. 2.1.1). These tracks differ in terms of years of schooling, curriculum, and, for the most part, teacher education. Each of them offers a single track of education and leads to a particular school-leaving qualification, which in turn largely determines students' educational and labor-market options (Maaz et al., 2008).

The academic track (i.e., *Gymnasium*) covers eight or nine years and offers the most demanding curriculum. After completion, students obtain the general university entrance qualification (the highest school-leaving qualification). The intermediate track (i.e., *Realschule*), which lasts six years, offers general education to prepare students for advanced vocational and professional training. In contrast, the vocational track (i.e., *Hauptschule*) provides a more fundamental general education over five years and prepares students for technical vocational training. After completing the intermediate or vocational track, students may have the option to advance to academic track schools under specific circumstances. These three school types are supplemented in several states with one or even two additional school types. One of them are the comprehensive schools which cover several courses of education in one school. In

comprehensive schools, students are usually streamed within the school with respect to certain subjects (course-by-course tracking). The comprehensive schools enable students to achieve all school-leaving qualifications. In some states, within-school tracking schools exist, which stream students within the school into different tracks for all subjects. In the eastern German states, students mainly attend either the academic track or within-school tracking schools and, to a smaller extent, comprehensive schools (see M. Becker et al., 2017).

Against a backdrop of demographic change, declining acceptance of the vocational track schools, low student achievement and pronounced socioeconomic disparities identified by large-scale international and national assessment studies, some states in particular faced substantial pressure to implement reforms.

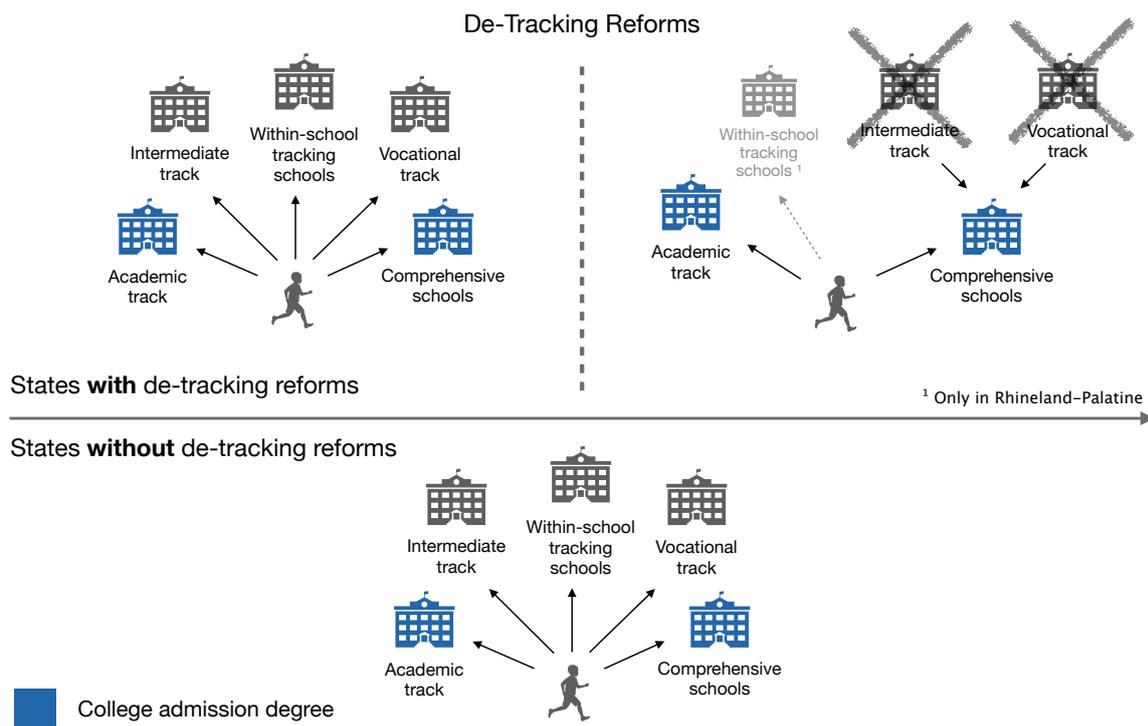
2.2.2 The *De-Tracking* Reforms of Some German States

The discussion about the “best” tracking practices has been going on in Germany for decades fueled by the results of international large-scale assessment studies, which have described above-average socioeconomic disparities in student achievement in Germany (cf. 2.1.1, see, e.g., Baumert et al., 2001; Baumert & Lehmann, 1997; Müller & Ehmke, 2016; OECD, 2019). Measures of socioeconomic disparities seem to be particularly high in school systems with between-school tracking practices (Ammermüller, 2005; Baumert et al., 2006b; Hanushek & Wößmann, 2006; Strello et al., 2021; Strello et al., 2022; Terrin & Triventi, 2023; Van de Werfhorst, 2018, 2019).

Further, vocational track schools have been widely criticized for depriving students of essential educational opportunities and reinforcing educational disparities (Baumert et al., 2006b; Hurrelmann, 2013; Ridderbusch, 2019; Tillmann, 2012; Tillmann & Wischer, 2006). Potential obstructing factors for efficient learning may be concentrated in vocational track schools, which is generally accompanied by a declining acceptance and stigmatization of this school type (Baumert et al., 2006b; Helbig & Nikolai, 2015; Knigge & Hannover, 2011; Weishaupt, 2009). Over the last fifteen years, several

states have implemented reforms and transitioned towards a two-path system in order to avoid the concentration of risk factors in schools, reduce disparities in achievement acquisition, and ultimately to raise achievement levels at secondary level and to improve participation in education (Baumert et al., 2013, 2017; Hurrelmann, 2013; Ridderbusch, 2019; Tillmann, 2012). By converting or merging vocational track, intermediate track, and comprehensive schools, the resulting comprehensive school is to take on a central role and - in the medium term - remain the only school type alongside the academic track (cf. Figure 2.1). Comprehensive schools integrate all three traditional educational tracks and thus also open up a second direct path to the general university entrance qualification in particular (i.e., a “two-path system” because both academic track and comprehensive schools offer students the opportunity to obtain all school-leaving qualifications) (cf. Hurrelmann, 2013).

Figure 2.1: School Structure in Germany Before and After De-Tracking Reforms



The states with *de-tracking* reforms include Berlin, Bremen, Hamburg, Rhineland-Palatinate, and Schleswig-Holstein. Among the states with reforms, Berlin established a two-path system in the 2010/11 school year by converting all

non-academic secondary schools into comprehensive schools, the “integrated secondary schools” (cf. Baumert et al., 2017). As early as in the 1970s, Bremen combined, but sustained the vocational and intermediate track schools in many cases organizationally at “school centers”. Since the 2009/10 school year, a new form of comprehensive school was introduced, which unites the former school centers and comprehensive schools and gives rise to a two-path system (cf. Maaz et al., 2019). In Hamburg, as of the 2010/11 school year, vocational and intermediate track schools were abolished and a new form of comprehensive school, the “Stadtteilschule”, was established, resulting in a two-path system (cf. Lenz et al., 2019). Since the 2009/10 school year, Rhineland-Palatinate has replaced vocational track and intermediate track with the “Realschule plus”, while at the same time significantly expanding the range of comprehensive schools, thus creating an expanded two-path system (see Landtag Rheinland-Pfalz, 2008; for a more detailed description of the *de-tracking* reforms in the individual five states, see Lenz, 2022). Schleswig-Holstein introduced two new school types in the 2008/09 school year with the “Regionalschule” as a within-school tracking school type and the “Gemeinschaftsschule” as a comprehensive school type. Vocational track and intermediate track schools were successively merged. As part of a further reform measure, the “Regionalschulen” were converted into “Gemeinschaftsschulen” since the 2014/15 school year, completing the transformation into a two-path system (see Jungmann, 2008; Wiechmann, 2013).

The decoupling of school type and school-leaving qualification, and the offering of all educational paths in the available school types, aim to contribute to a more balanced composition of the student body. Prior to the reforms, there were vocational track schools that were characterized by a student body composition identified as unfavorable for learning (Baumert et al., 2006b; Maaz et al., 2008; Trautwein et al., 2007). Many low-achieving students from socioeconomically disadvantaged backgrounds attended the vocational track schools, which largely determined their school-leaving qualification that lead them finish school after nine years (Buchmann & Park, 2009; Maaz et al., 2008). The *de-tracking* reforms are linked to the goal of reducing socioeconomic disparities in learning and educational participation, and ultimately raising achievement levels at the secondary level including through a

performance-based - and consequently also socioeconomic (and cultural) - mix of the student body (cf. Hurrelmann, 2013; Neumann, Maaz, & Becker, 2013; Ridderbusch, 2019; Tillmann, 2012). This dissertation aims to evaluate whether these *de-tracking* reforms contributed to reducing socioeconomic disparities by focusing on the SEC and its link to student achievement, as well as on students' educational aspirations, which have only been studied in individual states to date (cf. Helbig & Nikolai, 2017; Lenz et al., 2019; Lenz et al., 2022; Maaz et al., 2019; Neumann, Lühe, et al., 2017). In addition, the dissertation provides findings that can enrich the scholarly discussions of the interplay between tracking practices and these aspects of socioeconomic disparities (cf. 2.1.4, 2.1.5, and 2.1.6). The next three sections present how the *de-tracking* reforms might affect SEC and its association with student achievement as well as students' educational aspirations.

2.2.3 Socioeconomic Composition in the Course of the *De-Tracking* Reforms

Before considering how socioeconomic disparities in terms of SEC effects on student achievement may have changed in the course of the *de-tracking* reforms, this section describes the SEC in the states with and without *de-tracking* reforms before and after the introduction of the *de-tracking* reforms. Since studies on SEC were only available in individual states (cf. Helbig & Nikolai, 2017; Maaz et al., 2019; Neumann, Becker, et al., 2017), the studies of this dissertation were preceded by a study on this very question:

Lenz, S., Holtmann, M., Rjosk, C., & Stanat, P. (2019). Soziokulturelle Segregation an weiterführenden Schulen – Analysen zur Rolle der Gliederung des deutschen Schulsystems und schulstruktureller Reformmaßnahmen [Sociocultural segregation in Germany's secondary school systems – Effects of between-school track]. *Zeitschrift für Erziehungswissenschaft*, 22, 1333–1358.

This study used data from the IQB National Assessment Study 2009 before the introduction of the *de-tracking* reforms and the IQB Trends in Student Achievement Study 2015 after the introduction of the *de-tracking* reforms, thus covering the reform period. The representative samples include a total of 36,605 students at 1466 schools (2009) and 31,350 students at 1425 schools (2015). The socioeconomic segregation of the states with *de-tracking* reforms were compared with two further groups of federal states without *de-tracking* reforms: one is characterized by a three-tier school structure with academic, intermediate, and vocational track schools or a three-tier structure plus comprehensive schools (Bavaria, Baden-Württemberg, Hesse, Lower Saxony, and North Rhine-Westphalia); and the other by a two-tier structure, i.e., with comprehensive schools or within-school tracking schools in addition to the academic track schools (Brandenburg, Mecklenburg-Vorpommern, Saarland, Saxony, Saxony-Anhalt, Thuringia). The highest value of the socioeconomic index ISEI (Ganzeboom et al., 1992) of the parents (HISEI) served as an indicator for student SES. First, both measurement points were compared in terms of the proportion of students who belong to the bottom SES quarter at the various school types using t-tests (Welch tests). Second, dissimilarity indices for the three groups of states for 2009 and 2015 were computed (cf. O. Duncan & Duncan, 1955). This segregation measure can be used to calculate the proportion of a given group of students who would have to change the school type in order for the group to be distributed evenly across the different school types.

The results showed that the proportion of students from socioeconomically disadvantaged backgrounds increased at the comprehensive schools in the states with *de-tracking* reforms, probably because students who attended intermediate or vocational track schools before the reforms now mainly attend comprehensive schools. Vertical socioeconomic segregation was evident to a similar extent in all three groups of states and remained largely stable between 2009 and 2015, also in the states with *de-tracking* reforms. The findings are consistent with the persistence of socioeconomic segregation identified in various analyses focusing on individual states (e.g., Helbig & Nikolai, 2017; Lenz et al., 2022; Maaz et al., 2019; Neumann, Becker, et al., 2017). The vertical socioeconomic segregation in all groups of states was

almost entirely due to differences between academic and non-academic school types and therefore did not appear to be influenced by the changes of the non-academic school types in the states with *de-tracking* reforms (for a further discussion of the results, see Lenz et al., 2019). To summarize, the study indicated that vertical socioeconomic segregation did not decrease in the course of the *de-tracking* reforms.

Subsequently, another study by Lenz et al. (2022) examined the segregation in three German states (Berlin, Bremen, and Hamburg) before and after the introduction of the *de-tracking* reforms. In consistency with the study presented above, their results indicated that the vertical socioeconomic segregation has not changed in Berlin and Hamburg. In Bremen, a decrease in vertical socioeconomic segregation could be observed, which might be related to the special circumstances of limiting the number of academic track schools (Lenz et al., 2022). Overall, the observations in the course of the *de-tracking* reforms indicated that vertical socioeconomic segregation has hardly changed (Helbig & Nikolai, 2017; Lenz et al., 2019; Lenz et al., 2022; Maaz et al., 2019; Neumann, Becker, et al., 2017).

The studies described focused on changes in socioeconomic segregation between academic and non-academic schools in the states with *de-tracking* reforms. Horizontal and vertical socioeconomic segregation affect the SEC of school types and schools within a school type. The studies did not examine the SEC in the different non-academic school types. Because of the changes in the non-academic school types as a result of the *de-tracking* reforms, it is particularly important to examine the so far unanswered question of whether and how their SEC has changed. This dissertation addresses this desideratum. The following section presents theoretical considerations on how the *de-tracking* reforms might affect SEC effects.

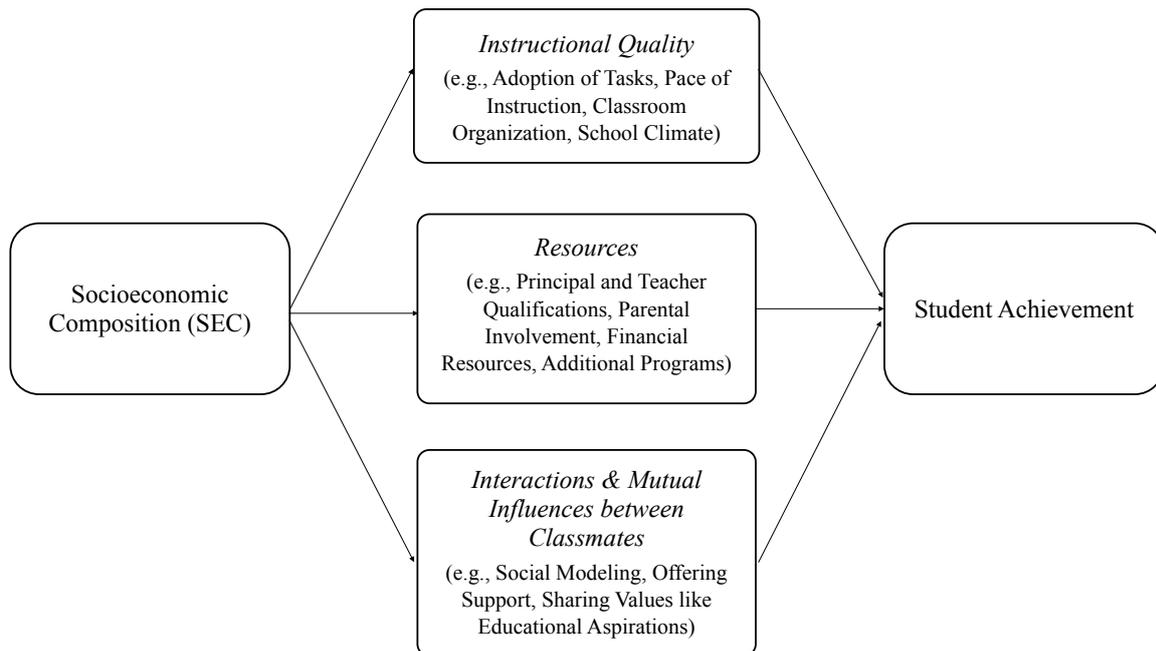
2.2.4 Socioeconomic Composition Effects on Student Achievement in the Course of the *De-Tracking* Reforms

In terms of how SEC might be related to student achievement when *de-tracking* reforms are introduced, processes in the three interrelated paths mentioned in section

2.1.6 could be affected (cf. Figure 2.2).¹ In general, the mediating processes of all three paths usually refer to socioeconomic and achievement composition because these characteristics are intertwined (see 2.1.4).

Instructional quality as one of the paths can include the adaptation of cognitively demanding learning tasks, pace of instruction, effective classroom organization, well thought-out instructional programs, and a student-oriented, positive school climate with higher levels of expectations depending on the class composition (Baumert et al., 2006b; Kunter et al., 2013; Mitchell et al., 2016; Pianta & Hamre, 2009; Reynolds et al., 2014), which have been empirically identified as being linked to student achievement, discipline problems, and motivational outcomes (e.g., Baumert et al., 2010; Holzberger et al., 2020; Kunter et al., 2013; Palardy, 2015; Rjosk et al., 2014). In systems with between-school tracking practices, lower-achieving students typically find themselves in lower tracks with less demanding curricula (Baumert et al., 2006b; Hattie, 2002). Moreover, teaching quality might be higher in academic track schools than in the other tracks due to differing teaching qualifications (Rjosk et al., 2014; Roloff et al., 2020). In the course of *de-tracking* reforms, instruction at comprehensive schools (compared to instruction at vocational or intermediate schools) could be made more challenging in line with the changing composition of the student body. Curricula could also be adapted to teach the competencies that students need to have in order to aim for the general university entrance qualification. These potential changes might, for example, contribute to weaken the associations of SEC and student achievement.

¹Here, it is referred to the distinction of three paths and the aspects of the other distinction are subsumed under it.

Figure 2.2: SEC Effects on Student Achievement (Mediation Model)

As another path, resources available to the class or school such as principal or teacher qualifications, parental involvement, and programs that promote learning are also discussed to be linked to the SEC and its association with student achievement (Akay & Karadag, 2019; Hattie, 2002; Muijs et al., 2014; Rumberger & Palardy, 2005; Tan et al., 2023). For example, the most qualified teachers (who generally have priority in transfer opportunities when positions become available) tend to go to schools with a more favorable composition, where they expect teaching to be easier due to fewer discipline problems and higher student motivation (Akay & Karadag, 2019; Betts et al., 2000; Harker & Tymms, 2004; Muijs et al., 2014). In addition, certain educational initiatives, like those allocating financial resources to schools based on their percentage of disadvantaged students, could be used to provide additional support for student achievement (cf. OECD, 2017). Furthermore, parental engagement is attributed to this path (Epstein, 2010; Tan, 2018, 2020). If, following the *de-tracking* reforms, comprehensive schools were also accepted as an alternative educational path to the academic track by students from socioeconomically advantaged backgrounds, it is conceivable, for example, that their parents, who are likely to value education,

would show involvement. In case of more favorable compositions at comprehensive schools after the introduction of the *de-tracking* reforms, some of the motivated and well-qualified teachers may choose to teach at comprehensive schools, as they could perceive the working conditions here as better than previously in the lower track school types. Yet, these processes might take time to unfold and show effects.

Interactions and mutual influences between classmates are categorized under a third path that might be affected more immediately by the *de-tracking* reforms. They capture a range of dynamics linked to peer interactions, which might affect student achievement and their motivational outcomes.² These processes include, for example, the passing on or demonstration of problem-solving techniques, the clarification of questions or the offer of support by higher-achieving students (more likely from socioeconomically advantaged backgrounds) (Bandura, 1971; Ben-Ari & Kedem-Friedrich, 2000; Kuzmina & Ivanova, 2018), which may mediate the association of SEC and student achievement or college choice (e.g., Hornstra et al., 2015; H. Liu et al., 2015; Palardy, 2015). Through mechanisms like social modeling or positive reinforcement, in classrooms with higher shares of students from socioeconomically advantaged backgrounds, shared beliefs, values, and behaviors towards learning and schooling, such as a consistent school attendance or positive school attitudes, are likely to become prevalent (see, e.g., Bandura, 1971; Dishion & Tipsord, 2011; Eccles & Roeser, 2011; Goldsmith, 2011). Following the introduction of the *de-tracking* reforms, by offering a pathway to the general university entrance qualification and potentially mixing higher- and lower-achieving students from different socioeconomic backgrounds at comprehensive schools, students may encourage each other and be more motivated to aspire higher school-leaving qualifications. According to the concept of social or peer contagion, peers influence each other's motivations and learning behavior, gradually aligning over time (Erbring & Young, 1979; Kelly, 2009). Higher-achieving students (more likely from socioeconomically advantaged backgrounds) can have a positive influence by modeling beneficial values and aspirations and sharing insights on the advantages of

²The peer processes addressed here do not quite correspond to the understanding of the term in economics, where it encompasses a broader range of characteristics (see Sacerdote, 2011)

education (Dupriez et al., 2012; Hornstra et al., 2015). These peer interactions could, in turn, facilitate student learning (e.g., Tenenbaum et al., 2020).

This section has reviewed theoretical assumptions and empirical findings about how SEC might affect student achievement and has provided some theoretical considerations about how SEC effects might be related to the introduction of *de-tracking* reforms. The empirical investigation of these theoretical considerations represents a desideratum of research.

In systems with between-school tracking practices, the students' educational aspirations are significantly shaped by the school type they attend. The next section focuses on how *de-tracking* reforms might affect students' educational aspirations and their socioeconomic disparities at the individual level.

2.2.5 Social Disparities in Students' Educational Aspirations in the Course of the *De-Tracking* Reforms

Research on differences in students' educational aspirations related to tracking practices began early (Kerckhoff, 1977; Spenner & Featherman, 1978). Beyond that, studies have explored the relationship between students' educational aspirations, their socioeconomic backgrounds, and between-school tracking practices (e.g., Bittmann & Schindler, 2021; Buchmann & Park, 2009; Lee, 2014; McDaniel, 2010). Buchmann and Dalton (2002) observed that in between-school tracking systems, the school type a student attends strongly determined students' educational aspirations, possibly because it might be perceived as an indicator of a student's abilities. Further, Buchmann and Park (2009) revealed that in between-school tracking systems, students' socioeconomic backgrounds played a pivotal role in deciding the school type a student ends up attending. Due to socioeconomic disparities in the transition from elementary to secondary level in between-school tracking systems (cf. 2.1.4), it is more likely that students from socioeconomically advantaged backgrounds attend the higher tracks than students from socioeconomically disadvantaged backgrounds with the same achievement (Buchmann & Dalton, 2002; Buchmann & Park, 2009; Lee,

2014; Parker et al., 2016; Van de Werfhorst & Hofstede, 2007). Students from socioeconomically disadvantaged backgrounds tended to be less likely to attend schools that allow a direct transition to university, when showing the same achievement (cf. 2.1.4; see also Triventi et al., 2020). The school type highlights the most probable educational trajectory, thereby largely shaping students' educational aspirations (Buchmann & Dalton, 2002). These observations aligned with the research of Lee (2014), which indicated that when accounting for the association of socioeconomic background and school type, the direct link between students' socioeconomic backgrounds and their educational aspirations became less pronounced. Lee (2014) posited that delaying the initiation of between-school tracking practices could attenuate the association between students' socioeconomic backgrounds and their educational aspirations.

In comprehensive schools, the school type does not determine students' educational aspirations (Buchmann & Dalton, 2002). There, students might adapt their educational aspirations to their performance over the course of their school careers (cf. Baumert et al., 2017; Buchmann & Park, 2009; Knigge, 2009; Maaz et al., 2008; Neumann, Lühe, et al., 2017). The opportunities at comprehensive schools and peer interactions might increase the educational aspirations of students from socioeconomically disadvantaged backgrounds beyond parental influence for those who perform well in school (cf. 2.2.4; Alexander & McDill, 1976; Buchmann & Dalton, 2002; Lorenz et al., 2020). A study by Matthewes (2021) showed that in German states with non-academic comprehensive schools in grade 5 and 6, the proportion of students who do not plan to obtain a general higher education entrance qualification is lower than in German states with between-school tracking practices. The *Berlin study*, which scientifically accompanied the *de-tracking* reforms in Berlin, observed a rise in educational aspirations among non-academic track students planning to obtain the general university qualifications after the implementation of the *de-tracking* reforms (Neumann, Lühe, et al., 2017). The correlation between students' socioeconomic backgrounds and their educational aspirations seemed to remain significant in Berlin (Maaz et al., 2017). However, student achievement did not appear to be controlled for in the analyses of students' educational aspirations in

the *Berlin study*. The statistical control of achievement is important in order to be able to assess whether students' educational aspirations are connected with their achievement. Bittmann and Schindler (2021) focused on high-achieving students in Germany, who indicated in grade 5 that they wish to obtain the general university entrance qualification, and observed the development of their aspirations up to grade 9. Their results suggested that students in comprehensive schools were more likely to lower their educational aspirations than students in academic track schools. The share of students who changed their educational aspirations towards a less ambitious option than the general university entrance qualification was higher in comprehensive schools than in academic tracks schools (see also Schindler et al., 2023). Moreover, the authors observed that students from socioeconomically disadvantaged backgrounds lowered their educational aspirations more frequently than students from socioeconomically advantaged backgrounds. They argued that this observations might be related to the fact that students from socioeconomically disadvantaged backgrounds more frequently attend non-academic tracks.

In summary, the findings described above suggest that changes in tracking practices might be associated with changes in students' educational aspirations as well as with socioeconomic disparities in students' educational aspirations. Since there are hardly any studies comparing the socioeconomic disparities in students' educational aspirations between different non-academic tracking practices, this dissertation addresses the question of whether the transformation of the non-academic multiple-track school structure into comprehensive schools (or the existence of comprehensive schools instead of a tracked non-academic school structure, respectively) might affect socioeconomic disparities in students' educational aspiration, when their achievement is taken into account. The *de-tracking* reforms in some German states provide an opportunity to extend this body of research by conducting a comprehensive comparison of the educational aspirations of students from all socioeconomic backgrounds, taking into account their achievement, between states with different non-academic tracking practices.

Against the theoretical background and the empirical findings presented, the next

section addresses the central debate about how reform effects can be estimated in the first place, before the objectives of the dissertation are summarized (Chapter 3) and the studies themselves are presented (Chapter 4 and Chapter 5).

2.3 Estimating the Effects of the *De-Tracking* Reforms

If policies are meant to improve on outcomes, they should not be based on ideology or wishful thinking but rather on proven effectiveness. But obtaining convincing evidence on the effects on specific education policies and practices is not an easy task. (Schlotter et al., 2011, p.28)

Generating empirical evidence that can effectively guide policy is a challenging endeavor. This difficulty arises from the need to establish causal relationships, which requires the use of specialized research methods and adequate data. The primary objective of these methods is to differentiate between accidental associations and genuine causation.

When factors other than the predictor itself contribute to the correlation between the predictor and the outcome, interpreting this correlation as the causal effect of the predictor is problematic; this is referred to as an endogeneity problem (see, e.g., Wooldridge, 2015). The term “endogeneity” originates from the notion that the predictor cannot be viewed as independent from the error term of the model. Instead, it is endogenously determined within the model, depending on or jointly determined with the outcome by a third factor. If this third factor affects both predictor and outcome and is unknown or unobservable, it is referred to as “omitted variable(s)”. The presence of a third factor introduces another potential explanation for the observed associations (see, e.g., Wooldridge, 2015). For example, in the case of an intervention or policy, if unobserved heterogeneity between the treated and untreated is associated with a variable that affects both the treatment of interest and the outcome, it is difficult to identify and isolate the true causal effects. A specific example of omitted variables related to this dissertation would be the

non-consideration of students' socioeconomic backgrounds when examining the relationship between the choice of school type in the transition from elementary to lower secondary school and student achievement (cf. 2.1.4). Another source of endogeneity is the concept of "reverse causality", where the outcome of interest may actually influence the treatment being studied, making it difficult to estimate causal effects (see, e.g., Gerstman, 2013). For example, reverse causality could be present if the effects of *de-tracking* reforms on student achievement are estimated, while student achievement might be a driving factor for the introduction of the *de-tracking* reforms (cf. 6.2.2).

Standard approaches attempt to address this issue by accounting for other potential sources of correlation and removing the impact of these observed differences on the outcomes. Multivariate models, such as classical ordinary least-squares (OLS) or multilevel modeling (also known as hierarchical linear models or HLM), estimate the effects of multiple variables on the outcome simultaneously. They estimate the association between predictor and outcome while considering the effects of other observed factors. However, often we cannot fully observe all relevant variables, and as long as some of them remain unobserved and omitted, the estimated conditional association cannot be interpreted causally (cf. Schlotter et al., 2011).

When it comes to estimating the effects of educational policy reforms, countries with different tracking practices are usually compared in order to identify the advantages and disadvantages of these practices. Apart from their tracking practices, these countries differ in a number of other characteristics such as their demographic structure, their teacher education or curricula, which makes reliable comparisons difficult. Ideally, we would like to observe the same individuals at the same time both in the treated and untreated countries. In such a scenario, one could argue that the difference in outcomes is genuinely attributable to the effects of the intervention. However, this is impossible because an individual cannot be simultaneously both treated and untreated. Therefore, the key challenge lies in estimating what would have happened in the counterfactual situation, specifically, the outcome that a treated individual would have had if they had not received the treatment (cf. Angrist &

Pischke, 2009). It is necessary to develop research designs that come as close as possible to constructing such a counterfactual comparison. The basic idea of those approaches that aim to identify causal effects is that treatment and control groups do not show systematic differences, provided that the division of the population into the two groups is completely random and a substantial number of individuals are observed that guarantees randomness (Schlotter et al., 2011).

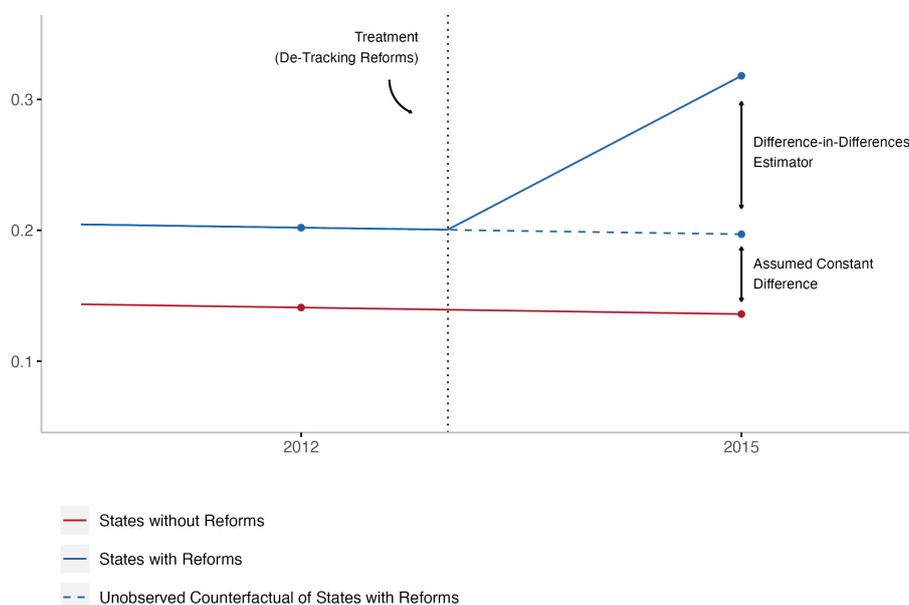
Due to the rarity of randomized experiments, methodological techniques are employed to approximate them as closely as possible. Examples of such techniques include the Instrumental-Variable Approach and the Regression-Discontinuity Approach (cf. Angrist & Pischke, 2009). These approaches seek to leverage variations in observational data that arise from sources external to the association of interest. In a way, they attempt to emulate the random assignment by exploiting instances where natural occurrences or institutional regulations and designs generate random variations (Schlotter et al., 2011).

Alternative approaches aim to mitigate the influence of unobserved factors, even when they are not directly observed in the data. This can be achieved if these unobserved factors remain “fixed” along the second dimension (such as time) of the data set, for example, if these unobserved factors remain constant over time (cf. Schlotter et al., 2011). That is why approaches like fixed effects estimations often rely on (extensive) panel data as a starting point. By disregarding variations in the outcome levels across individuals and focusing solely on changes over time – as in the Difference-in-Differences approach, which is related to fixed effects estimations – these approaches address unobserved heterogeneity resulting from factors that remain constant over time.

The Difference-in-Differences approach is a method for panel data analysis employed when certain groups are exposed to a treatment while others are not (cf. Angrist & Pischke, 2009; Athey & Imbens, 2017). The identification strategy involves computing two differences between group means: The first difference between the groups at each time point is calculated; the second difference is then determined from the difference between the differences calculated in the first step. The second difference

captures how the change in the outcome varies between the two groups, thus aiming to represent the causal effect of the treatment variable.

The underlying idea of this strategy is the following: The two groups may exhibit observable differences, meaning their group-specific means might diverge even in the absence of treatment. However, if this difference remains constant over time in the absence of treatment, it can be eliminated by subtracting the difference in the outcome variable before the treatment (cf. Figure 2.3). Consequently, the remaining difference between these group-specific differences is assumed to reflect the causal effect of interest. This aims to take into account both unobserved group characteristics and characteristics of the timing of the intervention that are common to both groups (Legewie, 2012; Zeldow & Hatfield, 2019). The crucial assumption required for the Difference-in-Differences approach to identify the causal effect is that the group-specific trends in the outcome of interest would be identical in the absence of treatment (Angrist & Pischke, 2009; Zeldow & Hatfield, 2019). In specific applications, it is necessary to critically discuss this untestable, crucial assumption of the Difference-in-Differences approach.

Figure 2.3: Identification of the Difference-in-Differences Estimator

Note: The figure is based on the application of the Difference-in-Differences approach in *Study 2*, but has been modified slightly for illustration purposes.

This dissertation addresses the challenge of identifying the potential associations of *de-tracking* reforms with socioeconomic disparities as adequately as possible. The data basis allows for a comparison of groups of states before and after the implementation of the *de-tracking* reforms, one group with *de-tracking* reforms and one group without *de-tracking* reforms. In this way, states without *de-tracking* reforms can serve as comparison target and time-invariant characteristics of the states are taken into account. However, the states with *de-tracking* reforms do not represent a comprehensive school system without between-school tracking practices, but a tracking system reduced to a two-path school structure. This means that the analyses cannot draw any conclusions about differences between a comprehensive secondary school system and between-school tracking system. Rather, the results aim to provide information on possible changes induced by *de-tracking* reforms from a multi-track to a comprehensive non-academic school structure (cf. Figure 2.3; Van de Werfhorst, 2018, 2019). This identification strategy is comparable to the one used by Piopiunik (2021) to analyzed the effects of *de-tracking* reforms in Germany on student

achievement. Van de Werfhorst (2018, 2019) applied a similar approach, but with comparisons between countries with and without reforms.

The analytical approach described differs from the approaches of many studies investigating tracking effects, e.g., in longitudinal analyses, which tend to take into account different developments in student achievement or socioeconomic disparities between different school tracks (e.g., M. Becker et al., 2006; Dockx et al., 2019; Pfof et al., 2010; Retelsdorf & Möller, 2008; Traini et al., 2021). Those studies compare learning developments of students in the different tracks and examine possible achievement gaps or the change in the association between educational outcomes and students' socioeconomic backgrounds. In doing so, they try to identify how the different tracks affect students' learning, e.g., by comparing student achievement in academic track schools and in the non-academic track schools. Another way of investigating tracking effects is based on cross-national comparisons. These studies describe differences between countries with different tracking practices (e.g., Dämmrich & Triventi, 2018; Duru-Bellat & Suchaut, 2005; Schütz et al., 2008; Skopek & Passaretta, 2021; Skopek et al., 2019; Van de Werfhorst & Mijs, 2010) or compare changes in socioeconomic disparities from elementary to lower secondary school level between countries with different tracking systems (e.g., Ammermüller, 2005; Hanushek & Wößmann, 2006; Strello et al., 2021). Since countries could be heterogeneous and differ in a number of other potentially relevant unobserved characteristics, the estimations of studies with cross-national comparisons could be biased due to omitted variables.

In summary, the analyses of the latter approaches aim to answer the question of how different tracking practices (e.g., between-school tracking practices in comparison to comprehensive schooling or an earlier compared to a later onset of between-school tracking practices) are associated with student outcomes and socioeconomic disparities. This dissertation aims to identify possible changes in student outcomes or socioeconomic disparities as a result of the introduction of *de-tracking* reforms – as opposed to the case where no *de-tracking* reforms were implemented.

Although both studies of this dissertation are interested in identifying the effects of

the *de-tracking* reforms, the focus and implementation of the analyses differed. Due to the complexity of the adequately specifying SEC effects (see 2.1.6), *Study 1* estimates and compares school type-specific SEC effects before and after the introduction of the reforms in the states with *de-tracking* reforms using multilevel-multigroup analyses. In addition, the difference in the estimators of SEC effects between states with *de-tracking* reforms and states without *de-tracking* reforms was calculated following the basic idea of the Difference-in-Differences approach. In *Study 2*, the Difference-and-Differences approach is applied to estimate the effects of the de-tracking reforms.

Using state-of-the-art identification strategies, the studies aimed to identify and estimate the effects of *de-tracking* reforms based on secondary analyses without planned experimental settings. Given the methodological considerations outlined above, the following section summarizes the objectives of this dissertation.

3

Objectives of the Dissertation

Previous research investigating the relationship between tracking practices and socioeconomic disparities has primarily been based on comparisons between countries with different tracking practices (cf. 2.1.4; e.g., Ammermüller, 2005; Hanushek & Wößmann, 2006; Skopek & Passaretta, 2021; Skopek et al., 2019; Strello et al., 2021; Terrin & Triventi, 2023; Van de Werfhorst & Mijs, 2010; Van de Werfhorst, 2018, 2019). However, these cross-national comparisons might to be subject to bias arising from omitted variables due to unobserved differences between countries (e.g., Piopiunik, 2021; Waldinger, 2007; Woessmann, 2016). Other than the comparison between countries, the similarities between the school systems of the German federal states are assumed to be accompanied with fewer unobserved differences due to agreements between the states (cf. KMK, 2020) that might affect the estimation of the association between tracking practices and socioeconomic disparities (cf. Piopiunik, 2021). As only some states have introduced *de-tracking* reforms, a comparison of states with and without *de-tracking* reforms before and after their introduction provides an opportunity to extend existing research findings on the relationship between tracking practices and socioeconomic disparities (cf. 2.3).

The way between-school tracking practices are implemented determines to a large extent how students are distributed across different school types (cf. 2.1.1; Baumert

et al., 2006b; Dumont, Neumann, Maaz, et al., 2013; Maaz et al., 2008). In turn, the school type a student attends decisively shapes his or her educational path and opportunities on the labor market (Buchmann & Dalton, 2002; Parker et al., 2016). The relationship between socioeconomic disparities and between-school tracking practices is intensively debated in Germany on the basis of international and national large-scale assessment studies, also due to above-average socioeconomic disparities in international comparisons (e.g., Kuhl et al., 2016; Mahler & Kölm, 2018; Müller & Ehmke, 2016; Niemietz et al., 2023; OECD, 2019). Socioeconomic disparities in systems with between-school tracking practices are largely explained by primary, secondary and tertiary effects of origin, which become particularly significant in the transition from elementary to lower secondary school (Baumert et al., 2006a; Boudon, 1974; Esser, 2016; Gresch, 2012; Maaz et al., 2008). In addition, between-school tracking practices have the side-effect of school-specific, vertical socioeconomic segregation, which leads to a different SEC of the student body at the individual school types (cf. 2.1.4 and 2.1.5; e.g., Baumert et al., 2006a; Hanushek & Wößmann, 2006; Maaz et al., 2014; Maaz et al., 2008; Strello et al., 2022; Van de Werfhorst & Mijs, 2010). A less favorable SEC of the student body in schools or classes appears to contribute to socioeconomic disparities in student achievement (cf. 2.1.6; e.g., Baumert et al., 2006b; Tan et al., 2023; Van Ewijk & Sleegers, 2010).

This dissertation examines whether the *de-tracking* reforms contribute to a change in socioeconomic disparities in terms of SEC effects on student achievement (cf. 2.2.3 and 2.2.4) as well as in terms of students' educational aspirations (cf. 2.2.5). To answer these questions, the effects of *de-tracking* reforms recently introduced in some German states, which transformed the up to four non-academic school types into a two-path school structure, are analyzed (cf. 2.2.2 and 2.3). In the two-path school structure, students can continue their school career until they acquire the general university entrance qualification in both academic track schools and comprehensive schools. These *de-tracking* reforms aim to reduce disparities in student achievement, their educational aspirations and educational attainment in the non-academic track schools, particularly with respect to students' socioeconomic backgrounds.

Due to changes in the non-academic tracks as a result of the *de-tracking* reforms, this dissertation addresses the question of whether the SEC in the non-academic school types has changed (cf. 2.2.3). In addition to answering this question, *Study 1* examines if *de-tracking* the non-academic tracks has affected the association of SEC and student achievement (cf. 2.2.4). Shared values among classmates are discussed to contribute to the explanation of SEC effects according to which students share norms, values and expectations facilitating student learning (e.g., Tenenbaum et al., 2020). Thus, *Study 1* also investigated whether shared values among students, measured by the level of students' educational expectations, contributed to the intended reduction of SEC effects on student achievement. After *de-tracking* reforms, the reformed school type offers students the opportunity to pursue all educational paths at this school type. *Study 1* therefore also examined whether the proportion of students expected to obtain the general higher education entrance qualification – as an indicator of shared values among students – contributed to a change of SEC effects.

In addition to student achievement, in systems with between-school tracking practices, students' educational aspirations are closely linked to the school types and the school-leaving qualifications they offer (cf. 2.1.3; e.g. Buchmann & Dalton, 2002; Buchmann & Park, 2009; Maaz et al., 2008; Zimmermann, 2020). Students' educational aspirations are – together with their achievement – an important determinant of the acquisition of educational qualifications and subsequent educational success (e.g., Beal & Crockett, 2010; Bozick et al., 2010; Sewell et al., 1970; Sewell et al., 1969). Between-school tracking practices tend to contribute to socioeconomic disparities in students' educational aspirations (cf. 2.2.5; e.g., Bittmann & Schindler, 2021; Buchmann & Dalton, 2002; Buchmann & Park, 2009; Maaz et al., 2008; Zimmermann, 2020). Therefore, *Study 2* analyzes how *de-tracking* of the non-academic tracks to comprehensive schools, which offer all educational paths, affects students' educational aspirations and the associated socioeconomic disparities.

4

Are Socioeconomic Composition Effects on Student Achievement Less Relevant after *De-Tracking* Reforms in Germany?

Holtmann, M., Rjosk, C., Lüdtke, O. & Stanat, P. (2023). *Are Socioeconomic Composition Effects on Student Achievement Less Relevant after De-Tracking Reforms in Germany?* [Manuscript submitted for publication]¹

Abstract

Recent *de-tracking* reforms in some German federal states converted or merged the two lower track secondary school types into a comprehensive one, where students can achieve all school-leaving qualifications. The goals of the reforms were to facilitate

¹For detailed information on the authors' contribution, see Appendix A.2.

desegregation of students from different socioeconomic backgrounds and to reduce educational disparities. We investigated whether socioeconomic composition (SEC) and SEC effects have, in fact, changed with the *de-tracking* reforms. Moreover, we examined whether a possible increase in students' educational expectations of achieving the general qualification for university entrance contributes to the intended reduction of SEC effects. We used data from two representative large-scale assessment studies carried out in 2009 and 2015 covering the period of reforms. The SEC changed in non-academic school types. SEC effects on students' reading achievement tended to decrease, but are still significant in the comprehensive schools. The level of students' educational expectations did not mediate the SEC effects.

Keywords

Socioeconomic composition, socioeconomic composition effects, de-tracking reforms, students' educational expectations, mediation, large-scale assessment studies

4.1 Introduction

Disparities in the educational system associated with students' socioeconomic backgrounds have been discussed for more than twenty years (e.g., Contini & Cugnata, 2020; Hanushek & Wößmann, 2006; Kim et al., 2019; J. Liu et al., 2022; Terrin & Triventi, 2023). Class and school characteristics related to student body composition are considered relevant for student learning (M. Becker et al., 2022; Coleman et al., 1966; Hamre & Pianta, 2010; Tan et al., 2023). In addition to the mean achievement of a class, the socioeconomic composition (SEC) is viewed as particularly important for students' learning development (i.e., Opdenakker & Van Damme, 2001; Palardy, 2013; Sciffer et al., 2022; Tan et al., 2023; Van Ewijk & Slegers, 2010). Studies on SEC effects indicate that students in classes with lower mean socioeconomic status (SES) perform less well, independent of their own background characteristics.

Classes differ in their mean achievement and SEC depending on, among other things, the tracking system, i.e., the extent to which the school system assigns students to different school types based on their achievement in elementary school (Baumert et al., 2006a; Dumont, Neumann, Maaz, et al., 2013; Sciffer et al., 2022). In Germany exists a prototypical between-school tracking system with an early start after 4-6 years of elementary schooling (M. Becker et al., 2017). Yet, de-tracking reforms have been carried out during the last years in some federal states. Their goals are to facilitate desegregation of students from different backgrounds and to reduce adverse composition effects and educational disparities (Baumert et al., 2017). By converting or merging vocational track, intermediate track, and comprehensive schools into comprehensive schools, all educational paths and school-leaving qualifications can be reached at comprehensive schools. The class composition might become more favorable especially for students from disadvantaged socioeconomic backgrounds who primarily attended the vocational track, i.e. the lowest track, before the reforms.

One theoretical explanation for the existence of SEC effects is that students in classes with higher mean SES share norms, values and expectations facilitating student learning (e.g., Tenenbaum et al., 2020). Before the reforms, each school type lead to a particular school-leaving qualification, which in turn strongly determined students' educational paths and labor-market options. In the comprehensive schools, students can strive for all school-leaving qualifications including the general university entrance qualification that previously could only be acquired at the academic track (cf. M. Becker et al., 2017). This change may lead classmates to encourage each other to develop higher educational expectations (Buchmann & Dalton, 2002), and provide low-SES students with environments more conducive to learning. This may subsequently reduce the association of SEC and student reading achievement, i.e., mediate the intended decline in SEC effects.

In the present study, we investigated whether the SEC and the SEC effects have, as intended, changed in the course of the recent de-tracking reforms in Germany by analyzing two national large-scale assessment studies before and after the implementation of the reforms. Moreover, we examined whether a possible increase

in students' educational expectations contributes to reducing the SEC effects. In order to approach these questions theoretically and empirically, we will address first how tracking and SEC are related and how de-tracking reforms may reduce this association. Next, we discuss tracking and SEC effects on students' achievement and, finally, how the de-tracking reforms might influence the SEC effects in the reformed school types.

4.2 Theoretical Background

4.2.1 Tracking, SEC and De-Tracking Reforms

In most school systems, students are in some way grouped according to their ability, with varying onset depending on the country (Chmielewski, 2014). Tracking practices in general are based on the assumption that forming more homogeneous learning groups by ability grouping facilitates student support according to their learning requirements (Ireson & Hallam, 2001; Van de Werfhorst, 2019).

How ability grouping is implemented differs between countries. In Germany – as in, for example, the Netherlands or Austria – between-school tracking is carried out, where students transfer to different school types after elementary school (after 4th or 6th grade). The decisive factor in the transition is the students' performance in elementary school. In Germany, education is a state matter, resulting in sixteen educational systems (M. Becker et al., 2017). Traditionally, students in the western German states attended one of the three school types: the academic, the intermediate or the vocational track. Each of these school types offered a single course of education in which students, including elementary school, attended school for either nine, ten, or thirteen years. The different school types led to a particular school-leaving qualification, thus shaping students' educational and labor-market path (Maaz et al., 2008). These three school types were supplemented in some states, on the one hand, by comprehensive schools which covered several courses of education by streaming students within schools with respect to certain subjects (course-by-course tracking)

and enabling students to get prepared for all school-leaving qualifications. On the other hand, they were supplemented by within-school tracking schools which streamed students within schools into different tracks for all subjects. In the eastern German states, students mainly attended either the academic track or within-school tracking schools and to a smaller extent comprehensive schools (see M. Becker et al., 2017, for further details).

Over the last ten to fifteen years, de-tracking reforms took place in some German federal states. These reforms reduced a multiple track system to a two-path school structure, mostly by converting or merging vocational track, intermediate track, and comprehensive schools. The reformed school type offers students the opportunity to pursue all educational paths at this school type (cf. M. Becker et al., 2017). Before the reform, it was mainly students at academic track schools who achieved the highest school-leaving qualification entitling them to study at university. In the reformed two-path school structure, students in comprehensive schools also have the opportunity to strive for this school-leaving qualification. However, in some states with de-tracking reforms, within-school tracking schools remain, to varying degrees, offering different tracks for all subject and leading to the corresponding school-leaving qualification. Since this school type was not affected by reforms, basically maintained the former tracked system and exists only in some states, we exclude it of our considerations.

In tracked school systems, students in the respective school types differ in their achievement and socioeconomic backgrounds because these student characteristics are typically highly correlated. Moreover, socioeconomic disparities occur at educational transitions (cf. primary and secondary effects of origin: Boudon, 1974; Maaz et al., 2008). As a result of these effects, the student bodies at different school types vary in mean achievement as well as in socioeconomic composition. The central goals of the de-tracking reforms were, amongst others, to reduce social segregation, adverse effects of student composition, and differences in learning trajectories at different school types. This was expected to ultimately raise the educational level, especially of lower performing students (cf. Baumert et al., 2013, 2017).

A few studies already investigated whether social segregation, SEC, and socioeconomic

disparities changed after the de-tracking reforms, mostly focusing on single states (Helbig & Nikolai, 2017; Lenz et al., 2019; Lenz et al., 2022; Maaz et al., 2019; Neumann, Becker, et al., 2017): A study by Lenz et al. (2019) using the same data the present study is based on showed that the proportion of students with low SES increased at the comprehensive schools in the states with reforms, probably because students who attended intermediate or vocational track schools before the reforms now mainly attend comprehensive schools. Social segregation between the academic track and non-academic school types mostly continued to exist (Lenz et al., 2022; Neumann, Becker, et al., 2017), with similar selection mechanisms in school choice observed as before the reform. The *Berlin study*, which scientifically accompanied the reform in Berlin, showed stable socioeconomic disparities in achievement and educational expectations, despite a significant increase in the expectation of achieving the general qualification for university entrance (Maaz et al., 2017). To date, we do not have a comprehensive picture across all states with reforms regarding how the SEC has changed in comprehensive schools. In addition, whether any change in SEC affected student achievement in the wake of de-tracking reforms remains unanswered so far.

4.2.2 Tracking and SEC Effects on Student Achievement

In tracked educational systems, the differences in achievement development between schools are due to both composition effects and institutional effects like differences in curricula or approaches to teaching (Baumert et al., 2006a; Maaz et al., 2008; Opdenakker & Van Damme, 2006; Terrin & Triventi, 2023). Composition effects refer to the influence of student body composition. They are defined as effects of aggregated individual characteristics at the group level controlling for the effects of the individual characteristics (Raudenbush & Bryk, 2002).

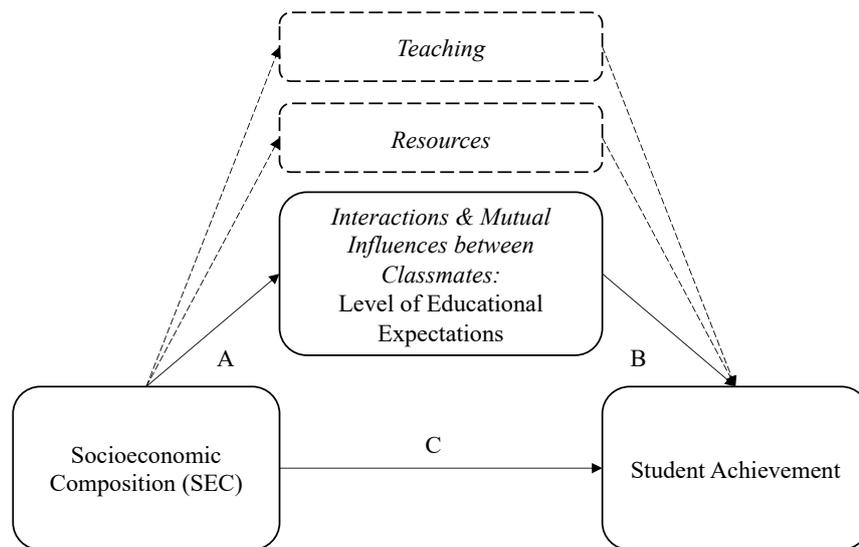
In the following, we focus on SEC effects that have consistently been identified in international and national studies (Hanushek et al., 2003; Opdenakker & Van Damme, 2001; Tan et al., 2023; Van Ewijk & Sleegers, 2010). The findings

suggest that students in classes with higher mean SES perform better regardless of their individual background characteristics. Taking institutional effects into account, Dumont, Neumann, Maaz, et al. (2013) summarized for Germany in a review that, after controlling for mean achievement, the SEC still has a positive association with students' achievement.

In tracked school systems, composition effects may differ between school types (Baumert et al., 2006a). Yet, there are only a few studies on school-type-specific composition effects in Germany. Small SEC effects were observed at the academic track schools while no significant SEC effects were observed at intermediate and vocational track schools (Baumert et al., 2006a; Dumont, Neumann, Nagy, et al., 2013). Stanat (2006) concluded from her analyses of vocational track schools that effects of SEC, mean achievement and the share of immigrant students are hard to disentangle showing overall less beneficial learning environments at those schools compared to schools in other tracks. For comprehensive schools, Baumert et al. (2006a) found significant SEC effects. Yet, changes in tracking and in the SEC at different school types may affect the extent to which SEC effects on student achievement occur.

4.2.3 De-Tracking Reforms and SEC Effects on Student Achievement

On the question of how the SEC affects student outcomes, three interrelated domains are commonly discussed (cf. Baumert et al., 2006a; Harker & Tymms, 2004; Palardy, 2013; Van Ewijk & Slegers, 2010): (1) teaching, (2) resources, and (3) interactions and mutual influences between classmates (see Figure 4.1). However, we will only briefly discuss the former two, as the focus of our study was on the latter domain.

Figure 4.1: Mediation Model of SEC Effects on Student Achievement

Concerning teaching, it is, for example, assumed that teachers adapt their instructional practice and organization of their lessons to the class composition (Dumay & Dupriez, 2007; Rumberger & Palardy, 2005). Teachers who teach in classes with lower mean SES and the associated lower mean achievement, are likely to lower their level of demands and pace more than necessary or deal more with discipline problems leading to lower student achievement in classes with low SEC (Palardy, 2015; Rjosk et al., 2014). In addition, the resources available to the class or school, such as parental involvement or the acquisition of adequate staff, may influence SEC effects (Baumert et al., 2006a). For example, teachers with good credentials and teaching skills typically have more employment options and might prefer working in schools with higher mean SES where they expect teaching to be easier due to fewer discipline problems and higher student motivation (Harker & Tymms, 2004).

More immediate effects could occur due to in the interactions and mutual influence between classmates. Several studies showed a possible mediation of SEC effects on achievement or college choice (e.g., Hornstra et al., 2015; H. Liu et al., 2015; Palardy, 2015). According to the notion of social or peer contagion, students influence each other's motivation and learning behavior and thus become more similar within a class over time (Erbring & Young, 1979; Kelly, 2009). In classes with higher mean SES, for example, classmates from advantaged socioeconomic backgrounds may have a positive

impact on the perception of schooling (Dupriez et al., 2012). They might exemplify corresponding values, aspirations, function as role models, and provide access to relevant information on the benefits of schooling (Dupriez et al., 2012). Their higher motivation and higher educational expectations can potentially influence classmates from disadvantaged socioeconomic backgrounds (Buchmann & Dalton, 2002; Hallinan & Williams, 1990). Positive attitudes towards school and achievement could, in turn, have positive effects on students' learning outcomes. In classes with lower mean SES, on the other hand, the appreciation of schooling might be less prevalent and the classmates may influence each other in their motivation and learning behavior in less favorable ways (Hornstra et al., 2015).

Before the reforms in Germany, vocational track schools in particular were characterized by student compositions with lower mean SES. Many low-achieving students from disadvantaged socioeconomic backgrounds attended this school type, which largely determined their educational path and their school-leaving qualification. In such a context, peers are less likely to encourage each other to strive for higher school-leaving qualifications. Also, students at vocational track schools are subject to stigmatization processes due to the unfavorable reputation of the school type, which can lead to negative collective identities and school motivation (Knigge & Hannover, 2011).

By merging vocational track, intermediate track, and comprehensive schools in some German states, direct links between the school type and the corresponding school-leaving qualification were dissolved. This decoupling aims at enabling students to adapt their educational expectations according to their achievement and to obtain higher levels of schooling (Neumann, Lühe, et al., 2017). After the reforms, students (from disadvantaged socioeconomic backgrounds) who would have attended a vocational track school now attend a comprehensive school, where they also have contact with higher-performing and higher-SES classmates. They could be encouraged and influenced by their fellow students. We do not expect a discouraging effect for students from advantaged socioeconomic backgrounds, as a general trend toward higher school-leaving qualifications emerged in Germany during the reform

period (Autorengruppe Bildungsberichterstattung, 2020). In this way, the association between SEC and educational expectations could decrease (see Figure 4.1: path A), which, in turn, could lead to the reform's intended reduction of SEC effects on students' achievement in the non-academic school types (path C).

Empirical analyses showed that educational expectations were higher in schools with higher mean SES (Frost, 2007; Meyer, 1970). Dupriez et al. (2012) found in 20 of the 28 countries studied higher educational expectations among students with comparable achievement and family backgrounds in schools with a more favorable SEC using data from PISA 2003. Buchmann and Park (2009) investigated the relationship between students' socioeconomic backgrounds, students' attendance of different school types, and their educational and occupational expectations with data from PISA 2003. In highly differentiated educational systems such as Germany, they found a strong correlation between students' socioeconomic backgrounds and the school type. The school type, in turn, was associated with students' educational expectations. Buchmann and Dalton (2002) found that interpersonal influences on students' educational expectations were not relevant in differentiated secondary school systems. However, interpersonal influences were relevant in undifferentiated secondary school systems, as the school type was not the primary determinant of students' educational expectations. The study accompanying the de-tracking reforms in Berlin found an increase in the educational expectations of non-academic track students (Neumann, Lühe, et al., 2017) suggesting that the de-tracking reforms led to changes in educational expectations. Thus, if the level of educational expectations among classmates mediated the SEC effects and the association between SEC and educational expectations diminished in the course of the de-tracking reforms (see Figure 4.1: path A), the SEC effects on students' achievement could decrease (path C).

Based on the theoretical considerations and the empirical findings presented, we pose the following research questions (RQ):

- 1) Has the SEC in the non-academic school type mainly covered by the de-tracking reforms changed after the reforms?

- 2) Have the SEC effects on student achievement decreased in the non-academic school type mainly covered by the de-tracking reforms?
- 3) Are the SEC effects on student achievement in the non-academic school type mainly covered by the de-tracking reforms mediated by the level of educational expectations in the classroom?

4.3 Materials and Methods

4.3.1 Participants

Our analyses are based on data from the National Assessment Study 2009 and the IQB Trends in Student Achievement Study 2015 carried out by the Institute for Educational Quality Improvement (IQB) in Germany (Köller et al., 2011; Stanat et al., 2018)². These large-scale assessment studies, in which one class per school was drawn, measured achievement in German, English, and French at the end of 9th grade in large representative samples for all 16 federal states in Germany. We investigated reading comprehension in German (i.e., the language of instruction) as an outcome because it represents a key requisite for acquiring competencies in all subjects.

We focused on the five German states with de-tracking reforms (Berlin, Bremen, Hamburg, Rhineland-Palatinate, and Schleswig-Holstein). Our analyses before the reforms are based on 7,052 students in 332 classes in 2009 (average number of students per classroom $M = 21$) and 4,438 students in 216 classes in 2015 after the reforms (average number of students per classroom $M = 21$). The sample included students who took part in the German language tests and who attended either vocational track, intermediate track or comprehensive schools. For a sample description, see Table 4.1.

²The datasets are available for non-commercial research at the research data centre at the IQB.

Table 4.1: Descriptive Statistics of Demographic Variables in the States With De-Tracking Reforms.

Variable	2009		2015	
	Individual Level <i>N</i> = 7,052	Classroom Level <i>N</i> = 332	Individual Level <i>N</i> = 4,438	Classroom Level <i>N</i> = 216
SES				
<i>M</i>	43.8	43.2	45.8	45.5
<i>SD</i>	19.1	8.9	19.3	8.4
Educational expectations (%)	9.8	8.80	23.6	22.3
Immigrant background (%)	30.0		37.8	
Gender (female (%))	47.0		45.6	
Age				
<i>M</i>	15.3		15.0	
<i>SD</i>	0.7		0.7	
School track (%) ^a				
Comprehensive Schools		14.7		41.2
Vocational Track		17.5		
Intermediate Track		23.8		0.01

Note. ^aThe remaining percentage of students attends either academic track or within-school tracking schools.

4.3.2 Measures

In our secondary analysis, we used information from standardized achievement tests, student and parent questionnaires, and the student participation list completed by the school coordinators of the study.

Outcome Variable.

Student achievement in reading comprehension. We used the five (2009) respectively fifteen (2015) plausible values estimated in the IQB assessment studies for each student based on standardized achievement tests (Böhme et al., 2010; Schipolowski et al., 2016). The competence estimates from 2009 were transferred to the metric of 2015 to allow for comparison. The metric of 2015 was scaled with a mean score of 500 and a standard deviation of 100 in the total sample of ninth graders in Germany (sample distribution in 2009: $M = 464.63$, $SD = 91.26$; in 2015: $M = 457.93$, $SD = 94.39$).

Student Level Background Variables.

SES. We used the highest “International socioeconomic Index of Occupational Status” (HISEI) as the measure of SES (Ganzeboom et al., 1992). This index can theoretically range from 11 to 89. The HISEI is based on a classification of the occupation of the students’ mother or father, whichever is higher.

Immigrant background. Immigrant background was categorized based on the country where the parents were born. To be classified as a student with an immigrant background, at least one parent had to be born abroad.

Gender. Dummy coding for student gender was used (0 = male, 1 = female).

General cognitive ability. The general cognitive ability of the students was only administered in 2015 and assessed with a subtest on reasoning of the Berlin test for fluid and crystalline intelligence for grades 8-10 (BEFKI 8-10: Wilhelm et al., 2014). In additional robustness checks, we took into account this measure at individual and aggregated at classroom level as a proxy for previous student achievement in these cross-sectional data (see Baumert et al., 2006a, on the validity of such a proxy).

Independent Classroom Level Variable.

SEC. At the classroom level, we aggregated the individual background information on SES as a measure for SEC.

Mediator.

Educational Expectations. The school coordinators indicated in the student participation list which school-leaving qualification each student is aiming for. We created a dichotomous variable (0= expectation of not achieving the general qualification for university entrance, 1= expectation of achieving the general qualification for university entrance). For addressing RQ3, we aggregated this dichotomous variable to the proportion of students who expect to achieve the highest school-leaving qualification to represent the level of educational expectations in the classroom.

In addition, in 2015, students were asked in the student questionnaire which school-leaving qualification they are likely to have attained when they leave school. The agreement between the student data and that of the school coordinators in 2015 was Cohen's Kappa = .70. In additional robustness checks, we calculate the analyses of RQ3 for 2015 using the student data.

4.3.3 Data Analysis

For data preparation and RQ1, we used the software *R* (v4.2.1, R Core Team, 2022). For RQ1, we estimated the changes of SEC in different school types with the *eatRep* package (Weirich et al., 2022). We analyzed the RQ2 and RQ3 conducting multilevel models using the software *Mplus* (Version 8.3; Muthén & Muthén, 1998 – 2007). We modeled the SEC effects on student achievement via mean SES, controlling for SES, immigrant background, and gender at the individual level. In order to obtain representative results, we used sample weights at student, classroom, and school

level (see Böhme et al., 2010; Schipolowski et al., 2016). Class level constructs were aggregated using the respective weights as well.

Metric background variables were standardized at the student level (SES, general cognitive ability) corresponding to centering on their grand mean. Categorical variables (immigrant background, gender, educational expectations) were not centered or standardized. We standardized class level constructs. In case of missing information on student level background variables, we used the five (2009) respectively fifteen (2015) values estimated by multiple imputation for each student in the IQB assessment studies (see Böhme et al., 2010; Schipolowski et al., 2016).

We tested whether the SEC effects differed before and after the reforms (between 2009 and 2015) conducting multilevel-multigroup analyses, with the assessment years representing the grouping variable.

For RQ1 we compared the mean SES at the classroom level, as a measure for SEC, in different non-academic school types before and after the reforms. For RQ2, we analyzed the changes in the SEC effects on students' reading comprehension in the different school types mainly covered by the de-tracking reforms (i.e. vocational track, intermediate track, and comprehensive schools). To analyze if the SEC effects on student achievement are mediated by the level of educational expectations in the classroom (RQ3), we carried out tests in four steps following the suggestion by MacKinnon et al. (2012). In all models, the change parameters (e.g., in Tables 4.3 to 4.4) were formed in *Mplus* as additional parameters via the difference of the regression weights of the corresponding years. We additionally calculated the mediation model in 2015 taking general cognitive ability into account (see Appendix, Table A.1).

4.4 Results

4.4.1 RQ1: SEC in Different School Types

When looking at the SEC in different school types, we observed a clearly lower mean SES in vocational track schools in 2009 compared to other tracks ($M = 37.76$, $SD = 8.58$; see Table 4.2). We found that the mean SES was comparable in intermediate track and comprehensive schools in 2009, but the dispersion was smaller in intermediate track schools (intermediate track: $M = 46.20$, $SD = 5.70$; comprehensive schools: $M = 46.68$, $SD = 9.22$). Across all tracks in the states with reforms, the mean SES in the class increased by $\Delta_{\text{compr}} = 2.33$ ($p < .01$). We observed a small decrease in comprehensive schools ($\Delta = -.98$, $p < .05$) and no differences between intermediate track schools in 2009 and comprehensive schools in 2015 ($\Delta_{\text{compr}} = -.50$, $p = .18$). The difference in the SEC is most pronounced in the comparison between vocational track schools in 2009 and comprehensive schools in 2015 ($\Delta_{\text{compr}} = 7.94$, $p = .00$). When comparing vocational track, intermediate track, and comprehensive schools in 2009 – the schools that were merged to comprehensive schools – to comprehensive school in 2015, the mean SES in the class increased. In these results, it should be emphasized that the situation has changed especially for former vocational track students who now attend comprehensive schools, which on average have a much more favorable SEC. The dispersion of the mean SES tended to be smaller in comprehensive schools in 2015 than in vocational track and comprehensive schools in 2009.

Table 4.2: RQ1: SEC in Different School Types in the States With De-Tracking Reforms.

	2009			2015			2009 to 2015	
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	Δ_{compr}	<i>p</i>
Vocational Track	37.76	8.58	1893				7.94	.00
Intermediate Track	46.20	5.70	2332				-.50	.18
Comprehensive	46.68	9.22	2827	45.70	7.97	4430	-0.98	.02

4.4.2 RQ2: SEC Effects in Different School Types

Next, we compared the SEC effects in comprehensive schools in 2015 with SEC effects in the three school types in 2009 mainly covered by the reforms (vocational track, intermediate track, and comprehensive schools). In comprehensive schools, a one standard deviation increase in mean SES was associated with an increase of about one third of a standard deviation in students' reading comprehension in 2009 ($\beta = 33.56$, $p = .00$, see Table 4.3)) and about one fourth of a standard deviation in 2015 ($\beta = 25.91$, $p = .00$). The estimator of SEC in comprehensive schools tended to be smaller in 2015 than in 2009, although the decline was not significant ($\Delta = -7.65$, $p = .19$). We additionally tested whether the changes of the SEC estimators in comprehensive schools differed between states with and without reforms by including dummies for groups of states and years in one model and calculating the differences as additional parameter in *MPlus*. These tests showed differences, although not significant, between the changes of SEC coefficients in the states with and without reforms (states with reforms: $\beta = -7.65$, $p = .19$; states without reforms: $\beta = 1.71$, $p = .82$, difference: $\Delta = -9.34$, $p = .33$ (additional analyses)). In the vocational track, the SEC was not associated with students' reading comprehension in 2009 ($\beta = 3.37$, $p = .86$). In the intermediate track, the SEC in 2009 was associated with students' reading comprehension to a similar extent as in comprehensive schools in 2015 ($\beta = 23.13$, $p = .01$). The SEC estimators in these two school types in 2009 were not significantly different from those in comprehensive schools in 2015, although the difference to the vocational track was relatively large (vocational track: $\Delta = 22.54$, $p = .26$; intermediate track: $\Delta = 2.78$, $p = .77$).

Table 4.3: RQ2: SEC Effects in Different School Types in the States With De-Tracking Reforms.

Dependent variable: Reading comprehension	VT			IT			CS		
	β (SE)	<i>p</i>	<i>N</i>	β (SE)	<i>p</i>	<i>N</i>	β (SE)	<i>p</i>	<i>N</i>
2009			1893			2332			2827
Classroom level									
SES (<i>M</i>)	3.37 (19.27)	.86		23.13 (8.26)	.01		33.56 (3.84)	.00	
Individual level									
SES	4.70 (3.80)	.22		0.94 (2.12)	.66		11.58 (1.87)	.00	
Immigrant Background	-19.48 (5.93)	.00		-22.43 (5.04)	.00		-28.01 (4.19)	.00	
Gender	4.21 (4.16)	.31		11.45 (4.05)	.01		8.73 (3.71)	.02	
2015									4430
Classroom level									
SES (<i>M</i>)							25.91 (4.70)	.00	
Individual level									
SES							10.4 (2.2)	.00	
Immigrant Background							-40.04 (4.29)	.00	
Gender							13.45 (3.26)	.00	
Δ SES (<i>M</i>): 2009 to 2015 (CS)	22.54 (19.81)	.26		2.78 (9.35)	.77		-7.65 (5.84)	.19	

Note. VT = Vocational Track, IT = Intermediate Track, CS = Comprehensive Schools.

4.4.3 RQ3: Mediation of SEC Effects by the Level of Educational Expectations

As students in the vocational and intermediate track generally do not aspire to the highest school-leaving qualification, we focused on the comprehensive schools in 2009. After the level of students' educational expectations was included as a mediator in the model of RQ2, SEC was significantly related to student achievement in both years, yet the size of the coefficients was reduced with no significant difference between 2009 and 2015 (see Table 4.4; 2009: $\beta = 26.39$, $p = .00$; 2015: $\beta = 22.95$, $p = .00$; $\Delta = -3.44$, $p = .59$; see Figure 4.1: path C). The SEC was positively associated with the level of students' educational expectations in both years (2009: $\beta = .33$, $p = .00$; 2015: $\beta = .24$, $p = .00$; path A). A one standard deviation increase in mean SES was associated with an increase in students' educational expectation of about one third (2009) respective one fourth (2015) of a standard deviation. This association showed a tendency to decrease ($\Delta = -.09$, $p = .26$). When both, SEC and the level of students' educational expectations were included in the regression model, the level of educational expectations was not related to students' reading comprehension (2009: $\beta = 5.94$, $p = .36$; 2015: $\beta = -6.07$, $p = .94$; $\Delta = -12.01$, $p = .25$; path B). The additional calculated product terms of the indirect paths (A x B) and their change were also not significant (2009: $\beta = 1.97$, $p = .36$; 2015: $\beta = -1.47$, $p = .48$; $\Delta = -3.44$, $p = .25$). However, at the individual level, students' educational expectations were strongly associated with their reading comprehension (2009: $\beta = 55.98$, $p = .00$; 2015: $\beta = 58.69$, $p = .00$).

Table 4.4: RQ3: Mediation of SEC Effects by the Level of Educational Expectations – Comprehensive Schools

	2009		2015		Change 2009 to 2015	
	β (SE)	<i>p</i>	β (SE)	<i>p</i>	Δ (SE)	<i>p</i>
Classroom level						
Reading Comprehension on						
SES (<i>M</i>)	26.39 (4.56)	.00	22.95 (4.98)	.00	-3.44 (6.38)	.59
Educational Expectations (%)	3.60 (6.48)	.58	-6.66 (8.48)	.43	-10.26 (10.43)	.33
Educational Expectations on						
SES (<i>M</i>)	0.33 (0.06)	.00	0.24 (0.06)	.00	-0.09 (0.09)	.26
Individual level						
SES	7.61 (1.80)	.00	7.72 (2.14)	.00	0.10 (2.77)	.97
Educational Expectations	56.90 (4.17)	.00	59.10 (5.14)	.00	2.20 (6.60)	.74
Immigrant Background	-26.21 (4.02)	.00	-37.54 (4.25)	.00	-11.33 (6.09)	.06
Gender	4.04 (3.45)	.24	9.59 (3.26)	.00	5.56 (4.66)	.23
<i>N</i> (Students)	2827		4438			
<i>N</i> (Classes)	126		216			

4.5 Discussion

The aim of this study was to examine how the SEC changed in the course of de-tracking reforms and whether SEC effects on student achievement were less pronounced after the de-tracking reforms in Germany. We focused on vocational track, intermediate track, and comprehensive schools, which were converted or merged from several school types to comprehensive schools in five German states. One of the key changes in these reforms is that students now have the option of continuing their school career at comprehensive schools like in the academic track until they reach the general university entrance qualification. By analyzing representative large-scale assessment studies from 2009 and 2015, we were able to cover the period of the reforms. Our findings suggest that the SEC in comprehensive schools after the reforms is at a comparable level to the SEC in intermediate track schools before the reforms and higher compared to the SEC at vocational tracks before the reforms. The SEC effects on student achievement tended to be smaller in the comprehensive schools after the reforms, although the decrease was not significant. The results further indicate that the SEC effects were not mediated by the level of educational expectations of the students in a classroom.

Before the reforms, students from disadvantaged socioeconomic backgrounds had a high probability of attending vocational track schools in 2009, where the SEC was unfavorable and the learning environment was less beneficial (cf. Baumert et al., 2006a; Knigge & Hannover, 2011; Stanat, 2006). After the reforms, students from disadvantaged socioeconomic backgrounds were likely to attend comprehensive schools. Our results of RQ1 showed that comprehensive schools, on average, had a more favorable SEC compared to vocational track schools in 2009. The SEC in comprehensive schools after the reforms was at a comparable level to the SEC in intermediate track schools before the reforms.

For RQ2, the results revealed a non-significant tendency for the SEC effects in comprehensive schools to decline to the level of the intermediate track before the reforms. Previous studies had difficulties detecting significant SEC effects at the

vocational track (cf. Baumert et al., 2006a; Dumont, Neumann, Nagy, et al., 2013), although the associations of unfavorable SEC (such as we found for the vocational track) are commonly discussed (cf. Baumert et al., 2006a; Dumont, Neumann, Maaz, et al., 2013; Tan et al., 2023; Van Ewijk & Slegers, 2010). The large standard errors in the vocational track in 2009 in our study indicate that the SEC effects in these school types can only be estimated with large uncertainty. Further analyses to explore the possible reasons for the high standard errors (e.g., the distribution of analysis variables, school or class weights, and sample sizes in the vocational track) only revealed that the average class size was smaller in vocational track schools compared with the other school types. Although class sizes are indeed sometimes smaller in vocational track schools, we conducted the analyses for RQ2 excluding classes smaller than 15 students. In these models, the standard error almost halved (but was still relatively large), and the SEC tended to be positively associated with student reading achievement.

In summary, our results of RQ1 and RQ2 suggest that the SEC and SEC effects after the reforms are roughly at the same level as the intermediate track before the reforms. This implies that students from disadvantaged socioeconomic backgrounds were likely to attend schools with a more favorable SEC after the reforms.

The de-tracking reforms in Germany decoupled school type and school-leaving qualifications at the non-academic school types. Studies on students' educational expectations in the course of the de-tracking reforms showed an increase in the educational expectations of non-academic track students (Neumann, Lühe, et al., 2017). The increase in the *Berlin study* was particularly evident at those comprehensive schools where students can directly achieve the highest school-leaving qualification rather than having to change to an associated academic-track school after secondary level I (Neumann, Lühe, et al., 2017). For the states with reforms, we found a tendency of SEC effects on students' educational expectations to decrease in comprehensive schools, yet they remained significant (see Figure 4.1: path A). This tendency to decrease could be due to the fact that the school type, i.e., comprehensive school, is no longer decisive for the school-leaving qualifications and,

thus, students' educational expectations. Instead, peer influences on students' educational expectations might have become more important (Buchmann & Dalton, 2002). Students may be more motivated and ambitious due to the possibility of achieving higher school-leaving qualifications and because classmates encourage each other (Dupriez et al., 2012). These peer interactions in a class, in turn, should facilitate student learning (e.g., Tenenbaum et al., 2020). However, the level of students' educational expectations was not associated with their reading achievement (path B). Thus, the SEC effects on students' achievement were not mediated by the level of educational expectations. These results are inconsistent with other studies showing a mediation of SEC effects on achievement or college choice by peer influences (e.g., H. Liu et al., 2015; Palardy, 2015). The fact that we did not find a significant mediation could be due to our indirect operationalization of the mediator which was derived by aggregating individual educational expectations. Direct class level constructs, such as academic climate or mutual support, might be a more appropriate indicator of beneficial peer interactions.

4.5.1 Limitations and Future Research

In addition to the limitations just mentioned with respect to the measurement of the level of students' educational expectations, our study has some limitations which should be addressed in future research.

First, students were not asked about their educational expectations in 2009. Therefore, we could only use information on students' educational expectation provided by the school coordinators. We additionally performed the analyses of RQ3 with information provided by students in 2015, which showed larger effects of SEC on students' educational expectations (see Table A.1). Thus, by using the data provided by school coordinators, we tend to underestimate the effect sizes of the association between SEC and the level of educational expectations. Moreover, our analyses showed a strong association of educational expectations and achievement at the individual level. Future studies should investigate the role of individual students' educational

expectations for socioeconomic disparities in the course of de-tracking reforms.

Second, the analyses were based on two cross-sectional datasets so that we could not draw causal inferences. The data covered the period in which the de-tracking reforms were introduced, although we could not rule out different developments in the individual states or other systematic differences. Furthermore, the restructuring measures may take longer to be fully implemented and take effect.

Third, our data did not include a measure of prior ability. Not including these abilities is thought to contribute to an overestimation of SEC effects (G. J. Duncan et al., 2004). Therefore, future studies, preferably longitudinal studies, should address this issue.

4.5.2 Conclusions

Different tracking systems are one of the starting points being investigated in terms of reducing socioeconomic disparities (Chmielewski, 2014; Sciffer et al., 2022; Strello et al., 2021; Terrin & Triventi, 2023). The de-tracking reforms can serve as a field study from which important observations can also be drawn for this ongoing debate on tracking systems and socioeconomic disparities. In conclusion, in our study of German secondary schools, we observe changes in SEC in the non-academic track school types that make students from disadvantaged socioeconomic backgrounds more likely to attend schools with a more favorable SEC after the reforms. The SEC effects on student achievement tended to become smaller, but still significant, in the comprehensive schools of the states with reforms. Our results suggest that the – compared to the academic track – more unfavorable SEC in non-academic schools was associated with lower achievement even after de-tracking reforms. Whether the two-path school structure may, as intended, reduce socioeconomic disparities in the medium- to long-term should be further investigated. In the reformed two-path school structure, which gives students the opportunity to strive for the general qualification for university entrance even at the non-academic type of school, we tested the level of educational expectations as a potential mediator of the SEC effects, which our results did not confirm. Future research should further explore the question how SEC effects

can be explained.

5

The Double-Edged Sword of *De-tracking* Reforms in Germany: Using a Difference-in-Differences Approach to Estimate Effects on Students' Educational Aspirations and their Association with Students' Socioeconomic Backgrounds

Holtmann, M., Rjosk, C., Jansen, M. & Lüdtke, O. (2024). The Double-Edged Sword of *De-tracking* Reforms in Germany: Using a Difference-in-Differences Approach to Estimate Effects on Students' Educational Aspirations and their Association with Students' Socioeconomic Backgrounds. *Studies in Educational Evaluation*, 81, 101350.¹

¹For detailed information on the authors' contribution, see Appendix B.2.

Abstract

Some German states recently reformed their tracking practices by reducing the number of secondary school track options from up to four to a two-path structure, in which students can continue their school career until they acquire the general university entrance qualification in both tracks (academic and comprehensive). We examine whether these *de-tracking* reforms were, as intended, associated with a change in students' educational aspirations and whether they reduced socioeconomic disparities in students' educational aspirations. We analyze representative data of around 65,000 ninth graders from states with and without reforms from pre- and post-reform periods (2012 to 2018), using a Difference-in-Differences approach and a regression of students' educational aspirations on their socioeconomic backgrounds. Our results suggest that students' educational aspirations have increased in the wake of the *de-tracking* reforms, but that socioeconomic disparities in students' educational aspirations have increased rather than weakened. Implications for tracking practices and socioeconomic disparities are discussed.

Keywords

De-tracking reforms, students' educational aspirations, students' socioeconomic backgrounds, Difference-in-Differences approach, socioeconomic disparities in students' educational aspirations

5.1 Introduction

In recent years, the Organisation for Economic Co-operation and Development (OECD) has emphasized the demand for cognitive and higher-level skills to prevent

unemployment and to meet the changing requirements of the labor market (OECD, 2021). On the way to acquiring high educational qualifications, the corresponding educational aspirations are an important determinant (Beal & Crockett, 2010; Bozick et al., 2010; Sewell et al., 1969). Socioeconomic disparities in students' educational aspirations contribute to students' from socioeconomically disadvantaged backgrounds lower chance of achieving the high qualifications.

Many studies identify students' socioeconomic background as a key factor in the development of their educational aspirations (Bodovski, 2014; Davies & Kandel, 1981; Sewell et al., 1969). Students from socioeconomically advantaged backgrounds perform better in school on average than their peers from disadvantaged backgrounds. Therefore, as one would expect, they also have higher educational aspirations (Beal & Crockett, 2010; Bozick et al., 2010; Zimmermann, 2020). But even independent of achievement, there are significant effects of students' socioeconomic backgrounds on educational aspirations (e.g., Beal & Crockett, 2010; Zimmermann, 2018). These effects contribute to socioeconomic disparities in educational success.

Beyond their family backgrounds, research has shown that students' educational aspirations are also shaped by structural features of the educational system such as the extent of tracking practices which is very plausible given that different school tracks are often intertwined with different school-leaving qualifications that students could aspire to achieve (Buchmann & Dalton, 2002; Kerckhoff, 1977). Changes in the tracking practices were made in some German federal states in the last 15 years. These states implemented *de-tracking* reforms which reduced their school systems of multiple types of secondary schools to a two-path school structure. Before the *de-tracking* reforms, each school type was connected to a particular school-leaving qualification, which in turn strongly determined students' educational and labor-market options. After the reforms, students can attain all school-leaving qualifications in the comprehensive track including the general university entrance qualification that previously could only be acquired at the academic track (cf. M. Becker et al., 2017). The goals of the *de-tracking* reforms were, amongst others, to reduce socioeconomic disparities in students' educational aspirations and success between different school

types.

In this study, we examine whether these *de-tracking* reforms were associated with a change in students' educational aspirations. In addition, we ask whether the reforms affected socioeconomic disparities in students' educational aspirations. We answer the second question in two steps: First, we investigate if educational aspirations of students from different socioeconomic backgrounds changed in the course of *de-tracking* reforms. Second, we examine if the relationship between students' educational aspirations and their socioeconomic backgrounds changed in the course of *de-tracking* reforms. We analyze representative data from federal states with and without reforms from the pre- and post-reform periods, using a Difference-in-Differences approach and a regression of students' educational aspirations on their socioeconomic backgrounds. We focus on the realistic educational aspirations of ninth graders, in which students are assumed to be able to make a mature assessment of their remaining school career given their actual academic performance. With our analyses, we contribute to the ongoing discussion of tracking practices and socioeconomic disparities.

5.2 Theoretical Background

5.2.1 Students' Educational Aspirations and Their Socioeconomic Backgrounds

Students' educational aspirations reflect their evaluation of their likely or possible educational attainment (Haller, 1968). Typically, they refer to distinct available educational alternatives in ascending order of difficulty (Haller, 1968). Student's realistic educational aspirations correspond to their educational intentions or plans, while student's idealistic educational aspirations represent their wishes regarding their educational career. In their realistic educational aspirations, students usually also take into account institutional characteristics and barriers (Finger, 2016; Kerckhoff, 1977).

In this study, we focus on students' realistic educational aspirations.

Following the theories of reasoned action and planned behavior, students' educational aspirations are an important predictor of decisions among possible alternatives in educational careers (Ajzen, 1985; Madden et al., 1992). Because they shape students' everyday behavior at school and thus their path through the educational system, students' educational aspirations are a key factor in educational success (e.g., Beal & Crockett, 2010; Bozick et al., 2010).

Educational aspirations have been suggested as one factor contributing to socioeconomic disparities in educational success. Many studies reveal a positive association between students' socioeconomic backgrounds and their educational aspirations (Bodovski, 2014; Sewell et al., 1969). On average, students from socioeconomically disadvantaged backgrounds tend to have lower educational aspirations than their more privileged peers. The link between students' socioeconomic backgrounds and their educational aspirations is partly due to differences in achievement (e.g., see Parker et al., 2016; Van de Werfhorst & Hofstede, 2007). On average, students from socioeconomically advantaged backgrounds perform better and have thus (reasonably) higher educational aspirations. However, in addition, two students attending the same type of secondary school and performing at the same level are likely to have different educational aspirations if they come from different socioeconomic backgrounds, as described within the framework of secondary effects of origin by Boudon (1974) (see also, e.g., Parker et al., 2016; Van de Werfhorst & Hofstede, 2007).

The discussion of secondary effects in shaping educational aspirations often refers to the Wisconsin Model of Status Attainment, which focuses on social influence processes of parents (e.g., Sewell et al., 1969). It assumes that students from socioeconomically disadvantaged backgrounds face lower educational norms, values, attitudes toward achievement, and beliefs about their educational development and success, for example, held by their parents (Stocké et al., 2011). Next to the status attainment theory, rational choice theory (and similar theories such as expectancy value theory) assumes that students' educational aspirations result from a calculation

about costs, benefits, and likelihood of success of an educational decision and that this calculation might differ based on socioeconomic backgrounds (for example with lower income families weighting the costs of education more strongly) (Breen & Goldthorpe, 1997; Wigfield & Eccles, 2000). Students' efforts to maintain the socioeconomic status of their parents can also be a central motive in explaining differences in educational aspirations due to students' socioeconomic backgrounds (Breen & Goldthorpe, 1997; Van de Werfhorst & Hofstede, 2007).

Both the status attainment and the rational choice theory and their determinants contribute to explaining the secondary effects of students' socioeconomic backgrounds on their educational aspirations (e.g., see Zimmermann, 2020). Zimmermann's results were found in Germany, a country with a between-school tracking system, which is of high relevance when considering the link between students' socioeconomic backgrounds and their educational aspirations (e.g., Buchmann & Dalton, 2002; Buchmann & Park, 2009; Maaz et al., 2008).

5.2.2 Students' Educational Aspirations in Tracked School Systems

In most school systems, students are grouped in some manner based on ability (Chmielewski, 2014). Ability grouping is rooted in the assumption that creating more homogeneous learning groups facilitates learning (Ireson & Hallam, 2001). How ability grouping is implemented differs between countries. In Germany, a between-school tracking has traditionally existed, where students transfer to different school types after elementary school (after 4th or 6th grade). Traditionally, these school types have directly led to different, hierarchically ordered school-leaving qualifications. In between-school tracking systems, the different pathways differ greatly in curricula. Students' changing between tracks is theoretically possible, but practically difficult (Maaz et al., 2008). In Germany, teachers provide a recommendation for the type of lower secondary school a student will attend in the future based on their performance (grades) in elementary school. Ultimately, the decision is generally up to the parents – that is, they can also derive it from the teacher's recommendation. At this transition,

aforementioned primary and secondary effects of origin are differentiated (cf. Boudon, 1974), and cost-benefit trade-offs related to students' socioeconomic backgrounds according to the rational choice theory or expectancy-value model come into play.

The placement into one of the school types gives students a clear signal about their abilities and future place in society (Buchmann & Dalton, 2002; Parker et al., 2016) and limits their scope for action in the educational pathway (Buchmann & Dalton, 2002; Parker et al., 2016). Between-school tracking systems are critically discussed as early perpetuators of socioeconomic disparities in students' educational aspirations, before students complete their education (Buchmann & Park, 2009).

Several studies address students' educational aspirations, their socioeconomic backgrounds, and between-school tracking (e.g., Buchmann & Park, 2009; Lee, 2014). Buchmann and Dalton (2002) found that school type strongly determines educational aspirations in between-school tracking systems. Thus, in between-school tracking systems the space for social influences of parents and peers is limited once students are in a track. However, Buchmann and Park (2009) found that in between-school tracking systems, students' socioeconomic backgrounds influence their assignment to different school types. Students from socioeconomically advantaged backgrounds are more likely to attend academic track schools that allow a direct transition to university. The school types in turn determine the educational aspirations of students by suggesting the most realistic and likely educational pathway. This is in line with a study by Lee (2014) which showed that the direct link between students' socioeconomic backgrounds and their educational aspirations was not significant when the association between students' socioeconomic backgrounds and school type was taken into account. She argues that the impact of students' socioeconomic backgrounds may be reflected in the choice of school type (Lee, 2014).

De-tracking, on the other hand, usually aims at reducing socioeconomic disparities in educational success, i.e., in the school-leaving certificates. Offering comparable educational pathways in the different school types might contribute to reducing socioeconomic disparities in students' educational aspirations. The *de-tracking* reforms in some German states provide a unique opportunity to examine the relationship

between students' socioeconomic backgrounds and their aspirations in the context of changing tracking policies more closely.

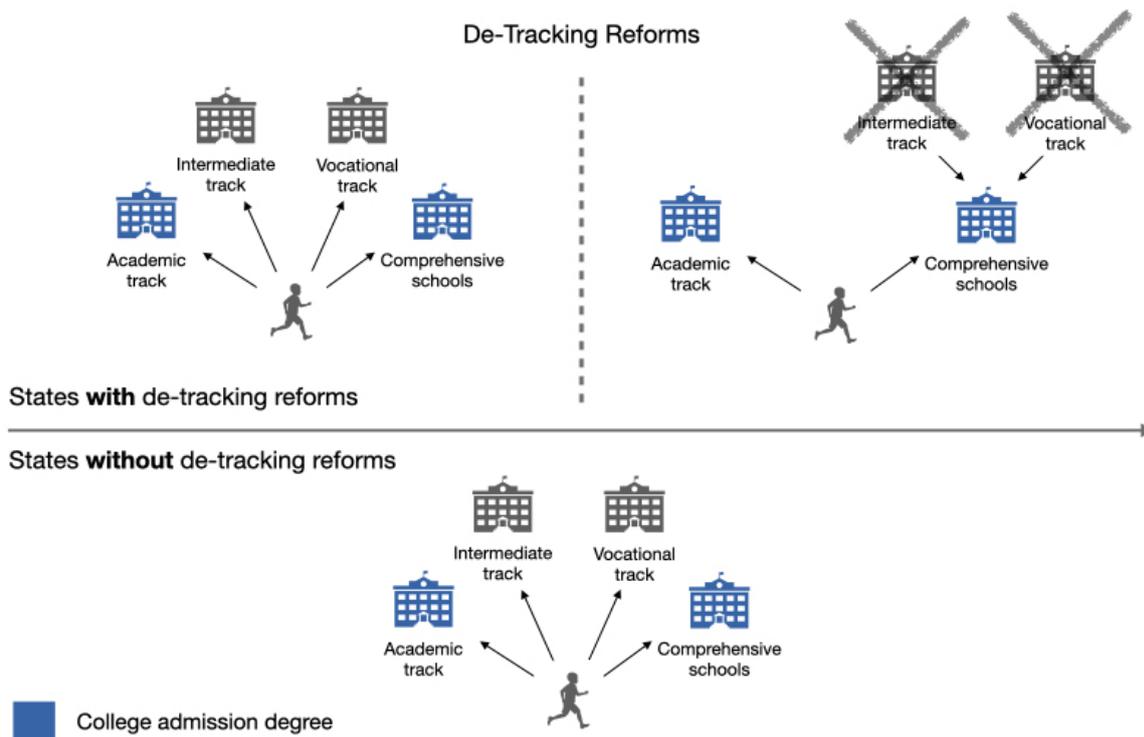
5.2.3 *De-tracking Reforms in Germany*

Germany is made up of 16 federal states, such as Bavaria or Saxony, comparable to the federal states of the USA. The federal states of Germany shape educational policy on the basis of federalism, which means that each state can, in principle, organize its school system on its own responsibility (M. Becker et al., 2017). The Standing Conference of the Ministers of Education and Cultural Affairs of the all federal states aligns educational policies between the states to a certain extent, e.g., with regard to basic structure, teaching methods, and curricula in order to ensure comparable educational standards and qualifications (KMK, 2020). Still, school policy in Germany is considered a state rather than a federal issue. Traditionally, students in the western German states attended one of the three school types: the academic, the intermediate or the vocational track. Each of these school types offered a single track of education and lead to a particular school-leaving qualification (Maaz et al., 2008). These three school types are supplemented in several states with one or even two additional school types. One of them is the comprehensive school type which covers several courses of education in one school. In comprehensive schools, students are usually streamed within the school with respect to certain subjects (course-by-course tracking). The comprehensive schools enable students to achieve all school-leaving qualifications. In some states, within-school tracking schools additionally exist. They stream students within the school into different tracks for all subjects. In the eastern German states, students mainly attend either the academic track or within-school tracking schools and to a smaller extent comprehensive schools (see M. Becker et al., 2017).

Over the last ten to fifteen years, *de-tracking* reforms took place in some German states. The *de-tracking* reforms aimed at raising the educational level, especially of lower achieving students (cf. Baumert et al., 2017). Reducing differences in learning trajectories at different school types and the link between students' socioeconomic

backgrounds and their educational aspirations should contribute to achieving the goals of the *de-tracking* reforms (cf. Baumert et al., 2017). The states with *de-tracking* reforms reduced their school systems with multiple school types to a two-path school structure, mostly by converting or merging vocational track, intermediate track, and comprehensive schools. The reformed comprehensive school type offers students the opportunity to reach all school-leaving qualifications (see Figure 5.1, cf. M. Becker et al., 2017). This means that students in these states can choose between two school types, academic track schools and comprehensive schools. In both, students can obtain the general university entrance qualification (the highest school-leaving qualification). It is an open question whether these changes have led to a change in students' educational aspirations.

Figure 5.1: Study Design: Tracking Practices in States With De-tracking Reforms and States Without Reforms



5.2.4 Students' Educational Aspirations and *De-tracking* Reforms in Germany

By merging vocational track, intermediate track, and comprehensive schools, direct links between school types and the corresponding school-leaving qualification were dissolved. This means that in a two paths system, the school type should no longer be the primary determinant of students' educational aspirations.

The decoupling of school type and school-leaving qualification aims at enabling students to increase their educational aspirations (or, at least, adapt them more flexibly throughout their school career based on their performance) and to obtain higher qualification levels (Baumert et al., 2017; Neumann, Lühe, et al., 2017). Before the reforms, students from socioeconomically disadvantaged backgrounds were likely to attend vocational track schools and develop the corresponding educational aspiration that lead them to finish their school career after nine years (Buchmann & Park, 2009; Maaz et al., 2008). They were subject to stigmatization processes due to the unfavorable reputation of the school type, which can lead to negative collective identities and school disengagement (Knigge & Hannover, 2011). After the reforms, they are likely to attend comprehensive schools that offer them the opportunity to obtain all school-leaving qualifications, which in itself is likely to broaden the educational options that students consider in their aspiration (e.g., because they do not have to change school type to acquire the general university entrance qualification). This openness may yield space for students to interact with diverse peers who can be an important alternative source of influence beyond the family (Buchmann & Dalton, 2002; Lorenz et al., 2020). Classmates from socioeconomically advantaged backgrounds may have a positive impact on the perception of schooling (Dupriez et al., 2012; Hornstra et al., 2015). They might exemplify corresponding values, aspirations, function as role models, and provide access to relevant information on the benefits of schooling (Dupriez et al., 2012). A discouraging effect for students from socioeconomically advantaged backgrounds in the direction of lower school-leaving qualifications seems less likely, as a general trend toward higher school-leaving

qualifications has emerged in Germany during the reform period (Autorengruppe Bildungsberichterstattung, 2020).

Buchmann and Park (2009) show that students' socioeconomic backgrounds strongly predict the school type they attend. Comprehensive schooling provides the chance to weaken this chain from students' socioeconomic backgrounds to an early choice of a (lower) secondary school type to the corresponding (lower) aspiration. Along these lines, Matthewes (2021) showed that achievement benefits for low-achieving students in systems with comprehensive schooling are large, while high-achieving students are not negatively affected. Moreover, he illustrates that comprehensive schooling diminishes the proportion of students with low educational aspirations (Matthewes, 2021). This is consistent with other empirical analyses showing students' educational aspirations are higher in comprehensive school systems (e.g., Buchmann & Park, 2009; Kerckhoff, 1977). The *Berlin study*, which scientifically accompanied the *de-tracking* reforms in the capital of Germany, found an increase in the educational aspirations of non-academic track students (who, on average, are more likely to come from socioeconomically disadvantaged backgrounds) to obtain the general university entrance qualification (Neumann, Lühe, et al., 2017). However, they also showed a stable, strong relationship between students' socioeconomic backgrounds and their educational aspirations, despite the general increase in aspirations (Maaz et al., 2017). In their analyses, they did not appear to control for student achievement. The trends of students' educational aspirations in the other states with *de-tracking* reforms, as well as the association with students' socioeconomic background, have not been studied yet.

5.2.5 Present Study

The *de-tracking* reforms in Germany aimed at reducing socioeconomic disparities in students' educational aspirations. In this study, we investigate whether these aims have been reached. We use a Difference-in-Differences approach to examine the development of students' educational aspirations and their relationship with students'

socioeconomic backgrounds in cohorts from the pre- and post-reform periods. Our analyses allow us to compare different tracking practices and their associations with socioeconomic disparities while controlling for existing differences between the states with and without reforms. In the wake of *de-tracking* reforms, the relationship between tracking practices and socioeconomic disparities in students' educational aspirations may be changing, partly because of expanded opportunities in comprehensive schools and partly because of smaller secondary effects of origin, as peer effects may become stronger in comprehensive schooling. We answer the following research questions:

RQ1: Did students' educational aspirations change in the course of *de-tracking* reforms?

RQ2: Did the educational aspirations of students from different socioeconomic backgrounds change in the course of *de-tracking* reforms?

RQ3: Did the relationship between students' educational aspirations and their socioeconomic backgrounds change in the course of *de-tracking* reforms?

5.3 Materials and Methods

5.3.1 Data

Our analyses are based on data from the IQB National Assessment Study 2012 and the IQB Trends in Student Achievement Study 2015 and 2018 carried out by the Institute for Educational Quality Improvement (IQB) in Germany (Pant et al., 2015; Stanat et al., 2018; Stanat et al., 2022)². The studies measured achievement in German, English, and French in 2015 and Mathematics and Science in 2012 and 2018 at the end of 9th grade. The large samples are representative for all 16 federal states in Germany.

We compare the eleven states without reforms with three states with *de-tracking*

²These datasets are available for non-commercial research at the research data centre at the IQB.

reforms. The states with *de-tracking* reforms were Berlin, Bremen, and Schleswig-Holstein. We excluded Hamburg and Rhineland-Palatinate from our analyses because they implemented *de-tracking* reforms right before 2012. The remaining states form the comparison group of states without reforms. Special education schools and academic track schools are not included in the analyses. For sample descriptions (including sample sizes), see Table 5.1.

Table 5.1: Descriptive Statistics of Dependent and Independent Variables and Sample Characteristics in States With De-Tracking Reforms and States Without De-Tracking Reforms.

		2012		2015		2018	
		with reforms	without reforms	with reforms	without reforms	with reforms	without reforms
Educational Aspirations ^a	%	20.20	19.10	31.80	18.60	42.70	28.20
Socioeconomic Backgrounds ^b							
Advantaged Group	%	44.50	46.50	48.60	46.20	48.40	48.10
Middle Group	%	34.10	31.60	30.10	31.00	30.30	30.20
Disadvantaged Group	%	21.40	21.90	21.40	22.80	21.30	21.70
Achievement	<i>M</i>	-0.72	-0.51	-0.57	-0.46	-0.81	-0.51
	<i>SD</i>	1.14	1.09	1.17	1.08	1.19	1.13
Gender (female)	%	45.20	47.70	46.80	48.30	47.70	48.20
Immigrant background	%	29.60	29.80	33.10	32.30	39.30	36.20
Age	<i>M</i>	15.32	15.33	15.05	15.21	15.65	15.71
	<i>SD</i>	0.69	0.68	0.74	0.71	0.75	0.74
School Type ^c							
Vocational Track	%	13.87	19.26	0.00	16.70	0.00	13.40
Intermediate Track	%	21.97	27.07	0.00	28.50	0.00	23.90
Comprehensive	%	20.37	7.82	53.60	10.20	57.20	13.60
Within-School Tracking	%	2.26	6.45	8.20	7.90	6.90	12.80
<i>N</i>		5,872	18,614	3,466	11,903	5,496	19,838

Note. ^aBinary variable (0= students who do not plan to obtain the general university entrance qualification, 1= students who plan to obtain the general university entrance qualification)

^bReferring to the EGP classification (Erikson et al., 1979) and assigning EGP classes to groups as in other large-scale assessment studies (e.g., Kuhl et al., 2016) ^cDue to an allocation error of comprehensive and within-school tracking schools in our data, we refer to the information provided by the Federal Statistical Office in 2012 (Statistisches Bundesamt, 2012).

5.3.2 Measures

In our secondary analyses, we used information from the student and the parent questionnaires.

Outcome Variable: Students' Educational Aspirations. Students' realistic educational aspirations are assessed by asking the students which school-leaving qualification they are likely to have attained when they leave school. We created a binary variable (0= students who do not plan to obtain the general university entrance qualification, 1= students who plan to obtain the general university entrance qualification).

Students' Socioeconomic Backgrounds. We use the EGP classification as an indicator of students' socioeconomic backgrounds (Erikson et al., 1979). The EGP classification determines the socioeconomic status of different socioeconomic groups according to the type of occupation (manual, non-manual, agricultural), the position in the occupation (self-employed, employed), the authority to issue instructions (none, low, high) and the required qualifications (none, low, high). Based on these characteristics, seven groups can be distinguished, ranging from the "upper service class" (EGP class I) to "unskilled and semi-skilled workers" (EGP class VII). Following the procedure of international large-scale assessment studies and the previous IQB studies (e.g., Kuhl et al., 2016), each child is assigned the EGP class that corresponds to the highest EGP score of the parents (mother or father). In accordance with the presentation in the IQB trends and the *Berlin study*, we assign the classes to three groups: The first group representing students from socioeconomically advantaged backgrounds includes EGP classes I and II, i.e., students whose parents work, for example, in independent academic professions or as executives or civil servants in the higher service and whose jobs are characterized by high responsibility, decision-making power and autonomy, referred to as socioeconomically advantaged group. The second group comprises EGP classes III and IV and thus the students whose parents are considered to be routine non-manual worker or small proprietors, referred to as socioeconomically middle group. The third group includes classes V to VII, i.e., students whose parents are,

for example, workers in predominantly manual work processes with a low level of requirements, referred to as socioeconomically disadvantaged group.

Student Achievement. The IQB studies include extensive assessments of students' achievement to examine if ninth grade students meet the proficiency expectations established in the national educational standards (For details on the test design and evaluation, see, e.g., B. Becker et al., 2019). We use the person scores provided in the data sets, specifically the Weighted Likelihood Estimates (WLE; Warm, 1989). For 2012 and 2018, we use students' WLE in mathematics and for 2015, students' WLE in German reading comprehension.

5.3.3 Analyses

We are interested in estimating the effect of the *de-tracking* reforms on students' educational aspirations. This is challenging, however, because we cannot observe the development of students' educational aspirations simultaneously with and without reforms. After a new policy is in effect, the Difference-in-Differences approach is one recommended way to estimate the impact of the new policy (see Angrist & Pischke, 2009). Simple comparisons of students' educational aspirations between the states with reforms and the states without reforms may be biased because differences may have existed before the reforms. To counter this, the Difference-in-Differences approach uses observed students' educational aspirations affected by the *de-tracking* reforms (as treated group) and students' educational aspirations not affected by the *de-tracking* reforms (as comparison group), both before and after the reforms. We estimate models in which we regress students' educational aspirations Y on (1) a dummy variable that indicates whether the state implemented *de-tracking* reforms (Reform; states with reforms = 1), (2) a dummy variable that indicates whether the student belongs to the pre- or post-reform period (Period; post-reforms period = 1), and (3) the interaction of the two dummy variables representing the Difference-in-Differences estimator:

$$Y_i = \gamma_0 + \gamma_1 \text{Reform} + \gamma_2 \text{Period} + \beta_{DD}(\text{Reforms} \times \text{Period}) + u_i \quad (5.1)$$

Y_i is the educational aspiration of student i (see Eq.5.1). γ_1 absorbs level differences in students' educational aspirations between states with reforms and states without reforms before the reforms. γ_2 estimates the trend in students' educational aspirations in the states without reforms. β_{DD} captures the differential development of educational aspirations for students in the states with reforms after the reforms and is thus the Difference-in-Differences estimator.

For unbiased Difference-in-Differences estimations, the key assumption is that trends in students' educational aspirations would be the same in states with and without reforms if there were no *de-tracking* reforms, the so-called common trend assumption (e.g., Angrist & Pischke, 2009). This is an assumption one cannot test because the treatment (reform) is in fact implemented. Instead, we discuss possible influencing factors that might affect the Difference-in-Differences estimation. The development of students' educational aspirations is decisively affected by student achievement (e.g., Zimmermann, 2020). Since student achievement develops differentially between states with reforms and states without reforms during the reform period and since it is associated with the treatment, it can be considered as a time-varying covariate (Zeldow & Hatfield, 2021). Our additional analyses show that, on average, mathematical achievement has decreased in the states with *de-tracking* reforms between 2012 and 2018, whereas we do not observe a change in student achievement in the states without *de-tracking* reforms, on average (see Table B.1 in the Appendix). If students' educational aspirations increase in the states with reforms, as in our case, the decrease in student achievement is more likely to weaken the reform effect. Therefore, we conclude that we do not need to constrain the interpretation of our Difference-in-Differences estimations based on the development of student achievement.

Because our outcome variable is binary, either logistic regression and linear probability models can be calculated. Without proportions close to 0 or 1 for the dependent variable, the linear and logistic models produce almost indistinguishable results,

except that the logistic estimates are more difficult to interpret (Hellevik, 2009). The comparison of logistic regression and linear probability models revealed similar results in our case. For a better interpretation of the coefficients, we report the results of the linear probability models. For binary outcome variables, Puhani (2012) showed that the Difference-in-Differences estimator can be taken directly from the interaction term in the model.

Our data allow us to estimate the effect of the reforms for two post-reform time points, 2015 and 2018. We report the Difference-in-Differences estimators for both time points separately to discuss direct and longer-term observations. We analyze our models with the software *Mplus* and take into account the hierarchical structure of students nested in classes to estimate standard errors correctly (Muthén & Muthén, 1998 – 2007). In order to obtain representative results, we used sample weights at student, classroom, and school level (e.g., Schipolowski et al., 2016). In case of missing information on students' educational aspirations (4.7 % (2012), 6.7 % (2015), and 14.6 % (2018)), their socioeconomic backgrounds (11.5 % (2015), 15.4 % (2018), and 28.0 % (2012)), their achievement (27.4 % (2015), 43.9 % (2018), and 44.3 % (2012))³, and the variables for the sample description, we estimate fifteen values by multilevel *Multivariate Imputation by Chained Equations* (with 50 iterations per imputation), which is implemented in the package *mice* (Van Buuren & Groothuis-Oudshoorn, 2011) for the statistical software *R* (R Core Team, 2022). For the multilevel imputation model, we use fully conditional specification and additionally include the aggregates of student level variables (e.g., class means of educational aspiration and SES) as predictors for the imputation of the missing information.

To answer RQ1, we analyze Difference-in-Differences models with 2015 and 2018 as post-reform time points to estimate the effect of the *de-tracking* reforms on students' educational aspirations.

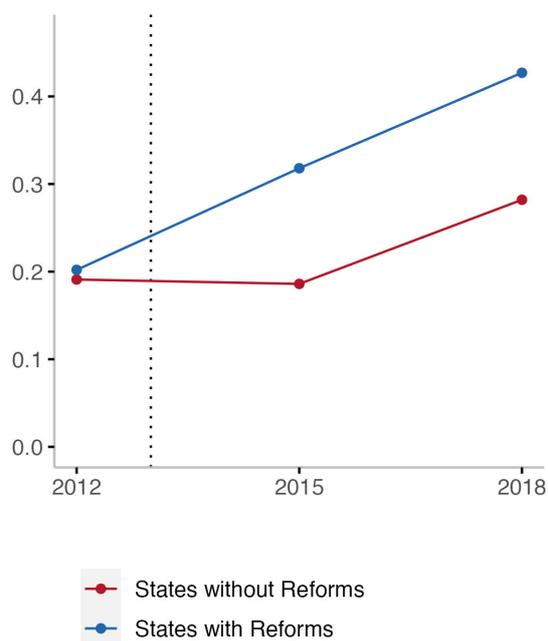
For RQ2, we estimate Difference-in-Differences models with 2015 and 2018 as post-reform time points for each socioeconomic group.

³In 2012 and 2018, students were randomly assigned to science or mathematics booklets based on a rotation design.

To answer RQ3, we analyze the secondary effects of students' socioeconomic backgrounds on their educational aspirations in which we regress students' educational aspiration on their socioeconomic backgrounds (dummy coding) under control of student achievement. In doing so, we conduct multi-group analyses in states with reforms and states without reforms separately, with the assessment years representing the grouping variable. We use the difference in the coefficients to make statements about whether the association between students' socioeconomic backgrounds and their educational aspirations has changed between the assessment years.

5.4 Results

Effects on students' educational aspirations. We present a figure showing the trends of the shares of students who plan to obtain the general university entrance qualification in the states with reforms and the states without reforms. We observe different trends in students' educational aspirations for states with and without *de-tracking* reforms (see Figure 5.2). Especially the development between 2012 and 2015 is diverging. The share of students who plan to obtain the general university entrance qualification increased by 11.6% from 20.2% in 2012 to 31.8% in 2015 in the states with reforms, whereas we observe no changes in the states without reforms. The increase between 2015 and 2018 was roughly comparable in states with reforms (10.9%) and states without reforms (9.6%) despite different starting points in 2015 (states with reforms: 31.8%; states without reforms: 18.6%).

Figure 5.2: Trends in Students' Educational Aspirations

Note. The vertical dotted line represents the implementation of *de-tracking* reforms.

The Difference-in-Differences analyses confirm the descriptive patterns (see Table 5.2). The interaction term, i.e., the Difference-in-Differences estimator, is significantly positive for both 2015 and 2018 as post reform time points (for 2015: $\beta_{DD} = .130, p < .01$; for 2018: $\beta_{DD} = .151, p < .01$). This suggests a positive effect of the *de-tracking* reforms on students' educational aspirations. In 2015, after the reforms, students have a 13% higher chance to plan to obtain the general university entrance qualification in the states with reforms than in the states without reforms in 2012 (respective 15% higher chance in 2018). The increase of students' educational aspirations between 2015 and 2018 in the states without reforms is reflected in the significant predictor of period ($\beta_{Period} = .086, p < .01$).

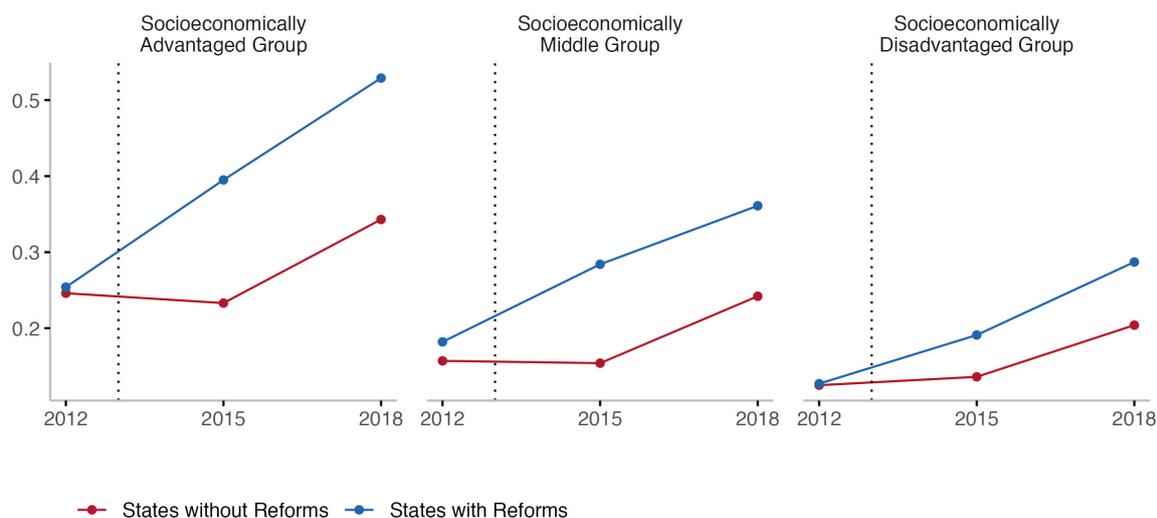
Table 5.2: Difference-in-Differences Analyses: Effects of De-Tracking Reforms on Students' Educational Aspirations

	2012 to 2015				2012 to 2018			
	β	SE	<i>p</i>	<i>N</i>	β	SE	<i>p</i>	<i>N</i>
Reform	0.004	0.019	.816	39,855	0.004	0.019	.823	49,820
Period	-0.005	0.016	.751		0.086	0.015	.000	
DiD	0.130	0.031	.000		0.151	0.027	.000	

Note. DiD = Difference-in-Differences.

Effects on students' educational aspirations according to socioeconomic groups.

The graphs of the trends in the shares of students who plan to obtain the general entrance qualification by socioeconomic groups reveal similar patterns in all groups, but different starting points and extents of increase (see Figure 5.3). In particular, between 2012 and 2015, we note different trends between the states with reforms and the states without reforms, with very little change in the states without reforms but larger changes in the states with reforms. The increases in the states with reforms range from about 6.5% in the socioeconomically disadvantaged group, about 10% in the socioeconomically middle group, and about 14% in the socioeconomically advantaged group. Between 2015 and 2018, the increases in states with reforms and states without reforms are about the same for all socioeconomic groups.

Figure 5.3: Trends in Students' Educational Aspirations According to Socioeconomic Groups.

Note. The vertical dotted line represents the implementation of *de-tracking* reforms.

For the socioeconomically advantaged group, we find significantly positive Difference-in-Differences estimators for both periods (see Table 5.3, for 2015: $\beta_{DD} = .162$, $p < .01$; for 2018: $\beta_{DD} = .192$, $p < .01$). In 2015, students have a 16.2% higher chance to plan to obtain the general university entrance qualification in the states with reforms (respective 19.2% higher chance in 2018) than students of this socioeconomic group in the states without reforms. As observed in the descriptive graph, the Difference-in-Differences estimators in the socioeconomically middle group are smaller but significant (2015: $\beta_{DD} = .109$, $p < .01$; 2018: $\beta_{DD} = .108$, $p < .01$). When comparing the three socioeconomic groups, the Difference-in-Differences estimator in the socioeconomically disadvantaged group is the smallest, but still significant for both post reform time points (2015: $\beta_{DD} = .066$, $p < .05$; 2018: $\beta_{DD} = .084$, $p < .01$). In summary, between 2012 and 2015, we find significantly positive Difference-in-Differences estimators in all socioeconomic groups indicating a positive effect of the *de-tracking* reforms on students' educational aspirations for students of all socioeconomic groups.

Table 5.3: *Difference-in-Differences Analyses: Effects of De-Tracking Reforms According to Socioeconomic Groups.*

	Socioeconomically Advantaged Group				Socioeconomically Middle Group				Socioeconomically Disadvantaged Group			
	β	SE	<i>p</i>	<i>N</i>	β	SE	<i>p</i>	<i>N</i>	β	SE	<i>p</i>	<i>N</i>
2012 to 2015												
Reform	-0.000	0.022	1.000	18,302	0.024	0.019	.205	13,027	0.001	0.017	.959	8,526
Period	-0.008	0.016	1.000		-0.002	0.013	1.000		0.006	0.015	.712	
DiD	0.162	0.034	.000		0.109	0.033	.001		0.066	0.030	.030	
2012 to 2018												
Reform	-0.001	0.022	1.000	23,390	0.023	0.019	.221	15,975	0.001	0.017	.975	10,454
Period	0.096	0.016	.000		0.083	0.014	.000		0.080	0.015	.000	
DiD	0.192	0.031	.000		0.108	0.027	.000		0.084	0.028	.003	

Note. DiD = Difference-in-Differences.

Socioeconomic disparities in students' educational aspirations. To answer RQ3, we regress students' educational aspirations on their socioeconomic backgrounds under control of their achievement. Students from the socioeconomically advantaged group serve as our reference group, the socioeconomically middle group and the socioeconomically disadvantaged group are included via dummy variables. These analyses allow us to examine whether *de-tracking* reforms have affected the secondary effects of students' socioeconomic backgrounds.

In 2012, we do not observe a difference between students from the socioeconomically middle group and students from the socioeconomically advantaged group in their educational aspirations under control of their achievement in the states with reforms (see Table 5.4, $\beta = -.024, p > .05$). However, in 2015 and 2018, students from the socioeconomically middle group have a 6.3% (2015) respective 10.4% (2018) lower chance to plan to obtain the general university entrance qualification than students from the socioeconomically advantaged group even if they show the same achievement (2015: $\beta = -.063, p < .01$; 2018: $\beta = -.104, p < .01$). These results indicate that students from the socioeconomically middle group had comparable chances to plan to obtain the general university entrance qualification in 2012 but lower chances after the reforms than students from the socioeconomically advantaged group with similar performance.

Students from the socioeconomically disadvantaged group have a between 4.9% (2012) and 14.9% (2018) lower chance to plan to obtain the general university entrance qualification than students from the socioeconomically advantaged group when performing similarly in school (2012: $\beta = -.049, p < .01$; 2015: $\beta = -.127, p < .01$; 2018: $\beta = -.149, p < .01$).

Contrary to the intended objectives of the reforms, for students from the socioeconomically middle group and from the socioeconomically disadvantaged group, the association between their socioeconomic backgrounds and their educational aspirations under control of their achievement increased significantly from 2012 to 2018 in the states with reforms (socioeconomically middle group: $\beta = .080, p < .01$; socioeconomically disadvantaged group: $\beta = .099, p < .01$).

In the states without reforms, students from the socioeconomically middle group and from the socioeconomically disadvantaged group have a lower chance to plan to obtain the general university entrance qualification than students from the socioeconomically advantaged group in all years as well (socioeconomically middle group in 2012: $\beta = -.055, p < .01$; socioeconomically middle group in 2015: $\beta = -.049, p < .01$; socioeconomically middle group in 2018: $\beta = -.069, p < .01$; socioeconomically disadvantaged group in 2012: $\beta = -.064, p < .01$; socioeconomically disadvantaged group in 2015: $\beta = -.055, p < .01$; socioeconomically disadvantaged group in 2018: $\beta = -.075, p < .01$). However, in contrast to the results in the states with reforms, the relationship between students' socioeconomic backgrounds and their educational aspirations did not increase (socioeconomically middle group between 2012 and 2015: $\beta = -.006, p > .05$; socioeconomically middle group between 2012 and 2018: $\beta = .014, p > .05$; socioeconomically disadvantaged group between 2012 and 2015: $\beta = -.009, p > .05$; socioeconomically disadvantaged group 2012 and 2018: $\beta = -.011, p > .05$).

Table 5.4: Regression Analyses: Students' Educational Aspirations on Their Socioeconomic Backgrounds.

	States with Reforms				States without Reforms			
	β	SE	p	N	β	SE	p	N
2012								
Student Achievement	0.086	0.010	.000	5,872	0.069	0.007	.000	18,614
Socioeconomically Middle Group	-0.024	0.015	.109		-0.055	0.010	.000	
Socioeconomically Disadvantaged Group	-0.049	0.016	.002		-0.064	0.011	.000	
2015								
Student Achievement	0.083	0.012	.000	3,466	0.041	0.007	.000	11,903
Socioeconomically Middle Group	-0.063	0.024	.008		-0.049	0.011	.000	
Socioeconomically Disadvantaged Group	-0.127	0.023	.000		-0.055	0.012	.000	
2018								
Student Achievement	0.135	0.010	.000	5,496	0.094	0.007	.000	19,838
Socioeconomically Middle Group	-0.104	0.019	.000		-0.069	0.010	.000	
Socioeconomically Disadvantaged Group	-0.149	0.022	.000		-0.075	0.014	.000	
Difference between								
2012 and 2015 Socioeconomically Middle Group	0.039	0.029	.184		-0.006	0.015	.692	
2012 and 2015 Socioeconomically Disadvantaged Group	0.078	0.027	.004		-0.009	0.016	.595	
2012 and 2018 Socioeconomically Middle Group	0.080	0.024	.001		0.014	0.014	.320	
2012 and 2018 Socioeconomically Disadvantaged Group	0.099	0.026	.000		0.011	0.019	.543	

5.5 Discussion

The purpose of this study was to examine whether students' educational aspirations and their association with students' socioeconomic background changed in the course of recently implemented *de-tracking* reforms in Germany. We used a Difference-in-Differences approach with representative data from states with reforms and states without reforms in the pre- and post-reform periods. Taking potential differences between states with and without reforms into account, our analyses allow us to draw conclusions about different tracking practices and their relationship with socioeconomic disparities.

In RQ1, we investigated whether the *de-tracking* reforms are related to a change in the educational aspirations of non-academic track students. Especially between 2012 and 2015, the period when ninth-graders are affected by the *de-tracking* reforms, we observe increased students' educational aspirations in the states with reforms and no change in the states without reforms. This implies that students have a higher chance to plan to obtain the general university entrance qualification in the states with reforms than in the others as intended by the *de-tracking* reforms. To encounter more general developments that are not associated with the *de-tracking* reforms, we checked the development of students' educational aspirations in each state separately. Unfortunately, the data usage agreement for our data sets doesn't allow us to show statistics on single state level. The state-by-state trends allowed us to keep interpretation of the results. Thus, we viewed all of our analyses as indicating that the *de-tracking* reforms had led to an increase in students' educational aspirations between 2012 and 2015. Between 2015 and 2018, we saw a general increase in students' educational aspirations, suggesting that this is not directly related to the *de-tracking* reforms.

In RQ2, we focused on the development of the educational aspirations of students from different socioeconomic backgrounds. We find positive Difference-in-Differences estimators in all three socioeconomic groups, but of different sizes. The Difference-in-Differences estimators in students' educational aspirations are largest for students

from the socioeconomically advantaged group and smallest for students from the socioeconomically disadvantaged group.

Finally, we regressed students' educational aspirations on students' socioeconomic backgrounds under control of student achievement in each assessment year to answer RQ3. Our last RQ aimed to examine whether the *de-tracking* reforms contributed to weaken secondary effects of students' socioeconomic backgrounds. These analyses complemented the patterns of RQ2. As expected based on decades of research, students from the socioeconomically middle and from the disadvantaged groups show lower educational aspirations on average, even when they perform similarly to students from socioeconomically advantaged backgrounds. When comparing the differences in the coefficients in socioeconomically disadvantaged group over the years, we found a significantly larger negative association of students' socioeconomic backgrounds and their educational aspirations in 2015 and 2018 in the states with reforms than in 2012. This implies that the increase in the chances to plan to obtain the general university entrance qualification is smaller in the group of students from socioeconomically disadvantaged backgrounds than in the group of students from socioeconomically advantaged backgrounds. For the socioeconomically middle group, we observe similar tendencies. We see our results as an indicator that the reforms did not manage to weaken the link between students' socioeconomic backgrounds and their aspirations. Rather, the secondary effects of students' socioeconomic backgrounds are still present and appear to have become stronger rather than weaker.

Our findings supplement observations made in the context of the *de-tracking* reforms in Berlin (Maaz et al., 2017; Neumann, Lühe, et al., 2017): In the states with reforms, we found students irrespective of their socioeconomic background having higher chances to plan to obtain the general university entrance qualification. Adjusting for achievement, we see a strong relationship, not weakened by the reforms, between students' socioeconomic backgrounds and their educational aspirations. Especially non-academic track students from socioeconomically advantaged backgrounds seem to immediately benefit from the reformed two-path school structure. One reason could be the cost-benefit calculations (Breen & Goldthorpe, 1997) respective the subjective

value (Wigfield & Eccles, 2000) regarding the school-leaving qualification. The *Berlin study* found that the proportion of students aspiring the general university entrance qualification is around twice as high at comprehensive schools with an integrated upper secondary school offer than at comprehensive schools without an integrated upper secondary school offer (Neumann, Lühe, et al., 2017). The salience of the path to the general entrance qualification might contribute to the fact that parents with a favorable socioeconomic status, who are likely to have attended school until the general university entrance qualification, consider it desirable for their children to obtain the general university entrance qualification. They might encourage their children to follow this path and aim for a comparable socioeconomic status to their own. Also the students might be motivated to achieve the educational status of their parents (cf. Breen & Goldthorpe, 1997). Bittmann and Schindler (2021) examined how the educational aspirations of high-achieving students developed from the 5th to the 9th grade, finding that students from socioeconomically disadvantaged backgrounds changed more often than students from socioeconomically advantaged backgrounds from aspiring the general university entrance qualification to aspiring a less ambitious school-leaving qualification. Students and parents from socioeconomically advantaged backgrounds might see and value the opportunities and benefits that the *de-tracking* reforms offer. On the other hand, it is conceivable that parents who attended the lower tracks of secondary schools and have achieved the associated school-leaving qualifications may not immediately see the benefits of this qualification and a longer school career. Instead, for example, the offer of vocational training after ten years of schooling and the associated job and earnings prospects could have a higher value in their calculations. Therefore, they may not encourage their children to follow this path. Andrew and Hauser (2011) show that students do not readily adjust their educational aspirations, but require a large increase in grade points to do so. These findings could also be results of risk-averse, downward adjustments of educational aspirations that Bittmann and Schindler (2021) observed among high-achieving students from socioeconomically disadvantaged backgrounds in non-academic track schools. It may take longer for the associations of the traditionally strictly tracked school system to break down for high-achieving students from socioeconomically disadvantaged

backgrounds.

In addition, motivational aspects of the students might also be important. If students attending the non-academic school type were less stigmatized after the *de-tracking* reforms, motivational gains would be conceivable (cf. Knigge & Hannover, 2011). In comprehensive schools after *de-tracking* reforms, students learn in classes with presumably wider achievement range. In the context of peer and reference group effects, both downward and upward comparisons may be more feasible after the reforms, which might affect motivational aspects of students, such as their academic self-concept, and then in turn their educational aspirations (cf. Bittmann & Schindler, 2021; Chmielewski et al., 2013; Tenenbaum et al., 2020). Although the *Berlin study* showed a significant increase in the number of students who meet the formal requirements for transferring to the upper secondary school and a shift of the transition decision to the end of 10th grade (Neumann, Lühe, et al., 2017), no indications of motivational changes of students in the comprehensive schools associated with the *de-tracking* reforms were found (Jansen et al., 2017).

In addition, other school structural characteristics could also play a role that we cannot examine with our data. The history of the school location, i.e., whether it emerged from a former comprehensive school and lower tracks school types, for example, is discussed as a factor for the persistence of socioeconomic segregation (Helbig & Nikolai, 2017; Lenz et al., 2019; Lenz et al., 2022; Neumann, Lühe, et al., 2017). Also, not all school locations have an integrated upper school offering (i.e., from the 11th grade to the general university qualification) despite the two-path school structure, but instead cooperate with other school locations (cf. Neumann, Lühe, et al., 2017).

5.5.1 Limitations and Future Research

The main criticism of our study may be that we base our analyses on three time points, with one pre-reform and two post-reform period time points. As a result, we cannot observe the (common parallel) trends between the groups of states for a longer time before the reforms are introduced (cf. Angrist & Pischke, 2009). With at least two

measurement points before the reforms, it would be possible to test whether the trend in the outcome (i.e., students' educational aspiration) is parallel in the two groups. For example, the placebo test suggested by Lechner (2011) where the "treatment time" is moved to the time before the actual treatment can serve as plausibility check for the common trend assumption. As a placebo, a Difference-in-Differences analysis is conducted before the real treatment. In general, despite some robustness and plausibility checks, the Difference-in-Differences approach relies on the strong, untestable common trend assumption, which one can only counter with theoretical considerations (cf. Zeldow & Hatfield, 2019).

Further criticism could be related to the fact that we have to assign the states to different groups. Grouping states does not allow to look at differences between single states. Characteristics such as demographic composition or whether a state is a city or an area state are not balanced at group level. We address this limitation through additional checks of whether the patterns of our results differ across states within each group confirming no different pattern between single states. Moreover, we include student achievement, which is distributed differently between states and is the main determinant of students' educational aspirations, in additional analyses.

If other states also decide to move to a two-path school structure, scientific monitoring and evaluation is strongly recommended (such as the *Berlin study*). This could include, for example, a survey on cost-benefit calculations (cf. Breen & Goldthorpe, 1997; Wigfield & Eccles, 2000), which might allow us to understand the reasons for possible changes in students' educational aspirations or the lack of changes in socioeconomic disparities in students' educational aspirations.

5.6 Conclusion

The present study contributes to the ongoing discussion on how structural features of the educational system such as the extent of tracking practices affect socioeconomic disparities in educational outcomes (cf. Hanushek & Wößmann, 2006; Maaz et al.,

2008; Terrin & Triventi, 2023). Some German states reformed their tracking practices towards a two-path school structure where students can acquire the general university entrance qualification in the academic as well as in the non-academic comprehensive schools. Our Difference-in-Differences analyses show that students' educational aspirations have increased in the wake of the recent *de-tracking* reforms, which could contribute to raising educational attainment demanded by the labor market (OECD, 2021). Furthermore, our results indicate that the *de-tracking* reforms did not reach their goal to weaken the link between students' socioeconomic backgrounds and their educational aspirations. Secondary effects of students' socioeconomic backgrounds are still present and appear to have become stronger rather than weaker.

6

Conclusion

6.1 Integration of the Key Findings Into the State of Research

For decades, educational research and policy has been concerned that students from socioeconomically disadvantaged backgrounds are less successful in school than their peers from socioeconomically advantaged backgrounds (see, e.g., Kim et al., 2019; J. Liu et al., 2022). Despite educational expansion, socioeconomic disparities in educational outcomes are stable or have even tended to increase in some countries (Chmielewski, 2019). One feature of the education system that has been empirically identified and therefore repeatedly discussed as contributing to socioeconomic disparities in students' educational outcomes is a country's tracking practices (cf. 2.1.4; e.g., Hanushek & Wößmann, 2006; Terrin & Triventi, 2023; Woessmann, 2016). The debate on tracking practices and socioeconomic disparities has been enriched, especially in recent years, by numerous studies, mainly focusing on student achievement as outcome variable (see, e.g., Schindler et al., 2023; Sciffer et al., 2022; Skopek & Passaretta, 2021; Skopek et al., 2019; Strello et al., 2021; Strello et al., 2022; Terrin & Triventi, 2023; Triventi et al., 2020). These studies are usually based on comparisons between countries with different tracking practices, which might be

biased due to unobserved, potentially relevant differences between the countries (Piopiunik, 2021; Woessmann, 2016).

In comparisons between German federal states, it is assumed that the unobserved differences are smaller than in cross-national comparisons, also due to agreements between the federal states regarding some central characteristics of the school system (cf. KMK, 2020; Piopiunik, 2021). This dissertation takes advantage of the introduction of *de-tracking* reforms in some of the German federal states by comparing states with and without *de-tracking* reforms before and after their introduction (cf. 2.2.2 and 2.3; see also, e.g., Lenz, 2022; Piopiunik, 2021). In this way, the results can expand the state of research on the relationship between tracking practices, *de-tracking* reforms, and socioeconomic disparities. The few studies that compared the effects of reforms which postponed the start or reduced the intensity of between-school tracking practices within countries tend to show a reduction in socioeconomic disparities in student achievement (cf. 2.1.4; Kerr et al., 2013; Meghir & Palme, 2005; Piopiunik, 2014, 2021). This dissertation extends existing research by examining whether *de-tracking* reforms reduced disadvantages in terms of student achievement due to a less favorable SEC (*Study 1*) and in terms of socioeconomic disparities in students' educational aspirations (*Study 2*) using pre- and post-reforms analyses strategies.

Study 1 focuses on school-type specific changes in the SEC and in their association with student achievement. Specifically, it examined whether the SEC and the association of SEC with student achievement in the non-academic tracks primarily addressed by the *de-tracking* reforms have changed in the course of the reforms. Because the two-path system also offers students in the non-academic track the path to the general university entrance qualification, a possible increase in students' educational expectations might have contributed to the intended reduction in SEC effects. To answer these questions, representative data from the IQB National Assessment Study 2009 before the introduction of the *de-tracking* reforms and the IQB Trends in Student Achievement Study 2015 after the introduction of the *de-tracking* reforms were used (Köller et al., 2011; Stanat et al., 2018).

In comparison to the SEC at vocational track, intermediate track, and comprehensive

schools in 2009 – the schools that were merged to comprehensive schools – the SEC at comprehensive schools in 2015 was more favorable, as measured by an increase of the mean SES. The SEC in comprehensive schools after the *de-tracking* reforms was comparable to the SEC in intermediate track schools before the reforms and more favorable compared to the SEC at vocational tracks before the reforms. This implies that students from socioeconomically disadvantaged backgrounds, on average, attended a school with a more favorable SEC after the *de-tracking* reforms.

The relationship between SEC and student achievement was significant in both 2009 and 2015. The estimators of SEC effects in comprehensive schools did not change significantly between 2009 and 2015, although they tended to be smaller after the *de-tracking* reforms. This means that the SEC is still significantly associated with student achievement after the *de-tracking* reforms. In line with existing research (cf. Baumert et al., 2006b; Dumont, Neumann, Becker, et al., 2013; Tan et al., 2023; Van Ewijk & Sleegers, 2010), the results thus seem to confirm the existence of SEC effects, regardless of whether the non-academic system is characterized by between-school tracking practices or comprehensive schools.

When comparing vocational track schools before the *de-tracking* reforms and comprehensive schools after the implementation, the more favorable SEC and the empirical confirmation of SEC effects (cf. 2.1.6; Tan et al., 2023; Van Ewijk & Sleegers, 2010) suggest that students from socioeconomically disadvantaged families might have benefited from the *de-tracking* reforms. Research indicates that learning with higher-achieving classmates could facilitate learning progress of lower-achieving students (who, on average, come from socioeconomically disadvantaged backgrounds) (e.g., Mathewes, 2021; Tenenbaum et al., 2020). In addition, motivational gains are also conceivable if comprehensive schools were less stigmatized as a school type (cf. Knigge & Hannover, 2011). However, the results of *Study 1* did not confirm this assumption. The estimation of SEC effects at vocational track schools before the *de-tracking* reforms was associated with some difficulties (see Discussion of *Study 1*). Future studies should investigate this question further, especially because the unfavorable composition of some vocational track schools has

been criticized as detrimental to learning (Baumert et al., 2017; Baumert et al., 2006b; Tillmann & Wischer, 2006; Trautwein et al., 2007).

To further investigate if learning-promoting shared values among students might have played a role in the way SEC affected student achievement, *Study 1* analyzed if the association of SEC and student achievement was mediated by the level of educational expectations in the classroom by following the steps suggested by MacKinnon et al. (2012): After the level of students' educational expectations was included as a mediator, SEC was significantly associated with student achievement in both years and the SEC was related to the level of students' educational expectations. When both, SEC and the level of students' educational expectations were included, the level of educational expectations was not associated with student achievement in both years. Accordingly, the product terms of the indirect paths were also not significant, implying that the level of educational expectations did not play a role in the relationship between SEC and student achievement.

Interactions and mutual influences between classmates are theoretically considered as a mediating pathway between SEC and student achievement (cf. 2.2.4; Baumert et al., 2006a; Harker & Tymms, 2004; Palardy, 2013; Van Ewijk & Slegers, 2010), which was supported by a few studies (cf. Hornstra et al., 2015; H. Liu et al., 2015; Palardy, 2015). Recent studies on SEC and student achievement seem to focus more on the aspects of instructional quality than interactions and mutual influences between classmates, identifying school or classroom climate (cf. Agirdag, 2018; Holzberger et al., 2020; Mitchell et al., 2016; Tan et al., 2023) or instructional practices (cf. Holzberger et al., 2020; Muijs et al., 2014; Tan et al., 2023) as relevant factors for student learning. The scarcity of empirical findings on mediating interactions and mutual influences between classmates makes it difficult to integrate the results into an overall picture and to conclude whether mediation could not be confirmed because of the indirect operationalization of shared values among classmates (cf. Discussion of *Study 1*). In future studies, more attention should be paid to the mediation of SEC effects through interactions and mutual influences between classmates – also with regard to tracking practices to examine their relationship with SEC effects. In addition,

all assumed mediation processes should be comprehensively investigated to identify the most important influences (cf. Tan et al., 2023).

By focusing on students' educational aspirations, *Study 2* addressed a central educational outcome in addition to student achievement (see 2.1.3). *Study 2* examined whether the *de-tracking* reforms were associated with a change in students' educational aspirations and their relationship with students' socioeconomic backgrounds. With a Difference-in-Differences approach, representative large-scale assessment data from states with and without reforms from the pre- and post-reform periods were analyzed (Pant et al., 2015; Stanat et al., 2018; Stanat et al., 2022). Additionally, students' educational aspirations were regressed on their socioeconomic backgrounds in the states with and without reforms and the results from the pre- and post-reform periods were compared.

Between 2012 and 2015, when ninth-graders were affected by the *de-tracking* reforms, the proportion of students who planned to obtain the general university entrance qualification in the states with reforms increased, while it has not changed in the states without reforms. These findings for students of comprehensive schools are consistent with the observations of the *Berlin study* (Maaz et al., 2017; Neumann, Lühe, et al., 2017). In addition, students' educational aspirations in all three socioeconomic groups (divided into socioeconomically disadvantaged, middle or advantaged groups according to the highest EGP class of one of the parents (Ehmke & Baumert, 2008; Erikson et al., 1979; Kuhl et al., 2016)) increased significantly more in states with *de-tracking* reforms than in the states without reforms. However, socioeconomic disparities in students' educational aspirations did not seem to have been reduced at comprehensive schools; on the contrary, they appeared to have increased. The increase in the proportion of students who plan to obtain the general university entrance qualification was higher in the group of students from socioeconomically advantaged backgrounds than in the group of students from socioeconomically disadvantaged backgrounds. When controlling for achievement, the association between students' socioeconomic backgrounds and their educational aspirations remained strong, unaffected by the *de-tracking* reforms. The results

suggest that the *de-tracking* reforms have not effectively mitigated the association between students' socioeconomic backgrounds and their educational aspirations - on the contrary, it seemed to have strengthened it. These findings for the five states with *de-tracking* reforms, which also took into account student achievement, confirm and complement the observations of the *Berlin study* (Maaz et al., 2017; Neumann, Lühe, et al., 2017).

The described result patterns might be connected to the considerations regarding the cost-benefit calculations and the subjective value of school-leaving qualifications, respectively (cf. 2.1.3 and 2.2.5; Breen & Goldthorpe, 1997; Wigfield & Eccles, 2000). For example, Bittmann and Schindler (2021) focused on the educational aspirations of high-achieving students between the 5th and 9th grade and observed that students from socioeconomically disadvantaged backgrounds switched more often than students from socioeconomically advantaged backgrounds from aspiring the general university entrance qualification to aspiring a less ambitious school-leaving qualification. In addition, Triventi et al. (2020) found that inter-track mobility appears to increase rather than decrease socioeconomic disparities, as it occurs more frequently among students from socioeconomically advantaged rather than socioeconomically disadvantaged backgrounds. Taking this into account, students from socioeconomically advantaged backgrounds might be more likely to recognize the opportunity and the value of the higher school-leaving qualifications now being offered than students from socioeconomically disadvantaged backgrounds – despite showing comparable achievement. Since students' educational aspirations are an important predictor of decisions among possible alternatives in educational careers according to the theories of reasoned action and planned behavior (Ajzen, 1985; Madden et al., 1992; Stocké, 2013), these developments seem to counteract the intended goal of raising educational attainment and reducing socioeconomic disparities (cf. Baumert et al., 2013, 2017). These findings could also be used to develop measures that target socioeconomic disparities in students' educational aspirations (cf. Easterbrook & Hadden, 2021).

Overall, the results of *Study 1* and *2* suggest that *de-tracking* reforms from a multiple

school type system towards a two-path system have indeed led to changes, such as those in SEC or in the increase in students' educational aspirations. Despite these changes, no reduction in socioeconomic disparities, either regarding the SEC effects or regarding students' educational aspirations were observed. This confirms and supplements the findings of other studies on these *de-tracking* reforms and socioeconomic disparities (see also Helbig & Nikolai, 2017; Lenz et al., 2019; Lenz et al., 2022; Maaz et al., 2017; Maaz et al., 2019; Neumann, Lühe, et al., 2017). These results seem to differ from the findings of other studies on *de-tracking* reforms, which tend to show an increase in student achievement among students from socioeconomically disadvantaged background (Kerr et al., 2013; Meghir & Palme, 2005; Piopiunik, 2014, 2021). Other studies comparing different tracking practices suggest that between-school tracking practices tend to reinforce rather than mitigate socioeconomic disparities in student achievement and students' educational aspirations (e.g., Ammermueller, 2013; Buchmann & Dalton, 2002; Buchmann & Park, 2009; Hanushek & Wößmann, 2006; Lee, 2014; Parker et al., 2016; Schütz et al., 2008; Strello et al., 2021; Van de Werfhorst & Mijs, 2010; Van de Werfhorst & Hofstede, 2007). One aspect that might contribute to the differing patterns of results could be the fact that the *de-tracking* reforms examined in the studies of this dissertation lead to a two-path school structure, while other studies generally refer to comprehensive schooling (cf. 2.1.4 and 2.3; Skopek et al., 2019). Another aspect that could contribute to this is the way in which comprehensive schools are implemented, e.g., in terms of teaching quality and school characteristics (Maaz et al., 2019; Neumann, Becker, et al., 2017). For example, in the *Berlin study* and in the accompanying evaluation study in Bremen, school characteristics such as the school history before the *de-tracking* reforms, the school profile and the question of whether schools have an integrated upper secondary school or cooperate with other school locations instead have been found to be relevant for the choice of a school type and a school within the school type (cf. Helbig & Nikolai, 2017; Maaz et al., 2017; Maaz et al., 2019; Neumann, Lühe, et al., 2017; Zunker & Neumann, 2020; Zunker et al., 2018). Whether increasing the attractiveness of comprehensive schools can contribute to achieving the reform goals - combined with the question of how it can

be increased - should be investigated further.

For the overall interpretation of the results, the methodological evaluation of the data basis and the identification strategies of the studies addressed in the next section should be taken into account.

6.2 Limitations and Methodological Evaluation

6.2.1 Data Basis

Four national large-scale assessment studies, the IQB National Assessment Study 2009, the IQB National Assessment Study 2012, the IQB Trends in Student Achievement Study 2015, and the IQB Trends in Student Achievement Study 2018 are the data basis for the present studies, which comprise samples of ninth graders in all German states. One strength of the studies is their representativeness, allowing valid and precise population estimates at the federal state and school type level. To achieve unbiased estimates, the specific features of the complex item and unit sampling design were taken into account when analyzing this data. Sampling weights were used for parameter estimation, and plausible values were used to estimate measurement errors.

The aim of studies such as the one in this dissertation is to provide unbiased estimates of population parameters or of the relationships between student achievement, other variables such as motivational constructs, and students' socioeconomic backgrounds based on fully observed information. However, missing data are common and often unavoidable in large-scale assessment studies (Grund et al., 2021). They arise, for example, if students do not answer certain questions in the background questionnaire, if a rotating questionnaire (or planned missing data) is used or if students do not take part in the examination (Graham et al., 2006; Rhemtulla & Hancock, 2016). Often, information on students' socioeconomic backgrounds is unavailable for a significant proportion of participants, complicating the analysis of their relationship

with educational outcomes (Grund et al., 2021). Multiple imputation is one of the recommended approaches for dealing with missing information and follows the basic idea of finding plausible substitutions for missing values based on the observed data and a statistical model (Enders, 2022; Rubin, 1987). In *Study 1*, five (2009) and fifteen (2015) values estimated by multiple imputation for each student provided in the IQB assessment studies were used, if information on student-level background variables was missing (Böhme et al., 2010; Schipolowski et al., 2016). In *Study 2*, missing data was treated by estimating fifteen values through multilevel *multivariate imputation through chained equations* with fully conditional specification, and additionally included the aggregates of student-level variables as predictors in the multilevel imputation model (cf. Van Buuren & Groothuis-Oudshoorn, 2011). Accordingly, all results were pooled following Rubin's rules (Rubin, 1987). In summary, the methods used in the studies to deal with missing data should provide the most reliable estimates of the population parameters possible (Enders, 2022; Grund et al., 2021).

Secondary analyses come with the challenge that the researcher cannot influence which information was collected. While the achievement assessment and the survey of some key background characteristics were designed to measure changes between cycles, there were also constructs such as students' educational aspirations that have not been consistently collected across the four studies. For example, the IQB National Assessment Study 2009 only included a question on educational aspirations in the student participating list, but not in the student questionnaires. The student participation list is filled out by the administration of a school and is therefore presumably significantly shaped by the school track a student attends – and less by individual characteristics. To address this limitation in *Study 1*, the inter-rater reliability between educational aspirations from the two available sources of information (student questionnaire and student participation list) in the IQB Trends in Student Achievement Study 2015 was calculated and the analysis results were additionally compared with both measures. These additional analyses showed that the use of the information contained in the student participation list tended to underestimate the effect sizes of the estimators, so the conclusions drawn from the

results should be robust.

Students' general cognitive abilities are another construct that is particularly important for estimating SEC effects, but was not measured in 2009. They are commonly used as a proxy for prior achievement and knowledge, which are often not available in large-scale cross-sectional studies (Baumert et al., 2006a). In countries with between-school tracking practices such as Germany, accounting for students' prior achievement is particularly important to address an endogeneity problem (cf. 2.1.6; Baumert et al., 2006a; G. J. Duncan et al., 2004). Differences in cognitive abilities or achievement level prior to the introduction of the tracking practices could (partially) explain later achievement differences in the tracks. If prior achievement is not taken into account, the estimate of composition effects could be overestimated. The results on SEC effects in *Study 1* were additionally checked by including students' general cognitive ability in 2015. Taking students' general cognitive abilities into account, the SEC was not related to students' reading achievement in 2015, which contradicts previous findings for Germany (Baumert et al., 2006a; Dumont, Neumann, Maaz, et al., 2013). Conversely, employing measures of students' general cognitive abilities could result in underestimating composition effects (G. J. Duncan et al., 2004). This is because the development of students' general cognitive abilities is inherently shaped by both schooling and composition factors (M. Becker et al., 2012). The problem of concurrency or reciprocity is particularly evident in cross-sectional studies (Baumert et al., 2006a; G. J. Duncan et al., 2004). It arises from the fact that students are not merely passive recipients of the influence of their classroom environment, but actively shape it. This dynamic affects the validity of analyses on SEC effects in cross-sectional studies. The problem of concurrency and reciprocity cannot be solved with the current data, so future studies on SEC effects and tracking practices should be based on the analysis of longitudinal data to address this limitation.

Another potentially omitted variable in analyses of SEC effects is the ethnic composition of the student body (Mickelson et al., 2013). It can be measured, for example, by the percentage of ethnic minority students, i.e., of students with at least one parent born abroad (cf. Rjosk et al., 2017). This composition characteristic

should be investigated further against the background of *de-tracking* reforms, particularly because of the increasing proportion of students with immigrant backgrounds and observed changes in the vertical ethnic segregation in the course of the *de-tracking* reforms (Henschel et al., 2023; Lenz et al., 2021).

While central constructs are often covered, the surveys rarely include constructs that allow underlying mechanisms to be identified. For example, in *Study 1*, the presumed mediation process could only be mapped indirectly. The mediator, students' level of educational expectations as an indicator for shared values among students, was derived by aggregating individual educational expectations. Class-level constructs such as academic climate or mutual support might be more appropriate indicators of meaningful interactions and mutual influences between classmates. *Study 2* also had to leave unanswered the question of why socioeconomic disparities in students' educational aspirations have widened instead of narrowed as intended. Information on the benefits and costs of particular educational pathway evaluations and choices (in line with Bittmann & Schindler, 2021; Breen & Goldthorpe, 1997; Eccles & Roeser, 2011) among students would be informative, including for identifying levers for achieving a key goal of *de-tracking* reforms, the reduction of socioeconomic disparities in educational attainment.

Although the IQB assessment studies were not designed to evaluate the *de-tracking* reforms, their survey dates just before and just after the reforms provide a suitable data set. When interpreting the results, it is important to keep in mind that the cohort of students, particularly in 2015, was among the first to go through the restructured secondary school systems. Moreover, in the implementation of the *de-tracking* reforms at these early points in time, there were also differences between and within states (Lenz et al., 2022; Maaz et al., 2019; Neumann, Becker, et al., 2017). Therefore, in addition to more rigorous evaluation, a medium-term replication of the findings is particularly valuable. Whether the methods used in the studies can be used to draw conclusions about effects of *de-tracking* reforms is discussed in the next section.

6.2.2 Grouping States and Estimating Effects of *De-tracking* Reforms

In an optimal scenario from a research perspective, the states in which the *de-tracking* reforms were implemented would have been randomly selected as part of a randomized controlled experiment – regardless of states’ characteristics. Randomization reduces bias in the estimation of causal inferences. Empirical research is most valuable for evidence-based policy-making processes when the data and the methods used can reveal causal relationships (see Angrist & Pischke, 2009). The fact that the city-states are among the states with *de-tracking* reforms – and that, overall, they tend to be lower-performing states rather than higher-performing states in national large-scale assessment studies – suggests that there may be some similarities among the states that have introduced the *de-tracking* reforms. This “selection bias” may also be related to the driving factors for adapting the *de-tracking* reforms. This means that the data base does not meet the criteria for a randomized experiment. Since this is often the case, researchers are looking for quasi-experimental designs that allow them to use statistical techniques and identification strategies to approximate randomized experiments and thus the estimation of causal relationships (Angrist & Krueger, 1999).

Both studies take advantage of the similarities between the federal states based on agreements between the states on key educational policy issues such as comparability and equivalence in curricula and educational qualifications (cf. KMK, 2020). The common features of the *de-tracking* reforms, which primarily relate to the fact that they lead to a two-path system, are another strength of the analyses (cf. 2.2.2; Baumert et al., 2013, 2017; Hurrelmann, 2013; Lenz, 2022; Ridderbusch, 2019; Tillmann, 2012). This definition of the group of states with *de-tracking* reforms is narrower than the definition used by Piopiunik (2021), which includes all federal states that have either merged two existing school types, offered more school-leaving qualifications in a particular school type or offered an additional comprehensive school. According to his definition, in some federal states with *de-tracking* reforms

(such as Baden-Wuerttemberg), lower and intermediate track schools can continue to exist. Comparing the two definitions, the results based on the definition of the group of states with *de-tracking* reforms used in *Study 1* and *Study 2* appear to be more informative regarding the associations of non-academic comprehensive schooling compared to non-academic between-school tracking practices with the outcomes studied.

However, a differing structural feature between the states with *de-tracking* reforms was the extent of within-school tracking schools. In 2015, 45 % of the schools in Rhineland-Palatine were labelled as within-school tracking schools. For *Study 1*, we additionally calculated the SEC and SEC effects in within-school tracking schools and ran the analyses excluding Rhineland-Palatine. These additional checks did not contradict the results and conclusions of *Study 1*. *Study 2* referred to the information provided by the Federal Statistical Office in 2012 (Statistisches Bundesamt, 2012) due to an allocation error of within-school tracking schools in the IQB National Assessment Study 2012. The proportion of students attending within-school tracking schools was below 10 % in 2012, 2015, and 2018 (Rhineland-Palatine was not part of the group of states with *de-tracking* reforms in *Study 2* (cf. *Materials and Methods* in *Study 2*)). If within-school tracking schools affected the results of *Study 2* cannot be answered yet and should be covered in future research.

In both studies, the results from the states with *de-tracking* reforms were compared with the results from the states without reforms and tested for significant differences. This implies that the results are intended to offer insights into potential changes induced by the implementation of *de-tracking* reforms, as opposed to a scenario where these reforms had not been introduced. However, the analyses do not allow any statements regarding distinctions between a comprehensive secondary school system and a system with between-school tracking practices. As the identification strategy of *Study 1*, multilevel-multigroup analyses were conducted in the states with and without *de-tracking* reforms separately, with the assessment years representing the grouping variable. The differences between the SEC coefficients in 2009 and 2015 indicated if the changes were significant. The results of states with *de-tracking* reforms were

compared with those of states without reforms. Additionally, the difference in changes in SEC coefficients between 2009 and 2015 in the states with and without reforms was calculated controlling for the variables required for an appropriate estimation of SEC effects (see 2.1.6). This approach – the difference in changes between the “treatment”, i.e., states with *de-tracking* reforms, and “control” group, i.e., states without *de-tracking* reforms – came close to the basic idea of the Difference-in-Differences analyses used in *Study 2*.

Study 2 used the Difference-in-Differences approach to estimate the effects of *de-tracking* reforms. The Difference-in-Differences approach takes into account the differences that may have existed between the groups of states before the *de-tracking* reforms were introduced (cf. 2.3). Although it was not random which states implemented the *de-tracking* reforms, the underlying assumption is that the changes in the students’ educational aspirations for the group without reforms, the “control group”, provide information about how students’ educational aspirations in the states with reforms, the “treatment group”, would have changed if the reforms had not been introduced (cf. Figure 2.3; Angrist & Pischke, 2009; Athey & Imbens, 2017). This common trend assumption is a strong, non-testable assumption (cf. Kahn-Lang & Lang, 2020; Zeldow & Hatfield, 2019), based on holding confounding factors constant due to time-invariant, i.e., fixed, differences between the groups of states in unobserved variables that contribute to predicting the students’ educational aspirations (see, e.g., Angrist & Pischke, 2009). All time effects are assumed to be the same for the states with and the states without *de-tracking* reforms. Plausibility and robustness checks such as looking at the trends in students’ educational aspirations before the implementation of the *de-tracking* reforms, i.e., the pre-trend, or the placebo test (cf. *Discussion of Study 2*; Lechner, 2011) can help justify this assumption (see, e.g., Angrist & Pischke, 2009; Zeldow & Hatfield, 2019), but were not possible here as no suitable data were available. But even these checks cannot assure that students’ educational aspirations would have developed in the same manner in both groups of states if *de-tracking* reforms had not been introduced (cf. Kahn-Lang & Lang, 2020).

The synthetic control method pioneered by Abadie et al. (2010, 2015) and Abadie and Gardeazabal (2003) aims to address the challenges of the parallel trend assumption. Building on the basic idea of the Difference-in-Differences approach, this method is based on the concept of weighting the outcomes of the control unit (Abadie et al., 2010, 2015; Abadie & Gardeazabal, 2003). Specifically, a synthetic control unit is formed as a time-invariant weighted average of control units selected based on their similarity to the treated unit in terms of observed characteristics and outcome trends prior to the treatment. This way, it tries to create a composite that resembles the treated unit before the intervention. Challenges for researchers applying this method include reasonably choosing the pool of control units and calculating appropriate weights (Abadie, 2021; Athey & Imbens, 2017; Doudchenko & Imbens, 2016).

Using a control group based on a weighted selection of the states without *de-tracking* reforms as proposed by the synthetic control method might have increased the validity of the results (cf. Arkhangelsky et al., 2021). Even with a weighted selected comparison group, however, the analyses would be based on grouping the states with *de-tracking* reforms. By averaging the states with *de-tracking* reforms, differences between the states are not reflected. If the synthetic control method were applied, the characteristics and results of the states with *de-tracking* would be aggregated, and potentially contradictory developments in the individual states could not be revealed. Analyses on single state level, which could reveal opposing developments, are only permissible with a special data use agreement of the education administrators of all states. But even if analyses at individual state level had been permissible, differences between the states that are relevant for the investigation of tracking practices and socioeconomic disparities might still be unobserved, such as the extent of horizontal segregation or other forms of achievement-based within-school differentiation (Betts, 2011; Waldinger, 2007). These are discussed further in the next two sections.

6.3 Implications

6.3.1 For Educational Policy

One of the challenges in educational policy making and practice is to make evidence-based decisions that are supposed to be based on identifying the effects of reforms (Athey & Imbens, 2017). Clearly defined objectives and identification strategies for evaluating the measures contribute to enrich these decisions (Athey & Imbens, 2017). If possible (even if this is often unrealistic), the implementation of (*de-tracking*) reforms should be designed so that other school policy changes are not introduced at the same time to allow for the clearest possible evaluation of the reforms (see also 2.3 and 6.2.2; Koerselman, 2013; Terrin & Triventi, 2023). However, with the exception of Berlin and Bremen, no specific evaluation of the reforms was carried out. The studies of this dissertation have taken up the challenge of contributing to this public debate with the help of secondary analyses, even without a planned evaluation.

A central guiding principle of the German Federal Ministry of Education and Research is to reduce disparities in terms of students' (socioeconomic) backgrounds throughout their educational careers (Bundesministerium für Bildung und Forschung, 2023). The design of the education system has fundamental implications for the education and health of children and adolescents and also matters for the perpetuation of socioeconomic disparities across generations (cf. 2.1.4; see, e.g., Black & Devereux, 2011; Lochner, 2011). Moreover, education is an important building block for the higher-level ability to develop competencies and thus autonomy in self-selected areas (cf. Nussbaum & Sen, 1993). Also at the societal level, an increase in educational attainment seems to be associated with an increase in societal welfare (Hanushek & Woessmann, 2020). Therefore, attention should be paid to potential factors during the school career that may perpetuate or reinforce socioeconomic disparities, such as the tracking practices (cf. Schütz et al., 2008; Temkin, 2016).

While it may be unrealistic to completely avoid socioeconomic disparities, the goal is

to reduce them as much as possible (Bundesministerium für Bildung und Forschung, 2023). However, educational success is still closely linked to students' socioeconomic backgrounds (Chmielewski, 2019; Kuhl et al., 2016; Mahler & Kölm, 2018; Weis et al., 2020). The much-referenced German “PISA shock”, which followed the first PISA study (Baumert et al., 2001), brought to the fore discussions, demands and policy objectives for reducing socioeconomic disparities, including those related to transforming, i.e., *de-tracking*, the school structure of lower secondary education (M. Becker et al., 2017; Ridderbusch, 2019; Tillmann, 2012, 2016; Waldow, 2009). Educational policy makers may be concerned about a potential trade-off between universal or unequal support for student achievement and socioeconomic disparities in the context of tracking practices (see also 2.1.2 and 2.1.4; Pfeffer, 2015; Skopek et al., 2019; Terrin & Triventi, 2023). Given the debates on tracking practices, achievement levels, and socioeconomic disparities among educational policy makers, research covering these topics is of great practical importance.

The results of *Study 1* and *2* suggest that *de-tracking* reforms did not reduce socioeconomic disparities in terms of SEC effects and students' educational aspirations. In conjunction with other studies on these *de-tracking* reforms that have focused on different aspects of socioeconomic disparities (see also Helbig & Nikolai, 2017; Lenz et al., 2019; Lenz et al., 2022; Maaz et al., 2017; Maaz et al., 2019; Neumann, Lühe, et al., 2017), the findings therefore indicate that the *de-tracking* reforms did not achieve the goal of reducing socioeconomic disparities, at least in the short term. Medium and long-term follow-up studies are highly recommended to confirm these conclusions – and to derive the consequences for possible adjustments and additions as well as for potential future reform initiatives. Weakening the primacy of academic track schools or strengthening the position of the comprehensive schools could be relevant in order to further develop the two-path model (with the exception of Bremen where the acceptance of comprehensive schools is high; Maaz et al., 2019; cf. Lenz, 2022). Yet this two-path system was a compromise solution that did not extend comprehensive schooling for all students and left the academic track untouched (Hurrelmann, 2013; Neumann, Maaz, & Becker, 2013; Nikolai, 2019; Ridderbusch, 2019). Whether an extension of comprehensive schooling (for all

students or for other states) would be conceivable is questionable when looking at the results of a referendum in Hamburg in 2010. There, the referendum, presumably mainly influenced by parents from socioeconomically advantaged backgrounds, prevented the extension of the common elementary school period from four to six years, which would also have affected the academic track (Helbig & Edelstein, 2010; Töller et al., 2011). It seems quite challenging to increase actual support for educational policies to reduce socioeconomic disparities, even if providing information about the actual extent of educational disparities is associated with an increase in public awareness – concerns about declining student achievement and doubts about the effectiveness of policies to reduce socioeconomic disparities may matter (Lergetporer et al., 2020).

Based on these observations and challenges, it seems advisable for educational policy makers to further improve the way how comprehensive schools are implemented. Comprehensive schools only have a chance of being accepted as a second path alongside the academic track schools, as intended by the *de-tracking* reforms, if they offer equally good preparation, well-trained teachers and convincing teaching quality and do not show a decline in student achievement, especially with regard to the preparation for the general university entrance qualification (cf. Maaz et al., 2019; Neumann, Becker, et al., 2017; Roloff et al., 2020). The establishment of an upper secondary school or a stable and externally visible cooperation is seen a central development task for schools without an integrated upper secondary school (cf. Maaz et al., 2019; Neumann, Becker, et al., 2017). The range of school profiles on offer can also have an influence on the choice of the individual school (cf. Zunker & Neumann, 2020; Zunker et al., 2018).

Regarding further implications that educational policy makers might draw from the studies examining the effects of the *de-tracking* reforms, particular attention should be paid to other forms of achievement-based differentiation that might not be addressed by the *de-tracking* reforms (Schindler et al., 2023; Triventi et al., 2020). Horizontal segregation based on school choice mechanisms due to different school location factors and the aforementioned school characteristics might play a more

important role (Drope & Jurczok, 2013; Helbig & Nikolai, 2017; Lenz et al., 2022; Maaz et al., 2019; Neumann, Kropf, et al., 2013; Neumann & Maaz, 2019; Zunker & Neumann, 2020; Zunker et al., 2018). In addition, consideration should be given to how to counter potential information asymmetries between families from different socioeconomic backgrounds regarding the benefits, costs, and risks of particular educational pathways (cf. 2.2.5; e.g., Bittmann & Schindler, 2021). In Bremen, mandatory information talks with parents about the two different school types were introduced before parents submit their school preferences (Maaz et al., 2019). The question of whether such measures to inform parents are useful should be examined further.

6.3.2 For Future Research

Beyond the study-specific and methodological implications discussed in section 6.1 and section 6.2, some overarching implications will be addressed in this section. From a methodological perspective, the studies of this dissertation can serve as an example of the identification strategies that can be used for examining the effects of *de-tracking* reforms in secondary analyses, i.e., without a planned experimental setting and corresponding evaluation. Comparable educational standards, final exams, and degrees harmonize educational policies in Germany apart from tracking practices across states to a certain extent (KMK, 2020). This, in turn, is an advantage of the studies of this dissertation over cross-national comparative approaches (see also Piopiunik, 2021). A lower degree of unobserved heterogeneity is assumed for comparisons between federal states than for cross-national comparisons, which might lead to a lower bias due to omitted variables (cf. Woessmann, 2016).

Still, differences in institutional features such as the share of private schools or school autonomy should be taken into account wherever possible (Alegre & Ferrer, 2010; Klemm et al., 2018; Woessmann, 2016). The variety of analytical approaches used to analyze the consequences of tracking practices might contribute to a heterogeneity of results (Skopek et al., 2019). For example, Esser and Seuring (2020) referred to the

extent to which teacher recommendations for the transition from elementary to lower secondary school are binding as a basis for distinguishing between different tracking practices. This way of differentiating between different tracking practices might have different implications for the way socioeconomic disparities could be explained than described in section 2.1.4. The conclusions drawn from the results of *Study 1* and *2* relate only to tracking practices defined by the lower secondary school structure. In future studies, other characteristics identified as relevant to the relationship between tracking practices, socioeconomic disparities, and students' educational outcomes should be investigated in relation to the *de-tracking* reforms.

Another important question when examining tracking practices is which comparison target or counterfactual situation is being referred to (Terrin & Triventi, 2023). In the studies of the dissertation, the results of states with *de-tracking* reforms were compared to the other states without *de-tracking* reforms (see 2.3). The states with *de-tracking* reforms do not offer a completely comprehensive secondary school system because they left the academic track untouched. In both groups of states, the systems are still characterized by between-school tracking practices in which the socioeconomic segregation between academic track and non-academic school types seems to persist (Helbig & Nikolai, 2017; Lenz et al., 2019; Lenz et al., 2022; Maaz et al., 2019; Neumann, Becker, et al., 2017). This indicates that the two paths are not yet on an equal standing. *Study 1* and *2* focused on the non-academic school types. Thus, the results of comprehensive schools after the *de-tracking* reforms, which aim to consistently offer a second path to the general university entrance qualification, were compared to systems with tracking practices including school types not consistently offering the path to this school-leaving qualification. Within the non-academic track, comprehensive schooling is therefore compared with between-school tracking practices based on hierarchically ordered school types. Similar to other studies on associations of tracking practices, the studies do not refer to a completely comprehensive secondary school system as comparison target after *de-tracking* reforms (Heisig & Matthewes, 2022; Terrin & Triventi, 2023). The results therefore do not provide any conclusions about the comparison of secondary school systems with between-school tracking practices and systems with comprehensive schooling.

One question that remains unanswered is what the medium- and long-term consequences of *de-tracking* reforms will be. Between-school tracking practices have a long tradition, especially in the western German states, including the states where the *de-tracking* reforms were implemented. The majority of the parents of the student cohorts studied probably grew up in these school structures. The parents play a decisive role in crucial educational decisions such as the choice of school after elementary school (see 2.1.4; Bittmann & Schindler, 2021). The acceptance of comprehensive schools might increase over time, presumably influenced by the quality of instruction they offer (cf. 6.3.1). In addition to the short-term consequences examined, medium and long-term consequences should therefore also be investigated in future.

One issue that is closely related to the quality of instruction is how to best promote student learning in a heterogeneous student body. The way socioeconomic disparities were examined in this dissertation refers to one possible conception of equity, the “equality of outcome” (Dumont & Ready, 2023, p. 2). According to this concept, equity means that students with different backgrounds achieve the same outcomes, such as academic achievement, which is understood as a pathway to relevant outcomes later in life such as income and health and intergenerational mobility (Schütz et al., 2008; Temkin, 2016). In addition to between-school tracking practices, achievement-based within-school differentiation should be taken into account when evaluating studies on tracking practices, students’ educational outcomes, and socioeconomic disparities (Schindler et al., 2023; Triventi et al., 2020). For example, the concept of personalized learning is discussed as one form of achievement-based within-school differentiation (cf. Dumont & Ready, 2023). The basic idea of personalized learning is to adapt instruction according to students’ individual differences and needs (Dumont, 2019; Dumont & Ready, 2023). Personalized learning at its core aims to help everyone reach their full potential. Since prior knowledge, formal and informal learning opportunities, and students’ socioeconomic backgrounds tend to be related, socioeconomic disparities in student achievement could persist or even widen as a result of personalized learning (Dumont & Ready, 2023; Sokolowski & Ansari, 2018; Stern, 2017). In addition to within-school achievement-based differentiation in the form of personalized learning

or course-by-course tracking, other relevant factors, such as horizontal segregation or further characteristics of the schools (such as school history or profiling of the schools) should also be considered, where possible, when drawing conclusions of tracking practices related to student achievement and socioeconomic disparities (cf. Helbig & Nikolai, 2017; Schindler et al., 2023; Triventi et al., 2020; Zunker & Neumann, 2020; Zunker et al., 2018).

Bibliography

- Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2), 391–425.
- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American Statistical Association*, 105(490), 493–505.
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), 495–510.
- Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1), 113–132.
- Agirdag, O. (2018). The impact of school SES composition on science achievement and achievement growth: Mediating role of teachers’ teachability culture. *Educational Research and Evaluation*, 24(3-5), 264–276.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 11–39). Springer.
- Akay, E., & Karadag, E. (2019). Multilevel analyses of student, parent, and school indicators of achievement in high school transition in Turkey. *School Community Journal*, 29(2), 31–62.
- Alegre, M. À., & Ferrer, G. (2010). School regimes and education equity: Some insights based on PISA 2006. *British Educational Research Journal*, 36(3), 433–461.
- Alexander, K. L., & McDill, E. L. (1976). Selection and allocation within schools: Some causes and consequences of curriculum placement. *American Sociological Review*, 963–980.

- Ammermueller, A. (2013). Institutional features of schooling systems and educational inequality: Cross-country evidence from pirls and pisa. *German Economic Review*, 14(2), 190–213.
- Ammermüller, A. (2005). Educational opportunities and the role of institutions. *ZEW-Centre for European Economic Research Discussion Paper*, (05-044).
- Andrew, M., & Hauser, R. M. (2011). Adoption? Adaptation? Evaluating the formation of educational expectations. *Social Forces*, 90(2), 497–520. <https://doi.org/10.1093/sf/sor005>
- Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. In *Handbook of labor economics* (pp. 1277–1366). Elsevier.
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12), 4088–4118.
- Armor, D., Marks, G., & Malatinszky, A. (2018). The impact of school SES on student achievement: Evidence from U.S. statewide achievement data. *Educational Evaluation and Policy Analysis*, 40(4), 613–630. <https://doi.org/10.3102/0162373718787917>
- Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2), 3–32.
- Autorengruppe Bildungsberichterstattung. (2020). *Bildung in Deutschland 2020. Ein indikatorgestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt*. wbv. <https://doi.org/10.3278/6001820gw>
- Bandura, A. (1971). *Social learning theory*. General Learning Press.
- Barz, H. (2012). Der PISA-Schock. Über die Zukunft von Bildung und Wissenschaft im Land der “Kulturnation”. In G. Besier (Ed.), *20 Jahre neue Bundesrepublik: Kontinuitäten und Diskontinuitäten* (pp. 215–238). LIT.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (2001). *PISA 2000: Basiskompetenzen von Schüülerinnen und Schülern im internationalen Vergleich*. Leske + Budrich.

- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133–180. <https://doi.org/10.3102/0002831209345157>
- Baumert, J., Maaz, K., Neumann, M., Becker, M., & Dumont, H. (2013). Die Berliner Schulstrukturreform: Hintergründe, Zielstellungen und theoretischer Rahmen [Berlin school structure reform: Background, objectives, and theoretical framework]. In K. Maaz, J. Baumert, M. Neumann, M. Becker, & H. Dumont (Eds.), *Die Berliner Schulstrukturreform: Bewertung durch die beteiligten Akteure und Auswirkungen des neuen Übergangsverfahrens von der Grundschule in die weiterführende Schule* (pp. 9–34). Waxmann.
- Baumert, J., Maaz, K., Neumann, M., Becker, M., & Dumont, H. (2017). Die Berliner Schulstrukturreform: Hintergründe, Zielstellungen und theoretischer Rahmen [Berlin school structure reform: Background, objectives, and theoretical framework]. In M. Neumann, M. Becker, J. Baumert, K. Maaz, & O. Köller (Eds.), *Zweigliedrigkeit im deutschen Schulsystem. Potenziale und Herausforderungen in Berlin* (pp. 9–38). Waxmann.
- Baumert, J., & Schümer, G. (2001). Schulformen als selektionsbedingte Lernmilieus. In J. Baumert et al. (Eds.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 454–467). Leske + Budrich.
- Baumert, J., Stanat, P., & Watermann, R. (2006a). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus [School structure and the development of differentiated learning and development environments]. In J. Baumert, P. Stanat, & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (pp. 95–188). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90082-7_4
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J., & Weiß, M. (Eds.). (2002). *PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich*. Leske + Budrich.

- Baumert, J., Bos, W., & Watermann, R. (1999). *TIMSS III: Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Zusammenfassung deskriptiver Ergebnisse* (2nd, überarbeitete Auflage). Max-Planck-Institut für Bildungsforschung.
- Baumert, J., & Lehmann, R. H. (1997). *TIMSS — Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Leske + Budrich.
- Baumert, J., Stanat, P., & Watermann, R. (2006b). *Herkunftsbedingte Disparitäten im Bildungswesen. Vertiefende Analysen im Rahmen von PISA 2000*. Springer.
- Beal, S. J., & Crockett, L. J. (2010). Adolescents' occupational and educational aspirations and expectations: Links to high school activities and adult educational attainment. *Developmental Psychology, 46*(1), 258.
- Becker, B., Weirich, S., Mahler, N., & Sachse, K. A. (2019). Testdesign und Auswertung des IQB-Bildungstrends 2018: Technische Grundlagen [Test design and evaluation of the IQB Education Trend 2018: Technical Basics]. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2018. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich* (pp. 409–430). Waxmann.
- Becker, M., Kocaj, A., Jansen, M., Dumont, H., & Lüdtke, O. (2022). Class-average achievement and individual achievement development: Testing achievement composition and peer spillover effects using five German longitudinal studies. *Journal of Educational Psychology, 114*(1), 177.
- Becker, M., Lüdtke, O., Trautwein, U., & Baumert, J. (2006). Leistungszuwachs in Mathematik: Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem? *Zeitschrift für Pädagogische Psychologie, 20*(4), 233–242.
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology, 104*, 682–699. <https://doi.org/10.1037/a0027608>

- Becker, M., Neumann, M., & Dumont, H. (2017). Recent developments in school tracking practices in Germany: An overview and outlook on future trends. *Orbis Scholae*, 10(3), 9–25.
- Ben-Ari, R., & Kedem-Friedrich, P. (2000). Restructuring heterogeneous classes for cognitive development: Social interactive perspective. *Instructional Science*, 28(2), 153–167. <https://doi.org/10.1023/A:1003914904913>
- Benner, D., & Sladek, H. (1994). Das Gesetz zur Demokratisierung der deutschen Schule und die unterschiedliche Auslegung seiner harmonistischen Annahmen zum Verhältnis von Begabung und Bestimmung in den Jahren 1946/47. *Pädagogik und Erziehungsalltag in der DDR: Zwischen Systemvorgaben und Pluralität*, 37–54.
- Betts, J. R., Rueben, K. S., & Danenberg, A. (2000). *Equal resources, equal outcomes? The distribution of school resources and student achievement in California*. Public Policy Institute of California. http://www.ppic.org/content/pubs/report/r_200jbr.pdf
- Betts, J. R. (2011). The economics of tracking in education. In E. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 341–381). Elsevier.
- Bittmann, F. (2022). Investigating the role of educational aspirations as central mediators of secondary school track choice in Germany. *Research in Social Stratification and Mobility*, 81, 100715.
- Bittmann, F., & Schindler, S. (2021). Analysing diversion processes in German secondary education: School-track effects on educational aspirations. *Kölner Zeitschrift für Soziologie & Sozialpsychologie*, 73(2), 231–257.
- Black, S. E., & Devereux, P. J. (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics*, 4, 1487–1541.
- Bodovski, K. (2014). Adolescents' emerging habitus: The role of early parental expectations and practices. *British Journal of Sociology of Education*, 35(3), 389–412.
- Böhme, K., Leucht, M., Schipolowski, S., Porsch, R., Knigge, M., & Köller, O. (2010). Anlage und Durchführung des Ländervergleichs [Structure and Implementation

- of the IQB National Assessment Study 2009]. In O. Köller, M. Knigge, & B. Tesch (Eds.), *Sprachliche Kompetenzen im Ländervergleich* (pp. 65–85). Waxmann.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in western society*. Wiley.
- Bozick, R., Alexander, K., Entwisle, D., Dauber, S., & Kerr, K. (2010). Framing the future: Revisiting the place of educational expectations in status attainment. *Social forces*, 88(5), 2027–2052.
- Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and society*, 9(3), 275–305.
- Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annu. Rev. Sociol.*, 31, 223–243.
- Brunello, G., & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic policy*, 22(52), 782–861.
- Buchmann, C., & Dalton, B. (2002). Interpersonal influences and educational aspirations in 12 countries: The importance of institutional context. *Sociology of Education*, 99–122.
- Buchmann, C., & Park, H. (2009). Stratification and the formation of expectations in highly differentiated educational systems. *Research in Social Stratification and Mobility*, 27(4), 245–267.
- Bundesministerium für Bildung und Forschung. (2023). *Bekanntmachung: Richtlinie zur Förderung des Programms „Integration durch Bildung“*. Retrieved September 22, 2023, from <https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2023/04/2023-04-28-Bekanntmachung-Integration-Bildung.html>
- Card, D., & Giuliano, L. (2016). Can tracking raise the test scores of high-ability minority students? *American Economic Review*, 106, 2783–2816.
- Chmielewski, A. K. (2014). An international comparison of achievement inequality in within-and between-school tracking systems. *American Journal of Education*, 120(3), 293–324.
- Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *American Sociological Review*, 84(3), 517–544.

- Chmielewski, A. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal*, 50(5), 925–957.
- Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Government Printing Office.
- Contini, D., & Cugnata, F. (2020). Does early tracking affect learning inequalities? Revisiting difference-in-differences modeling strategies with international assessments. *Large-Scale Assessments in Education*, 8(1), 14. <https://doi.org/10.1186/s40536-020-00094-x>
- Dahrendorf, R. (1965). *Bildung ist Bürgerrecht: Plädoyer für eine aktive Bildungspolitik*. Nannen.
- Dämmrich, J., & Triventi, M. (2018). The dynamics of social inequalities in cognitive-related competencies along the early life course—a comparative study. *International Journal of Educational Research*, 88, 73–84.
- Davies, M., & Kandel, D. B. (1981). Parental and peer influences on adolescents' educational plans: Some further evidence. *American Journal of Sociology*, 87(2), 363–387.
- Dishion, T. J., & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology*, 62, 189–214. <https://doi.org/10.1146/annurev.psych.093008.100412>
- Ditton, H., & Krüsken, J. (2006). Der Übergang von der Grundschule in die Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 9(3), 348–372.
- Ditton, H. (2013). Kontexteffekte und Bildungsungleichheit: Mechanismen und Erklärungsmuster. In R. Becker & A. Schulze (Eds.), *Bildungskontexte* (pp. 173–206). Springer.
- Dockx, J., De Fraine, B., & Vandecandelaere, M. (2019). Does the track matter? A comparison of students' achievement in different tracks. *Journal of Educational Psychology*, 111(5), 827.

- Doudchenko, N., & Imbens, G. W. (2016). *Balancing, regression, difference-in-differences and synthetic control methods: A synthesis*. National Bureau of Economic Research (NBER Working Paper No. 22791).
- Drewek, P. (1994). *Die Entwicklung des Bildungssystems in den West-Zonen und in der Bundesrepublik von 1945/49 bis 1990. Strukturelle Kontinuität und Reformen, Bildungsexpansion und Systemprobleme*.
- Drope, T., & Jurczok, A. (2013). Weder gleichwertig noch gleichartig. Besonderheiten und Problemlagen Integrierter Sekundarschulen in einem sozio-ökonomisch schwachen Stadtteil Berlins. *Zeitschrift für Pädagogik*, 59(4), 496–507.
- Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *The American Economic Review*, 101(5), 1739–1774. <https://doi.org/10.1257/aer.101.5.1739>
- Dumay, X., & Dupriez, V. (2007). Accounting for class effect using the TIMSS 2003 eighth-grade database: Net effect of group composition, net effect of class process, and joint effect. *School Effectiveness and School Improvement*, 18(4), 383–408. <https://doi.org/10.1080/09243450601146371>
- Dumont, H., Neumann, M., Maaz, K., & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen [The composition of the student body as a factor influencing school performance]. *Psychologie in Erziehung Und Unterricht*, 60(3), 163–183. <https://doi.org/10.2378/peu2013.art14d>
- Dumont, H., Neumann, M., Nagy, G., Becker, M., Rose, N., & Trautwein, U. (2013). Einfluss der Klassenkomposition auf die Leistungsentwicklung in Haupt- und Realschulen in Baden-Württemberg [Influence of class composition on the achievement development in the vocational and intermediate track in Baden-Wuerttemberg]. *Psychologie in Erziehung Und Unterricht*, 60(3), 198–213. <https://doi.org/10.2378/peu2013.art16d>
- Dumont, H. (2019). Neuer Schlauch für alten Wein? Eine konzeptuelle Betrachtung von individueller Förderung im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 22(2), 249–277.

- 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. In *Die Berliner Schulstrukturreform: Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen* (pp. 133–207). Waxmann.
- Dumont, H., & Ready, D. D. (2023). On the promise of personalized learning for educational equity. *npj Science of Learning*, 8(1), 26.
- Duncan, G. J., Magnuson, K. A., & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, 1, 59–80. <https://doi.org/10.1080/15427609.2004.9683330>
- Duncan, G. J., & Raudenbush, S. W. (1999). Assessing the effects of context in studies of child and youth development. *Educational Psychologist*, 34, 29–41.
- Duncan, O., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20, 210–217.
- Dupriez, V., Monseur, C., Van Campenhout, M., & Lafontaine, D. (2012). Social inequalities of post-secondary educational aspirations: Influence of social background, school composition and institutional context. *European Educational Research Journal*, 11(4), 504–519. <https://doi.org/10.2304/eeerj.2012.11.4.504>
- Duru-Bellat, M., & Suchaut, B. (2005). Organisation and context, efficiency and equity of educational systems: What PISA tells us. *European Educational Research Journal*, 4(3), 181–194.
- Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers*, 56(2), 209–230.
- Easterbrook, M. J., & Hadden, I. R. (2021). Tackling educational inequalities with social psychology: Identities, contexts, and interventions. *Social Issues and Policy Review*, 15(1), 180–236.
- Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, 21, 225–241. <https://doi.org/10.1111/j.1532-7795.2010.00725.x>

- Ehmke, T., & Baumert, J. (2008). Soziale Disparitäten des Kompetenzerwerbs und der Bildungsbeteiligung in den Ländern: Vergleiche zwischen PISA 2000 und 2006 [Social disparities in competency acquisition and educational participation across countries: Comparisons between PISA 2000 and 2006]. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme, & R. Pekrun (Eds.), *PISA 2006 in Deutschland: Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (pp. 319–342). Waxmann.
- Enders, C. K. (2022). *Applied missing data analysis*. Guilford Publications.
- Epple, D., Newlon, E., & Romano, R. (2002). Ability tracking, school competition, and the distribution of educational benefits. *Journal of Public Economics*, 83(1), 1–48.
- Epstein, J. L. (2010). School/family/community partnerships: Caring for the children we share. *Phi delta kappan*, 76(9), 701–712. <https://doi.org/10.1177/0031721711009200326>
- Erbring, L., & Young, A. A. (1979). Individuals and social structure: Contextual effects as endogenous feedback. *Sociological Methods and Research*, 7(4), 396–430. <https://doi.org/10.1177/004912417900700404>
- Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *The British Journal of Sociology*, 30(4), 415–441.
- Esser, H. (2016). Bildungssysteme und ethnische Bildungsungleichheit. In C. Diehl, C. Hunkler, & C. Kristen (Eds.), *Ethnische Ungleichheiten im Bildungsverlauf* (pp. 331–396). Springer.
- Esser, H., & Seuring, J. (2020). Kognitive Homogenisierung, schulische Leistungen und soziale Bildungsungleichheit. *Zeitschrift für Soziologie*, 49(5-6), 277–301.
- Fend, H. (1982). *Gesamtschule im Vergleich: Bilanz der Ergebnisse des Gesamtschulversuchs*. Beltz.
- Figlio, D. N., & Page, M. E. (2002). School choice and the distributional effects of ability tracking: Does separation increase inequality? *Journal of Urban Economics*, 51, 497–514.

- Finger, C. (2016). Institutional constraints and the translation of college aspirations into intentions—evidence from a factorial survey. *Research in Social Stratification and Mobility, 46*, 112–128. <http://hdl.handle.net/10419/218969>
- Frost, M. B. (2007). Texas students' college expectations: Does high school racial composition matter? *Sociology of Education, 80*(1), 43–65. <https://doi.org/10.1177/003804070708000103>
- Ganzeboom, H. B. G., 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)
- Gerstman, B. B. (2013). *Epidemiology kept simple: An introduction to traditional and modern epidemiology*. John Wiley & Sons.
- Glaser, E. (2017). Pädagogik und Politik. Der Deutsche Ausschuss für das Erziehungs- und Bildungswesen und seine Empfehlungen als ein Beitrag zur Wissensgeschichte der Erziehungswissenschaft in der frühen Bundesrepublik. In S. Reh, E. Glaser, B. Behm, & T. Drope (Eds.), *Wissen machen: Beiträge zu einer Geschichte erziehungswissenschaftlichen Wissens in Deutschland zwischen 1945 und 1990* (pp. 88–107). Beltz Juventa.
- Goldsmith, P. R. (2011). Coleman revisited: School segregation, peers, and frog ponds. *American Educational Research Journal, 48*, 508–535. <https://doi.org/10.3102/0002831210392019>
- Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11*, 323–343. <https://doi.org/10.1037/1082-989X.11.4.323>
- Gresch, C. (2012). *Der Übergang in die Sekundarstufe I: Leistungsbeurteilung, Bildungsaspiration und rechtlicher Kontext bei Kindern mit Migrationshintergrund*. Springer VS.
- Grund, S., Lüdtke, O., & Robitzsch, A. (2021). On the treatment of missing data in background questionnaires in educational large-scale assessments: An evaluation of different procedures. *Journal of Educational and Behavioral Statistics, 46*(4), 430–465. <https://doi.org/10.3102/1076998620959058>

- Günther, K.-H. (1988). Profiles of educators: Wilhelm von Humboldt (1767–1835). *Prospects*, 18(1), 127–136.
- Guyon, N., Maurin, E., & McNally, S. (2012). The effect of tracking students by ability into different schools: A natural experiment. *Journal of Human Resources*, 47, 684–721.
- Haag, N., Kocaj, A., Jansen, M., & Kuhl, P. (2017). Soziale Disparitäten. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich, & N. Haag (Eds.), *IQB-Bildungstrend 2016: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (pp. 213–236). Waxmann.
- Haller, A. O. (1968). On the concept of aspiration. *Rural sociology*, 33(4), 484–487.
- Hallinan, M. T., & Williams, R. A. (1990). Students' characteristics and the peer-influence process. *Sociology of Education*, 63(2), 122–132. <https://doi.org/10.2307/2112858>
- Hamre, B. K., & Pianta, R. C. (2010). Classroom environments and developmental processes. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of Research on Schools, Schooling and Human Development*. Routledge. <https://doi.org/10.4324/9780203874844.ch3>
- Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5), 527–544. <https://doi.org/10.1002/jae.741>
- Hanushek, E. A., & Woessmann, L. (2020). A quantitative look at the economic impact of the european union's educational goals. *Education Economics*, 28(3), 225–244. <https://doi.org/10.1080/09645292.2020.1719980>
- Hanushek, E. A., & Wößmann, L. (2006). Does early tracking affect educational inequality and performance? Differences-in-differences evidence across countries. *Economic Journal*, 116(510), C63–C76. <https://doi.org/10.1111/j.1468-0297.2006.01076.x>
- Harker, R., & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15(2), 177–199. <https://doi.org/10.1076/sesi.15.2.177.30432>

- Hattie, J. A. C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, 37, 449–481. [https://doi.org/10.1016/S0883-0355\(03\)00015-6](https://doi.org/10.1016/S0883-0355(03)00015-6)
- Hauser, R. M. (1970). Context and consex: A cautionary tale. *American Journal of Sociology*, 75, 645–664.
- Heisig, J. P., & Matthewes, S. H. (2022). No evidence that strict educational tracking improves student performance through classroom homogeneity: A critical reanalysis of Esser and Seuring (2020). *Zeitschrift für Soziologie*, 51(1), 99–111.
- Helbig, M., & Edelstein, B. (2010). *Debatte Privatschule. Hamburg bleibt sitzen*. Retrieved September 21, 2023, from <https://taz.de/Debatte-Primarschule/!5138764/>
- Helbig, M., & Nikolai, R. (2015). *Die Unvergleichbaren: der Wandel der Schulsysteme in den deutschen Bundesländern seit 1949*. Julius Klinkhardt.
- Helbig, M., & Nikolai, R. (2017). *Alter Wolf im neuen Schafspelz? Die Persistenz sozialer Ungleichheiten im Berliner Schulsystem [Old wolf in new sheep's clothing? The persistence of social inequalities in Berlin's school system]*. WZB Discussion Paper.
- Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, 43, 59–74.
- Henschel, S., Heppt, B., & Weirich, S. (2023). Zuwanderungsbezogene Disparitäten [immigration-related disparities]. In P. Stanat, S. Schipolowski, R. Schneider, S. Weirich, S. Henschel, & K. A. Sachse (Eds.), *IQB-Bildungstrend 2022. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich* (pp. 299–344). Waxmann.
- Herrlitz, H.-G., Hopf, W., Titze, H., & Cloer, E. (2009). *Deutsche Schulgeschichte von 1800 bis zur Gegenwart: Eine Einführung* (5., aktualisierte Aufl.). Juventa.
- Hill, M. S., & Jenkins, S. P. (2001). Poverty among british children: Chronic or transitory? In B. Bradbury, S. P. Jenkins, & J. Micklewright (Eds.), *The Dynamics of Child Poverty in Industrialised Countries* (pp. 174–195). Cambridge University Press.

- Hindriks, J., Verschelde, M., Rayp, G., & Schoors, K. (2010). *School tracking, social segregation, and educational opportunity: Evidence from Belgium*. CORE Discussion Paper (81).
- Höffken, G. (2006). *Zur Institutionalisierung und Entwicklung der Mittelschule in Preußen 1872 bis 1945 unter besonderer Berücksichtigung des Chemieunterrichts* (Doctoral dissertation). Universität Potsdam.
- Holzberger, D., Reinhold, S., Lüdtke, O., & Seidel, T. (2020). A meta-analysis on the relationship between school characteristics and student outcomes in science and math – evidence from large-scale studies. *Studies in Science Education*, 56(1), 1–34. <https://doi.org/10.1080/03057267.2020.1735758>
- Hornstra, L., van der Veen, I., Peetsma, T., & Volman, M. (2015). Does classroom composition make a difference: Effects on developments in motivation, sense of classroom belonging, and achievement in upper primary school. *School Effectiveness and School Improvement*, 26(2), 125–152. <https://doi.org/10.1080/09243453.2014.887024>
- Hurrelmann, K. (2013). Das Schulsystem in Deutschland: Das „Zwei-Wege-Modell“ setzt sich durch. *Zeitschrift für Pädagogik*, 59(4), 455–468.
- Ireson, J., & Hallam, S. (2001). *Ability grouping in education*. Sage.
- Jackson, M. (2013). *Determined to succeed?: Performance versus choice in educational attainment*. Stanford University Press.
- Jansen, M., Stäbler, F., Becker, M., & Neumann, M. (2017). Motivationale Orientierungen und Aspekte schulischen Wohlbefindens vor und nach der Berliner Schulstrukturreform. In M. Neumann, M. Becker, J. Baumert, K. Maaz, & O. Köller (Eds.), *Zweigliedrigkeit im deutschen Schulsystem: Potenziale und Herausforderungen in Berlin* (pp. 295–320). Waxmann.
- Jenkins, S. P., Micklewright, J., & Schnepf, S. V. (2008). Social segregation in secondary schools: How does England compare with other countries? *Oxford Review of Education*, 34(1), 21–37.
- Jungmann, C. (2008). *Die Gemeinschaftsschule: Konzept und Erfolg eines neuen Schulmodells*. Waxmann.

- Kahn-Lang, A., & Lang, K. (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, 38(3), 613–620.
- Kelly, S. (2009). Social identity theories and educational engagement. *British Journal of Sociology of Education*, 30(4), 449–462. <https://doi.org/10.1080/01425690902954620>
- Kerckhoff, A. C. (1977). The realism of educational ambitions in England and the United States. *American Sociological Review*, 563–571.
- Kerr, S., Pekkarinen, T., & Uusitalo, R. (2013). School tracking and development of cognitive skills. *Journal of Labor Economics*, 31, 577–602.
- Kim, S. W., Cho, H., & Kim, L. Y. (2019). Socioeconomic status and academic outcomes in developing countries: A meta-analysis. *Review of Educational Research*, 89(6), 875–916.
- Klemm, K., Hoffmann, L., Maaz, K., & Stanat, P. (2018). *Privatschulen in Deutschland: Trends und Leistungsvergleiche*. Friedrich Ebert Stiftung.
- KMK. (2020). Ländervereinbarung über die gemeinsame Struktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen (Beschluss der Kultusministerkonferenz vom 15.10.2020).
- Knigge, M., & Hannover, B. (2011). Collective school-type identity: Predicting students' motivation beyond academic self-concept. *International Journal of Psychology*, 46(3), 191–205. <https://doi.org/10.1080/00207594.2010.529907>
- Knigge, M. (2009). *Hauptschüler als Bildungsverlierer?: Eine Studie zu Stigma und selbstbezogenem Wissen bei einer gesellschaftlichen Problemgruppe*. Waxmann Verlag.
- Koerselman, K. (2013). Incentives from curriculum tracking. *Economics of Education Review*, 32, 140–150.
- Köller, O., & Baumert, J. (2001). Leistungsgruppierungen in der Sekundarstufe I: Ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. *Zeitschrift für Pädagogische Psychologie*, 15(2), 99–110.

- Köller, O., Knigge, M., & Tesch, B. (2011). IQB-Ländervergleich Sprachen 2008/2009 (IQB-LV 2008-9) [IQB National Assessment Study 2008/2009] (Version 2) [Data set]. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. https://doi.org/http://doi.org/10.5159/IQB_LV_2008_v2
- Köller, O. (2005). Gesamtschule – Erweiterung statt Alternative. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer, & L. Trommer (Eds.), *Das Bildungswesen in der Bundesrepublik Deutschland: Strukturen und Entwicklungen im Überblick* (2nd, pp. 458–486). Rowohlt Taschenbuch Verlag.
- Kuhl, P., Haag, N., Federlein, F., Weirich, S., & Schipolowski, S. (2016). Soziale Disparitäten [Social disparities]. In P. Stanat, K. Böhme, S. Schipolowski, & N. Haag (Eds.), *IQB- Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (pp. 409–430). Waxmann.
- 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*, 805–820. <https://doi.org/10.1037/a0032583>
- Kuzmina, Y., & Ivanova, A. (2018). The effects of academic class composition on academic progress in elementary school for students with different levels of initial academic abilities. *Learning and Individual Differences, 64*, 43–53. <https://doi.org/10.1016/j.lindif.2018.04.004>
- Landtag Rheinland-Pfalz. (2008). Gesetzentwurf der Landesregierung. Landesgesetz zur änderung der Schulstruktur (Drucksache 15/2514).
- Lauen, D. L., & Gaddis, S. M. (2013). Exposure to classroom poverty and test score achievement: Contextual effects or selection? *American Journal of Sociology, 118*, 943–979. <https://doi.org/https://doi.org/10.1086/668408>
- Lavrijsen, J., & Nicaise, I. (2016). Educational tracking, inequality and performance: New evidence from a differences-in-differences technique. *Research in Comparative and International Education, 11*(3), 334–349.
- Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics, 4*(3), 165–224. <https://doi.org/10.1561/08000000014>

- Lee, B. (2014). The influence of school tracking systems on educational expectations: A comparative study of Austria and Italy. *Comparative Education*, 50(2), 206–228.
- Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule. *Kölner Zeitschrift für Soziologie & Sozialpsychologie*, 64(1).
- Lenz, S. M., Holtmann, M., Rjosk, C., & Stanat, P. (2019). Soziokulturelle Segregation an weiterführenden Schulen — Analysen zur Rolle der Gliederung des deutschen Schulsystems und schulstruktureller Reformmaßnahmen [Sociocultural segregation in Germany's secondary school systems — Effects of between-school tracking and de-tracking reforms]. *Zeitschrift für Erziehungswissenschaft*, 22(6), 1333–1358. <https://doi.org/10.1007/s11618-019-00913-0>
- Lenz, S. M., Rjosk, C., & Stanat, P. (2022). Schulische Segregation und ihre Veränderung im Zuge von Schulstrukturreformen in Berlin, Bremen und Hamburg. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 42(1), 54–72. <https://doi.org/10.3262/ZSE2201054>
- Lenz, S. M., Rjosk, C., Lorenz, G., & Stanat, P. (2021). Ethnische Segregation zwischen Schularten in mehrgliedrigen Schulsystemen und im “Zwei-Wege-Modell”. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73(1), 59–84.
- Lenz, S. M. (2022). *Schulstruktur und schulische Segregation: Allgemeine Zusammenhänge und Veränderungen mit der Umstellung auf zweigliedrige Schulsysteme in deutschen Bundesländern* (Doctoral dissertation). <https://refubium.fu-berlin.de/handle/fub188/35509>
- Lergetporer, P., Werner, K., & Woessmann, L. (2020). Educational inequality and public policy preferences: Evidence from representative survey experiments. *Journal of Public Economics*, 188. <https://doi.org/https://doi.org/10.1016/j.jpubeco.2020.104226>
- Liu, H., Van Damme, J., Gielen, S., & Van Den Noortgate, W. (2015). School processes mediate school compositional effects: Model specification and estimation. *British Educational Research Journal*, 41(3), 423–447. <https://doi.org/10.1002/berj.3147>

- Liu, J., Peng, P., Zhao, B., & Luo, L. (2022). Socioeconomic status and academic achievement in primary and secondary education: A meta-analytic review. *Educational Psychology Review*, 34(4), 2867–2896.
- Lochner, L. (2011). Non-production benefits of education: Crime, health, and good citizenship.
- Lorenz, G., Lenz, S. M., & Rjosk, C. (2023). Effizienz und soziale Ungleichheit in strikt leistungsdifferenzierenden Bildungssystemen. Eine kritische Betrachtung des Model of Ability Tracking (MoAbiT). *Zeitschrift für Soziologie*, 52(4), 404–424. <https://doi.org/10.1515/zfsoz-2023-2028>
- Lorenz, G., Boda, Z., Salikutluk, Z., & Jansen, M. (2020). Social influence or selection? Peer effects on the development of adolescents' educational expectations in Germany. *British Journal of Sociology of Education*, 41(5), 643–669.
- Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2 x 2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods*, 16, 444–467. <https://doi.org/10.1037/a0024376>
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13, 203–229. <https://doi.org/10.1037/a0024376>
- Lüdtke, O., Robitzsch, A., & Köller, O. (2002). Statistische Artefakte bei Kontexteffekten in der pädagogisch-psychologischen Forschung. *Zeitschrift für Pädagogische Psychologie*, 16, 217–231. <https://doi.org/10.1024//1010-0652.16.34.217>
- Maaz, K., Neumann, M., & Baumert, J. (Eds.). (2014). *Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter: Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive*. Springer.
- Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives*, 2(2), 99–106. <https://doi.org/10.1111/j.1750-8606.2008.00048.x>

- Maaz, K. (2006). *Soziale Herkunft und Hochschulzugang: Effekte institutioneller Öffnung im Bildungssystem*. VS Verlag für Sozialwissenschaften.
- Maaz, K., Bachsleitner, A., Baumert, J., Dumont, H., Neumann, M., Radmann, S., & Becker, M. (2017). Soziokulturelle Disparitäten des Bildungserfolgs vor und nach der Schulstrukturreform in Berlin [Socio-cultural disparities in educational achievement before and after the Berlin school structure reform]. In M. Neumann, M. Becker, J. Baumert, K. Maaz, & O. Köller (Eds.), *Zweigliedrigkeit im deutschen Schulsystem: Potenziale und Herausforderungen in Berlin* (pp. 361–424). Waxmann Verlag.
- Maaz, K., Hasselhorn, M., Idel, T.-S., Klieme, E., Lütje-Klose, B., Stanat, P., Neumann, M., Bachsleitner, A., Lühe, J., & Schipolowski, S. (Eds.). (2019). *Zweigliedrigkeit und Inklusion im empirischen Fokus. Ergebnisse der Evaluation der Bremer Schulreform*. Waxmann.
- MacKinnon, D. P., Cheong, J., & Pirlott, A. G. (2012). Statistical mediation analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA Handbook of Research Methods in Psychology, Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological* (pp. 313–331). American Psychological Association. <https://doi.org/10.1037/13620-018>
- Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin*, 18(1), 3–9.
- Mahler, N., & Kölm, J. (2018). Soziale Disparitäten [Social disparities]. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2018. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich* (pp. 265–291). Waxmann.
- Marks, G. N. (2015). Are school-SES effects statistical artefacts? Evidence from longitudinal population data. *Oxford Review of Education*, 41(1), 122–144. <https://doi.org/10.1080/03054985.2015.1006613>
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual

- and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, 47, 106–124. <https://doi.org/10.1080/00461520.2012.670488>
- Matthewes, S. H. (2021). Better together? Heterogeneous effects of tracking on student achievement. *The Economic Journal*, 131(635), 1269–1307.
- McDaniel, A. (2010). Cross-national gender gaps in educational expectations: The influence of national-level gender ideology and educational systems. *Comparative Education Review*, 54(1), 27–50.
- Meghir, C., & Palme, M. (2005). Educational reform, ability, and family background. *American Economic Review*, 95(1), 414–424.
- Meyer, J. W. (1970). High school effects on college intentions. *American Journal of Sociology*, 76(1), 59–70. <https://doi.org/10.1086/224906>
- Mickelson, R. A., Bottia, M. C., & Lambert, R. (2013). Effects of school racial composition on K-12 mathematics outcomes: A metaregression analysis. *Review of Educational Research*, 83, 121–158. <https://doi.org/10.3102/0034654312475322>
- Mitchell, R., Mendiola, B., Schumacker, R., & Lowery, X. (2016). Creating a school context of success. *Journal of Educational Administration*, 54(6), 626–646. <https://doi.org/10.1108/JEA-02-2015-0018>
- Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art – teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>
- Müller, K., & Ehmke, T. (2016). Soziale Herkunft und Kompetenzerwerb. In *PISA 2015: Eine Studie zwischen Kontinuität und Innovation* (pp. 285–316). Waxmann.
- Müller-Benedict, V. (2007). Wodurch kann die soziale Ungleichheit des Schulerfolgs am stärksten verringert werden? *Kölner Zeitschrift für Soziologie & Sozialpsychologie*, 59(4), 615–639.
- Muthén, L. K., & Muthén, B. (1998 – 2007). *Mplus User's Guide. Eighth Edition*. Muthén & Muthén.
- 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. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 41(4), 449–466.
- Neumann, M., Becker, M., Baumert, J., Maaz, K., & Köller, O. (2017). *Zweigliedrigkeit im deutschen Schulsystem: Potenziale und Herausforderungen in Berlin*. Waxmann.
- Neumann, M., Kropf, M., Becker, M., Albrecht, R., Maaz, K., & Baumert, J. (2013). Die Wahl der weiterführenden Schule im neu geordneten Berliner Übergangsverfahren. In *Die Berliner Schulstrukturreform: Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen* (pp. 87–132). Waxmann.
- Neumann, M., & Maaz, K. (2019). Nachfrageunterschiede zwischen weiterführenden Schulen. Eine Untersuchung auf Basis schulstatistischer Daten im Land Bremen. In D. Fickermann & H. Weishaupt (Eds.), *Bildungsforschung mit Daten der amtlichen Statistik* (pp. 197–214). Waxmann.
- Neumann, M., Lühe, J., Becker, M., & Baumert, J. (2017). Schulbiografische Verläufe und Abschlussaspirationen vor und nach der Berliner Schulstrukturreform [School biographies and graduation aspirations before and after the Berlin school structure reform]. In M. Neumann, M. Becker, J. Baumert, K. Maaz, & O. Köller (Eds.), *Zweigliedrigkeit im deutschen Schulsystem: Potenziale und Herausforderungen in Berlin* (pp. 127–154). Waxmann Verlag.
- Neumann, M., Maaz, K., & Becker, M. (2013). Die Abkehr von der traditionellen Dreigliedrigkeit im Sekundarschulsystem: Auf unterschiedlichen Wegen zum gleichen Ziel? *Recht der Jugend und des Bildungswesens*, 61(3), 274–292.
- Niemietz, J., Jindra, C., Schneider, R., Schumann, K., S., S., & Sachse, K. A. (2023). Soziale Disparitäten [Social disparities]. In P. Stanat, S. Schipolowski, R. Schneider, S. Weirich, S. Henschel, & K. A. Sachse (Eds.), *IQB-Bildungstrend 2022. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich* (pp. 261–298). Waxmann.
- Nikolai, R. (2019). Zweigliedriges Schulsystem. In M. Haring, C. Rohlf, & M. Gläser-Zikuda (Eds.), *Handbuch Schulpädagogik* (pp. 94–102). Waxmann.

- Nussbaum, M., & Sen, A. (1993). *The quality of life*. Clarendon Press.
- OECD. (2010). *Education at a glance 2010: OECD indicators*.
- OECD. (2017). *The funding of school education: Connecting resources and learning*.
<https://doi.org/10.1787/9789264276147-en>
- OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed*. <https://doi.org/10.1787/b5fd1b8f-en>
- OECD. (2021). *Education policy outlook 2021: Shaping responsive and resilient education in a changing world*.
- Opdenakker, M.-C., & Van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal*, 27(4), 407–432. <https://doi.org/10.1080/01411920125573>
- Opdenakker, M.-C., & Van Damme, J. (2006). Differences between secondary schools: A study about school context, group composition, school practice, and school effects with special attention to public and catholic schools and types of schools. *School Effectiveness and School Improvement*, 17(1), 87–117. <https://doi.org/10.1080/09243450500264457>
- Palardy, G. J. (2013). High school socioeconomic segregation and student attainment. *American Educational Research Journal*, 50(4), 714–754. <https://doi.org/10.3102/0002831213481240>
- Palardy, G. J. (2015). High school socioeconomic composition and college choice: Multilevel mediation via organizational habitus, school practices, peer and staff attitudes. *School Effectiveness and School Improvement*, 26(3), 329–353. <https://doi.org/10.1080/09243453.2014.965182>
- Pant, H. A., Stanat, P., Hecht, M., Heitmann, P., Jansen, M., Lenski, A. E., Penk, C., Pöhlmann, C., Roppelt, A., Schroeders, U., & Siegle, T. (2015). *IQB-Ländervergleich Mathematik und Naturwissenschaften 2012 (IQB-LV 2012) [IQB National Assessment Study 2012 (IQB-LV 2012)] (Version 4) [Data set]*. IQB – Institut zur Qualitätsentwicklung im Bildungswesen.
- Parker, P. D., Jerrim, J., Schoon, I., & Marsh, H. W. (2016). A multinational study of socioeconomic inequality in expectations for progression to higher education:

- The role of between-school tracking and ability stratification. *American Educational Research Journal*, 53(1), 6–32.
- Pfeffer, F. T. (2015). Equality and quality in education. a comparative study of 19 countries. *Social science research*, 51, 350–368.
- Pfost, M., Karing, C., Lorenz, C., & Artelt, C. (2010). Schereneffekte im ein- und mehrgliedrigen Schulsystem. *Zeitschrift für Pädagogische Psychologie*.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109–119. <https://doi.org/10.3102/0013189X09332374>
- Picht, G. (1964). *Die deutsche Bildungskatastrophe: Analyse und Dokumentation*. Walter-Verlag.
- Piopiunik, M. (2014). The effects of early tracking on student performance: Evidence from a school reform in Bavaria. *Economics of Education Review*, 42, 12–33. <https://doi.org/10.1016/j.econedurev.2014.06.002>
- Piopiunik, M. (2021). *How does reducing the intensity of tracking affect student achievement? Evidence from German state reforms* (CESifo Working Paper No. 9214). Center for Economic Studies and ifo Institute (CESifo). Munich.
- Pischke, J.-S., & Manning, A. (2006). Comprehensive versus selective schooling in england in wales: What do we know?
- Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear “difference-in-differences” models. *Economics Letters*, 115(1), 85–87.
- R Core Team. (2022). *R: A language and environment for statistical computing* [R version 4.2.1]. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Rauch, D., Mang, J., Härtig, H., & Haag, N. (2016). Naturwissenschaftliche Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In *PISA 2015: Eine Studie zwischen Kontinuität und Innovation* (pp. 317–348). Waxmann.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd). Sage.

- Retelsdorf, J., & Möller, J. (2008). Entwicklungen von Lesekompetenz und Lesemotivation. Schereneffekte in der Sekundarstufe? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *40*(4), 179–188.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (eer): A state-of-the-art review. *School Effectiveness and School Improvement*, *25*(2), 197–230. <https://doi.org/10.1080/09243453.2014.885450>
- Rhemtulla, M., & Hancock, G. R. (2016). Planned missing data designs in educational psychology research. *Educational Psychologist*, *51*, 305–316. <https://doi.org/10.1080/00461520.2016.1208094>
- Ridderbusch, J. (2019). *Deutschland auf dem Weg zum zweigliedrigem Schulsystem: Transfer- und Lernprozesse in der Bildungspolitik*. Springer.
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction*, *32*, 63–72. <https://doi.org/10.1016/j.learninstruc.2014.01.007>
- Rjosk, C., Richter, D., Lüdtke, O., & Eccles, J. S. (2017). Ethnic composition and heterogeneity in the classroom: Their measurement and relationship with student outcomes. *Journal of Educational Psychology*, *109*(8), 1188.
- Roemer, J. E., & Trannoy, A. (2016). Equality of opportunity: Theory and measurement. *Journal of Economic Literature*, *54*(4), 1288–1332.
- Roloff, J., Klusmann, U., Lüdtke, O., & Trautwein, U. (2020). The predictive validity of teachers' personality, cognitive and academic abilities at the end of high school on instructional quality in Germany: A longitudinal study. *Aera Open*, *6*(1), 2332858419897884.
- Rubin, B. D. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.
- Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, *107*(9), 1999–2045. <https://doi.org/10.1111/j.1467-9620.2005.00583.x>

- Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 249–277). Elsevier. <https://doi.org/10.1016/B978-0-444-53429-3.00004-1>
- Scharenberg, K., Rollett, W., & Bos, W. (2014). Zur Domänenspezifität differenzieller Lern- und Entwicklungsmilieus. In H. G. Holtappels, A. S. Willems, M. Pfeifer, W. Bos, & N. McElvany (Eds.), *Jahrbuch der Schulentwicklung. Band 18. Daten, Beispiele und Perspektiven* (pp. 239–269). Beltz Juventa.
- Schindler, S., Bar-Haim, E., Barone, C., Fels Birkelund, J., Boliver, V., Capsada-Munsech, Q., Erola, J., Facchini, M., Feniger, Y., Heiskala, L., et al. (2023). Educational tracking and social inequalities in long-term labor market outcomes: Six countries in comparison. *International Journal of Comparative Sociology*. <https://doi.org/https://doi.org/10.1177/00207152231151390>
- Schipolowski, S., Haag, N., & Böhme, K. (2016). Anlage und Durchführung [structure and implementation]. In P. Stanat, K. Böhme, S. Schipolowski, & N. Haag (Eds.), *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (pp. 95–119). Waxmann.
- Schlotter, M., Schwerdt, G., & Woessmann, L. (2011). Econometric methods for causal evaluation of education policies and practices: A non-technical guide. *Education Economics*, 19(2), 109–137.
- Schütz, G., Ursprung, H. W., & Wößmann, L. (2008). Education policy and equality of opportunity. *Kyklos*, 61(2), 279–308.
- Sciffer, M. G., Perry, L. B., & McConney, A. (2022). Does school socioeconomic composition matter more in some countries than others, and if so, why? *Comparative Education*, 58(1), 37–51.
- Sewell, W. H., Haller, A. O., & Ohlendorf, G. W. (1970). The educational and early occupational status attainment process: Replication and revision. *American Sociological Review*, 1014–1027.
- Sewell, W. H., Haller, A. O., & Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review*, 82–92.

- Sewell, W. H., Hauser, R. M., Springer, K. W., & Hauser, T. S. (2003). As we age: A review of the Wisconsin Longitudinal Study, 1957–2001. *Research in Social Stratification and Mobility*, 20, 3–111.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Skopek, J., & Passaretta, G. (2021). Socioeconomic inequality in children's achievement from infancy to adolescence: The case of Germany. *Social Forces*, 100(1), 86–112.
- Skopek, J., Triventi, M., & Buchholz, S. (2019). How do educational systems affect social inequality of educational opportunities? The role of tracking in comparative perspective. In R. Becker (Ed.), *Research Handbook on the Sociology of Education* (pp. 214–232). <https://doi.org/10.4337/9781788110426.00022>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.
- Sokolowski, H. M., & Ansari, D. (2018). Understanding the effects of education through the lens of biology. *npj Science of Learning*, 3(1), 17.
- Spenner, K. I., & Featherman, D. L. (1978). Achievement ambitions. *Annual Review of Sociology*, 373–420.
- Stanat, P. (2006). Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft. In J. Baumert, P. Stanat, & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (pp. 189–219). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90082-7_5
- Stanat, P., Böhme, K., Schipolowski, S., Haag, N., Weirich, S., Sachse, K. A., Hoffmann, L., & Federlein, F. (2018). *IQB-Bildungstrend Sprachen 2015 (IQB-BT 2015) [IQB Trends in Student Achievement 2015 (IQB-BT 2015)] (Version 5) [Data set]*. IQB – Institut zur Qualitätsentwicklung im Bildungswesen.

- Stanat, P., Schipolowski, S., Mahler, N., Weirich, S., Henschel, S., Holtmann, M., Becker, B., & Kölm, J. (2022). *IQB-Bildungstrend Mathematik und Naturwissenschaften 2018 (IQB-BT 2018) [IQB Trends in Student Achievement 2018 (IQB-BT 2018)] (Version 1) [Data set]*. IQB – Institut zur Qualitätsentwicklung im Bildungswesen.
- Statistisches Bundesamt. (2012). *Fachserie 11: Bildung und Kultur, Reihe 1: Allgemeinbildende Schulen, Schuljahr 2011/12 [Reference Series 11: Education and Culture, Series 1: General Education Schools, School Year 2011/12]*. Retrieved January 27, 2023, from https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00010180/2110100127004.pdf
- Stern, E. (2017). Individual differences in the learning potential of human beings. *Npj Science of Learning*, 2(1), 2. <https://doi.org/10.1038/s41539-017-0002-2>
- Stocké, V. (2013). Bildungsaspirationen, soziale Netzwerke und Rationalität. In R. Becker & A. Schulze (Eds.), *Bildungskontexte* (pp. 269–298). Springer.
- Stocké, V., Blossfeld, H.-P., Hoenig, K., & Sixt, M. (2011). Social inequality and educational decisions in the life course. *Zeitschrift für Erziehungswissenschaft*, 14(2), 103–119.
- Strello, A., Strietholt, R., Steinmann, I., & Siepman, C. (2021). Early tracking and different types of inequalities in achievement: Difference-in-differences evidence from 20 years of large-scale assessments. *Educational Assessment, Evaluation and Accountability*, 33, 139–167.
- Strello, A., Strietholt, R., & Steinmann, I. (2022). Does tracking increase segregation? International evidence on the effects of between-school tracking on social segregation across schools. *Research in Social Stratification and Mobility*, 78, 100689.
- Stubbe, T. C., Kasper, D., & Jentsch, A. (2020). Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen. In K. e. a. Schwippert (Ed.), *TIMSS 2019: Mathematische und naturwissenschaftliche Kompetenzen von*

- Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 315–330). Waxmann.
- Stubbe, T. C., Krieg, M., Beese, C., & Jusufi, D. (2020). Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässlerinnen und Viertklässlern. In K. Schwippert et al. (Eds.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 263–290). Waxmann.
- Stubbe, T. C., Lorenz, J., Bos, W., & Kasper, D. (2016). Kapitel XIII. Der Übergang von der Primar- in die Sekundarstufe. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert, & D. Kasper (Eds.), *TIMSS 2015: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 351–366). Waxmann.
- Tan, C. Y. (2018). Socioeconomic status, involvement practices, and student science achievement: Insights from a typology of home and school involvement patterns. *American Educational Research Journal*, 56(3), 899–924. <https://doi.org/10.3102/0002831218807146>
- Tan, C. Y. (2020). *Family cultural capital and student achievement: Theoretical insights from PISA*. Springer.
- Tan, C. Y., Hong, X., Gao, L., & Song, Q. (2023). Meta-analytical insights on school SES effects. *Educational Review*. <https://doi.org/10.1080/00131911.2023.2184329>
- Temkin, L. S. (2016). The many faces of equal opportunity. *Theory and Research in Education*, 14(3), 255–276.
- Tenenbaum, H. R., Winstone, N. E., Leman, P. J., & Avery, R. E. (2020). How effective is peer interaction in facilitating learning? A meta-analysis. *Journal of Educational Psychology*, 112(7), 1303–1319. <https://doi.org/10.1037/edu0000436>
- Terhart, E. (2001). Bildungsforschung, Bildungsadministration, Bildungswirklichkeit: eine systematische Annäherung. *Politikberatung durch Bildungsforschung: Das Beispiel: Schulentwicklung in Hamburg*, 17–32.

- Terrin, É., & Triventi, M. (2023). The effect of school tracking on student achievement and inequality: A meta-analysis. *Review of Educational Research*, 93(2), 236–274.
- Thrupp, M., Lauder, H., & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research*, 37, 483–504. [https://doi.org/10.1016/S0883-0355\(03\)00016-8](https://doi.org/10.1016/S0883-0355(03)00016-8)
- Tillmann, K.-J. (2012). Das Sekundarschulsystem auf dem Weg in die Zweigliedrigkeit: Historische Linien und aktuelle Verwirrungen. *Pädagogik*, 64(5), 8–12.
- Tillmann, K.-J. (2016). Neue Argumente auf einem alten Kampffeld? Die Schulstruktur-Diskussion in kritisch-erziehungswissenschaftlicher Sicht. In A. Schippling, C. Grunert, & N. Pfaff (Eds.), *Kritische Bildungsforschung: Standortbestimmungen und Gegenstandsfelder* (pp. 211–226). Leske + Budrich.
- Tillmann, K.-J., & Wischer, B. (2006). Heterogenität in der Schule: Forschungsstand und Konsequenzen. *Pädagogik*, 58(3), 44–48.
- Töller, A. E., Pannowitsch, S., Kuscheck, C., & Mennrich, C. (2011). Direkte Demokratie und Schulpolitik. Lehren aus einer politikfeldanalytischen Betrachtung des Scheiterns der Hamburger Schulreform. *Zeitschrift für Parlamentsfragen*, 42(3), 503–523.
- Traini, C., Kleinert, C., & Bittmann, F. (2021). How does exposure to a different school track influence learning progress? Explaining scissor effects by track in Germany. *Research in Social Stratification and Mobility*, 76, 100625.
- Trautwein, U., Baumert, J., & Maaz, K. (2007). Hauptschulen = Problemschulen? *Aus Politik und Zeitgeschichte*, 57(28), 3–9.
- Triventi, M., Skopek, J., Kulic, N., Buchholz, S., & Blossfeld, H.-P. (2020). Advantage ‘finds its way’: How privileged families exploit opportunities in different systems of secondary education. *Sociology*, 54(2), 237–257.
- Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45, 1–67.
- Van de Werfhorst, H. G., & Mijs, J. J. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36, 407–428.

- Van de Werfhorst, H. G. (2018). Early tracking and socioeconomic inequality in academic achievement: Studying reforms in nine countries. *Research in Social Stratification and Mobility*, 58, 22–32.
- Van de Werfhorst, H. G. (2019). Early tracking and social inequality in educational attainment: Educational reforms in 21 European countries. *American Journal of Education*, 126(1), 65–99.
- Van de Werfhorst, H. G., & Hofstede, S. (2007). Cultural capital or relative risk aversion? Two mechanisms for educational inequality compared. *The British Journal of Sociology*, 58(3), 391–415.
- Van Ewijk, R., & Slegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5(2), 134–150. <https://doi.org/10.1016/j.edurev.2010.02.001>
- Waldinger, F. (2007). *Does tracking affect the importance of family background on students' test scores?* [Working Paper, London School of Economics].
- Waldow, F. (2009). What PISA did and did not do: Germany after the “PISA shock”. *European Educational Research Journal*, 8(3), 476–483.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.
- Weirich, S., Hecht, M., & Becker, B. (2022). *eatRep: Educational assessment tools for replication methods* [R package version 0.14.5]. <https://github.com/weirichs/eatRep>
- Weis, M., Müller, K., Mang, J., Heine, J.-H., Mahler, N., & Reiss, K. (2020). Soziale Herkunft, Zuwanderungshintergrund und Lesekompetenz. In *PISA 2018. Ergebnisse der aktuellen Erhebungsrunde mit dem Schwerpunkt Lesen*. Cornelsen Verlag.
- Weishaupt, H. (2009). Die Hauptschule – ein Auslaufmodell? Eine Bestandsaufnahme im föderalen Vergleich. *Die Deutsche Schule*, 101(1), 20–31.
- Wiechmann, J. (2013). Der Einfluss dezentraler Entscheidungsebenen auf den schulischen Systemwandel in Schleswig-Holstein. In M. Rürup & I. Bormann (Eds.), *Innovationen im Bildungswesen* (pp. 111–124). Springer.

- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.
- Wilhelm, O., Schroeders, U., & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8–10)*. Hogrefe.
- Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record*, 112(4), 1008–1037.
- Woessmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives*, 30(3), 3–32.
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.
- Zeldow, B., & Hatfield, L. (2019). *Difference-in-differences*. Retrieved September 13, 2023, from <https://diff.healthpolicydatascience.org/#robustness>
- Zeldow, B., & Hatfield, L. A. (2021). Confounding and regression adjustment in difference-in-differences studies. *Health Services Research*, 56(5), 932–941.
- Zhou, H., & Ma, X. (2023). A partial simulation study of phantom effects in multilevel analysis of school effects: The case of school socioeconomic composition. *Sociological Methods & Research*, 52(3), 1321–1339. <https://doi.org/10.1177/0049124120986195>
- Zimmermann, T. (2018). Die Bedeutung signifikanter Anderer für eine Erklärung sozial differenzierter Bildungsaspirationen. *Zeitschrift für Erziehungswissenschaft*, 21(2), 339–360.
- Zimmermann, T. (2020). Social influence or rational choice? Two models and their contribution to explaining class differentials in student educational aspirations. *European Sociological Review*, 36(1), 65–81.
- Zunker, N., & Neumann, M. (2020). Stratifizierung durch Profilierung? Das Zusammenspiel von inhaltlichen Profilen und Merkmalen der Schülerkomposition an nicht-gymnasialen weiterführenden Schulen in Berlin. In C. I. Keßler & S. Nonte (Eds.), *(Neue) Formen der Differenzierung: Schul- und*

Klassenprofilierung im Spannungsfeld gesellschaftlicher Teilhabe und sozialer Ungleichheit (pp. 35–53). Waxmann.

Zunker, N., Neumann, M., & Maaz, K. (2018). Angebot und Nachfrage bei der Einzelschulwahl. Der Einfluss von Schulmerkmalen und der Zusammensetzung der Schülerschaft auf die Nachfrage nach weiterführenden Schulen in Berlin. *Zeitschrift für Pädagogik*, 64(5), 586–611.

A

**Socioeconomic Composition Effects
and *De-Tracking* Reforms**

A.1 Tables

Table A.1: RQ4: Mediation of SEC Effects by the Level of Educational Expectations in the States with Reforms – Comprehensive Schools (2).

2015		
Student data		
	β (SE)	<i>p</i>
Classroom level		
Reading Comprehension on		
SES (<i>M</i>)	22.45 (5.75)	.00
Educational Expectations (%)	-7.32 (8.68)	.40
Classroom level		
Educational Expectations on		
SES (<i>M</i>)	0.42 (0.04)	.00
Individual level		
SES	6.72 (2.29)	.00
Educational Expectations	41.86 (4.26)	.00
Immigrant Background	-42.42 (4.29)	.00
Gender	10.72 (3.30)	.00
<i>N</i> (Students)	3,850	
<i>N</i> (Classes)	215	

A.2 Authors' contributions

CRedit author statement: Marlen Holtmann: Conceptualization, Methodology, Data Curation, Writing - Original Draft, Visualization; Camilla Rjosk: Conceptualization, Methodology, Formal Analysis, Writing – Review & Editing, Visualization, Supervision, Project Administration; Oliver Lüdtke: Methodology, Formal Analysis, Writing – Review & Editing, Supervision; Petra Stanat: Conceptualization, Writing – Review & Editing, Supervision, Project Administration.

B

***De-Tracking* Reforms, Aspirations, and
SES**

B.1 Tables

Table B.1: *Mathematical Achievement Between 2012 and 2018.*

	2012		2018		Difference 2012 to 2018	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	Δ	<i>p</i>
States with Reforms	-0.722	1.138	-0.809	1.192	-0.087	0.002
States without Reforms	-0.509	1.089	-0.510	1.130	-0.001	0.940

B.2 Authors' contributions

CRedit author statement: Marlen Holtmann: Conceptualization, Methodology, Data Curation, Writing - Original Draft, Visualization; Camilla Rjosk: Conceptualization, Methodology, Formal Analysis, Writing – Review & Editing, Supervision, Project Administration; Malte Jansen: Conceptualization, Methodology, Formal Analysis, Writing – Review & Editing; Oliver Lüdtke: Methodology, Formal Analysis, Writing – Review & Editing, Supervision.

