

IfE



Erziehungswissenschaftliche
Studien
Band 11

Maria Krieg

Interessenentwicklung im Kontext von
Klassenprofilierung in der Sekundarstufe I

MINT
Klassenprofil
Fach- und Sachinteresse
Fähigkeitsselbstkonzept
Cross-Lagged-Panel-Modelle
Messinvarianz
Musik

Universitätsverlag Göttingen

Maria Krieg

Interessenentwicklung im Kontext von Klassenprofilierung
in der Sekundarstufe I

Dieses Werk ist lizenziert unter einer

[Creative Commons](https://creativecommons.org/licenses/by-sa/4.0/)

[Namensnennung - Weitergabe unter gleichen Bedingungen](https://creativecommons.org/licenses/by-sa/4.0/)

[4.0 International Lizenz.](https://creativecommons.org/licenses/by-sa/4.0/)



erschieden als Band 11 der Reihe „Erziehungswissenschaftliche Studien“
im Universitätsverlag Göttingen 2022

Maria Krieg

Interessenentwicklung
im Kontext von
Klassenprofilierung
in der Sekundarstufe I

Erziehungswissenschaftliche
Studien

Band 11

Universitätsverlag Göttingen
2022

Bibliografische Information der Deutschen Nationalbibliothek

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

Herausgeber:innen der Reihe „Erziehungswissenschaftliche Studien“:

Prof. Dr. Ariane Willems (Schriftleitung), Prof. Dr. Klaus-Peter Horn,
Prof. Dr. Katharina Kunze, Prof. Dr. Jessica Löser, Prof. Dr. Kerstin Rabenstein,
Prof. Dr. Tobias C. Stubbe, Prof. Dr. Hermann Veith

Georg-August-Universität Göttingen
Institut für Erziehungswissenschaft
Waldweg 26
37073 Göttingen
(<http://www.uni-goettingen.de/ife>)

Kontakt

Maria Krieg
E-Mail: mkrieg@uni-goettingen.de

Die vorliegende Arbeit wurde von der Sozialwissenschaftlichen Fakultät der Georg-August-Universität Göttingen unter demselben Titel als Dissertation angenommen.

Dieses Buch ist auch als freie Onlineversion über die Homepage des Verlags sowie über den Göttinger Universitätskatalog (GUK) bei der Niedersächsischen Staats- und Universitätsbibliothek Göttingen (<https://www.sub.uni-goettingen.de>) erreichbar. Es gelten die Lizenzbestimmungen der Onlineversion.

Satz und Layout: Maria Krieg
Umschlaggestaltung: Margo Bargheer



© 2022 Universitätsverlag Göttingen
<https://univerlag.uni-goettingen.de>
ISBN: 978-3-86395-544-1
DOI: <https://doi.org/10.17875/gup2022-1950>
ISSN: 2199-5133
eISSN: 2512-6024

Zusammenfassung

Profilierungsmaßnahmen auf Klassenebene (sogenannte Profilklassen) stellen ein zentrales Element aktueller Schulentwicklungsprozesse dar und gelten als prominenteste Form der praktischen Umsetzung zunehmender Schulautonomie. In der Sekundarstufe I gewinnen vor allem Profilklassen im musikalischen sowie naturwissenschaftlichen Bereich (Musik- bzw. MINT-Klassen) an Popularität. Einzelne Gestaltungsmerkmale des Profilverrichts (u.a. zusätzliche Unterrichtsstunden, innovative und praxisbezogene Lehr-Lern-Formate sowie der Einbezug außerschulischer Lernorte) werden mit vielfältigen pädagogischen Potentialen verbunden – insbesondere für das Interesse sowie das Fähigkeitsselbstkonzept der Schülerinnen und Schüler. Hinsichtlich einer differenzierten Analyse der Entwicklung des Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts der Schülerinnen und Schüler in Profilklassen der Sekundarstufe I besteht aktuell jedoch ein Forschungsdesiderat.

Die Ergebnisse der vorliegenden Arbeit geben Aufschluss über die Entwicklung des durchschnittlichen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts zwischen der fünften und siebten Jahrgangsstufe in den Domänen Musik und Naturwissenschaften. Dabei wird der Effekt des Besuchs einer Musik- oder MINT-Profilklassen herausgestellt. Anhand latenter Cross-Lagged-Panel-Modelle werden darüber hinaus die Stabilität sowie die längsschnittliche Zusammenhangsstruktur

des Fach- beziehungsweise Sachinteresses und des Fähigkeitsselbstkonzepts untersucht und mögliche Unterschiede zwischen Profilklassen und deren Parallelklassen ohne Profil beleuchtet. Die Befunde ermöglichen die Evaluation des Potentials von Profilklassen zur Förderung des Interesses sowie des Fähigkeitsselbstkonzepts und bieten vielfältige Anknüpfungspunkte für die pädagogische Praxis.

Inhaltsverzeichnis

Zusammenfassung.....	5
1 Einleitung.....	11
2 Interesse.....	17
2.1 Historische Grundlagen pädagogisch-psychologischer Interesstheorien	18
2.2 Das Interessenkonstrukt in der Person-Gegenstands-Relation	20
2.3 Situationales und individuelles Interesse	21
2.4 Fach- und Sachinteresse	24
2.5 Interessengnese und -differenzierung – theoretische Annahmen.....	27
2.6 Empirische Befunde zur Interessenentwicklung in der Sekundarstufe I	29
3 Fähigkeitsselbstkonzept	35
3.1 Theoretische Konzeption.....	36
3.2 Struktur des Selbstkonzepts	37
3.3 Die Rolle des Vergleichs für die Selbstkonzeptgenese	39

3.4	Empirische Befunde zur Entwicklung des Selbstkonzepts in der Sekundarstufe I.....	41
4	Erwartungs-Wert Theorie.....	43
4.1	Das Modell – theoretische Annahmen.....	44
4.2	Empirische Befunde zum Zusammenhang der Wert- und Erwartungskomponente	47
5	Schul- und Klassenprofilierung	53
5.1	Profilbildung im Kontext verstärkter Schulautonomie und Wettbewerbsorientierung	54
5.2	Verbreitung von Schul- und Klassenprofilen.....	58
5.3	Wahl von Profilangeboten.....	61
5.4	Rechtliche Rahmenbedingungen zur Unterrichtsgestaltung in Profilklassen	66
5.5	Musik- und MINT-Profilklassen – Unterrichtsgestaltung und Unterrichtsqualität	68
5.6	Potentiale von Profilklassen – empirische Befunde.....	74
5.7	Entwicklung des Interesses und des Fähigkeitsselbstkonzepts in Profilklassen	85
6	Zwischenfazit und Forschungsdesiderate.....	89
7	Forschungsfragen und Hypothesen.....	95
8	Datengrundlage und methodisches Vorgehen	101
8.1	Die Studie ProBiNi.....	101
8.1.1	Stichprobenbeschreibung	104
8.1.2	Erhebungsinstrumente.....	107
8.2	Ordinales Skalenniveau – forschungsmethodische Aspekte.....	113
8.3	Umgang mit fehlenden Werten	116
8.4	Methodisches Vorgehen	118
8.4.1	Strukturgleichungsmodellierung.....	119
8.4.2	Model-Fit.....	120
8.4.3	Konfirmatorische Faktorenanalyse	124
8.4.4	Messinvarianz	126
8.4.5	Cross-Lagged-Panel-Modelle	132
8.4.6	Software und analytisches Vorgehen	136

9	Empirische Analysen.....	139
9.1	Die Konstrukte: Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept	140
9.1.1	Überprüfung der Faktorenstruktur	140
9.1.2	Messinvarianz im Längsschnitt.....	141
9.1.3	Messinvarianz im Gruppenvergleich	147
9.2	Deskriptive Befunde.....	150
9.2.1	Entwicklung des Fachinteresses in Profilklassen und Klassen ohne Profil	151
9.2.2	Entwicklung des Sachinteresses in Profilklassen und Klassen ohne Profil	153
9.2.3	Entwicklung des Fähigkeitsselbstkonzepts in Profilklassen und Klassen ohne Profil	155
9.3	Cross-Lagged-Panel-Modelle	158
9.3.1	Musikalisches Interesse und Fähigkeitsselbstkonzept	159
9.3.2	Naturwissenschaftliches Interesse und Fähigkeitsselbstkonzept	162
9.3.3	Musikalisches Interesse und Fähigkeitsselbstkonzept in Musik-Klassen.....	166
9.3.4	Naturwissenschaftliches Interesse und Fähigkeitsselbstkonzept in MINT-Klassen.....	170
10	Zusammenfassung und Diskussion der Ergebnisse.....	177
10.1	Konstruktvalidität und Messinvarianz.....	178
10.2	Entwicklung des Interesses und des Fähigkeitsselbstkonzepts in Profilklassen	179
10.3	Zusammenhangsstruktur des Interesses und des Fähigkeitsselbstkonzepts	185
10.4	Implikationen für die Praxis.....	190
10.5	Forschungsdiesiderate	192
11	Verzeichnisse	195
11.1	Literaturverzeichnis	195
11.2	Abbildungsverzeichnis	231
11.3	Tabellenverzeichnis	233

1 Einleitung

Wer kein Interesse hat, ist nicht gebildet.

Hans Schiefele, 1986¹

Das Zitat von Hans Schiefele verdeutlicht den besonderen Stellenwert von interessen geleiteten Lernprozessen, der damals wie heute Bestand hat. Denn das Interesse von Schülerinnen und Schülern wird nicht nur als einer der einflussreichsten Bedingungsfaktoren des schulischen Lernens bezeichnet (Krapp, 1998), die Förderung von Interessen wird auch als entscheidendes Ziel schulischer Bildung formuliert (Willems, 2018). Interessierte Schülerinnen und Schüler sind motivierter, sich mit bestimmten Unterrichtsinhalten zu beschäftigen (Harackiewicz & Knogler, 2017) und weisen ein höheres Maß an Aufmerksamkeit auf (Hidi, Renninger & Krapp, 2004). Im schulischen Kontext sind durch Interessen angeregte Lernprozesse auch deshalb besonders bedeutsam, da diese die Schülerinnen und Schüler dazu veranlassen, sich neues Wissen über einen Gegenstand anzueignen oder bereits vorhandenes Wissen zu vertiefen (Krapp, 1998). Wiederholt haben empirische For-

¹ Schiefele, H. (1986). Interesse – Neue Antworten auf ein altes Problem. *Zeitschrift für Pädagogik*, 32(2), 153–162.

schungsergebnisse jedoch fächerübergreifend einen erheblichen Interessenrückgang im Laufe der Schulzeit aufgezeigt – insbesondere nach dem Übergang in die Sekundarstufe I (vgl. u.a. Anderhag et al., 2016; Daniels, 2008; Ferdinand, 2014; Fredricks & Eccles, 2002; Gottfried, Fleming & Gottfried, 2001; Häußler & Hoffmann, 1995; Potvin & Hasni, 2014; Potvin, Hasni, Sy & Riopel, 2020; Watt, 2004; Wigfield et al., 1997), sodass stetig nach Erklärungen für die Interessenabnahme sowie nach Förderungsmöglichkeiten des Interesses von Schülerinnen und Schülern gesucht wird (vgl. u.a. Ehrtmann, Wolter & Hannover, 2019; Harackiewicz & Knogler, 2017; Heß, 2011a; Krieg & Stubbe, 2018; Potvin et al., 2020; Willems, 2018). Vor allem die mathematisch-naturwissenschaftlichen Fächer (vgl. Daniels, 2008; Frenzel, Goetz, Pekrun & Watt, 2010; Gottfried et al., 2001; Häußler & Hoffmann, 1995; Krapp, 1996; Lazarides, Dicke, Rubach & Eccles, 2020) sowie das Fach Musik (Wigfield et al., 1997) scheinen vom Interessenrückgang betroffen zu sein. Die negative Entwicklung des Interesses geht zudem häufig mit einem negativen Trend des Fähigkeitsselbstkonzepts im Verlauf der Sekundarstufe I einher (vgl. u.a. Dickhäuser, Janke, Praetorius & Dresel, 2017; Lazarides et al., 2020; Marsh & Ayotte, 2003; Scherrer & Preckel, 2019; Wigfield et al., 1997).

Eine Möglichkeit, der Verringerung des Interesses und des Fähigkeitsselbstkonzepts nach dem Übergang in die Sekundarstufe entgegenzuwirken, könnte der Besuch von Profilklassen sein. Besonders musikalische und naturwissenschaftliche Profilklassen (sogenannte Musik- und MINT-Klassen) gehören zu den beliebtesten Profilangeboten (Weber, Nonte & Stubbe, 2021) und eröffnen neue Möglichkeiten einer innovativen Unterrichtsgestaltung. Die spezifischen Gestaltungsmerkmale von Profilklassen (u.a. zusätzliche Unterrichtsstunden, praxisbezogene Lehr-Lern-Formate sowie Kooperationen mit außerschulischen Institutionen und Lernorten) werden in der Forschungsliteratur mit vielfältigen pädagogischen Potentialen verbunden. Dazu zählen neben der Förderung der allgemeinen schulischen Leistungen und des Sozialverhaltens insbesondere positive Effekte für das Interesse sowie das Fähigkeitsselbstkonzept der Schülerinnen und Schüler (vgl. Eder, 2011; Göllner, 2017; Heß, 2017a; Seidel et al., 2016). Während einzelne Gestaltungsmerkmale der Profilangebote hinsichtlich ihrer pädagogischen Potentiale für das Interesse und Fähigkeitsselbstkonzept vergleichsweise gut beforscht sind (vgl. u.a. Fritzsche, Kröner & Pfeiffer, 2011; Heß, 2011a; Mokhonko, 2016; Nonte & Schurig, 2019; Nonte & Schwippert, 2012; Pabst-Krueger, 2020; Potvin & Hasni, 2014; Simon, 2019; Steegh, Höffler, Höft & Parchmann, 2021b), fehlt es jedoch aktuell an differenzier-ten Untersuchungen zur Entwicklung des domänenspezifischen Interesses und des Fähigkeitsselbstkonzepts der Schülerinnen und Schüler in Profilklassen der Sekundarstufe I (vgl. Krieg & Stubbe, 2018; Koivuhovi, Vainikainen & Kalalahti, 2021). Dieses Forschungsdesiderat wird in der vorliegenden Arbeit aufgegriffen und mögliche positive Effekte des Besuchs einer Profilklassen auf die Interessen- und Fähigkeitsselbstkonzeptentwicklung der Schülerinnen und Schüler empirisch überprüft.

Ein geeigneter theoretischer Rahmen für die integrierende Analyse des Interesses und des Fähigkeitsselbstkonzepts im Kontext spezifischer Bedingungen des Sozialisationsumfeldes, wie wir sie in Profilklassen vorfinden, stellt die Expectancy-Value Theory (EVT; dt. Erwartungs-Wert Theorie; Eccles & Wigfield, 2002) dar. Nach der EVT können die Erwartungs- (Fähigkeitsselbstkonzept) und Wertkomponente (Interesse) als positiv miteinander verknüpfte Konstrukte betrachtet werden (vgl. Wigfield, Tonks & Klada, 2016). Die EVT ist in der pädagogisch-psychologischen Forschung aktuell eines der am weitesten verbreiteten Modelle (Arens, 2021; Gaspard, Hasselhorn, Nagengast & Trautwein, 2019) und wurde bereits auf vielfältige Fragestellungen im schulischen Kontext angewendet (vgl. u.a. Arens, 2021; Gaspard, Laueremann, Rose, Wigfield & Eccles, 2019; Guo, Marsh, Parker, Morin & Dicke, 2017; Köller, Meyer, Saß & Baumert, 2019; Lowe & Coy, 2016; Nagy, Trautwein, Baumert, Köller & Garrett, 2006; Nonte, Krieg & Stubbe, 2021). Längsschnittliche Analysen, die Fragen der Veränderbarkeit und Zusammenhänge der Erwartungs- und Wertkomponente am Übergang in die Sekundarstufe I adressieren, sind hingegen selten. Zudem wurden vor allem die Fächer Naturwissenschaften und Musik in den bisherigen längsschnittlichen Analysen ausgespart. Erste empirische Evidenzen für die Domänen Mathematik und Englisch verweisen auf positive Effekte des vorherigen Fähigkeitsselbstkonzepts auf das spätere fachspezifische Interesse (Arens, Schmidt & Preckel, 2019).

Aus diesen zwei Forschungsperspektiven – einerseits des angenommenen positiven Effekts des Besuchs einer Profilklasse auf die Entwicklung des Interesses und des Fähigkeitsselbstkonzepts sowie andererseits der Frage nach der längsschnittlichen Beziehungsstruktur der Erwartungs- (Fähigkeitsselbstkonzept) und Wertkomponente (Interesse) zu Beginn der Sekundarstufe I – ergibt sich die übergeordnete Kernfrage der vorliegenden Arbeit:

Wie entwickelt sich das musikalische und naturwissenschaftliche Fach- beziehungsweise Sachinteresse sowie das domänenspezifische Fähigkeitsselbstkonzept der Schülerinnen und Schüler während der Sekundarstufe I und welche Rolle spielt dabei der Besuch einer Musik- beziehungsweise MINT-Profilklasse?

Zur Untersuchung dieser Forschungsfrage ist die Arbeit wie folgt gegliedert: Die erste Säule dieser Arbeit besteht aus der theoretischen Einführung der für diese Arbeit zentralen Konstrukte des Interesses (Kapitel 2) und des Fähigkeitsselbstkonzepts (Kapitel 3) sowie deren Einbettung in den nationalen und internationalen Forschungsstand. Dabei wird ein besonderer Fokus auf Entwicklungsprozesse des Interesses und des Fähigkeitsselbstkonzepts während der Sekundarstufe I gelegt. Die zwei zunächst parallel entwickelten Forschungsstränge zum Interessenkonstrukt sowie zum Fähigkeitsselbstkonzept können anschließend im vierten Kapitel im Rahmen der Expectancy-Value Theory (EVT; Eccles & Wigfield, 2002) miteinander verknüpft werden. Nach der Skizzierung der theoretischen Annahmen der EVT

werden empirische Befunde zu längsschnittlichen Zusammenhängen der Erwartungs- (Fähigkeitsselbstkonzept) und Wertkomponente (Interesse) zusammenfassend dargestellt. Zudem werden Unterschiede in der Operationalisierung der Wertkomponente als Interesse am Unterrichtsfach (Fachinteresse) oder Interesse an musikalischen beziehungsweise naturwissenschaftlichen Themen (Sachinteresse) gegenüberstellend diskutiert.

Die zweite zentrale Säule dieser Arbeit widmet sich den Profilierungsprozessen auf Schul- und Klassenebene (Kapitel 5). Neben der inhaltlichen Definition des Profilbegriffs und der Verortung des Themas Profilierung im Kontext der Schulautonomie-Debatte wird die Profilbildung als nationales ebenso wie internationales Phänomen der Schulentwicklung beschrieben. Die zunehmende Beliebtheit von Profilierungsmaßnahmen wird aus der Darstellung zur Verbreitung von Schul- und Klassenprofilen deutlich. Zudem werden einerseits die Bedingungsfaktoren der Profilverwahl sowie andererseits mögliche Aufnahmekriterien für eine Profilklass thematisiert. Neben den rechtlichen Rahmenbedingungen zur Einrichtung von Profilklassen und der Gestaltung von Profilverunterricht werden darüber hinaus konkrete organisatorische und didaktische Gestaltungsmerkmale des musikalischen sowie naturwissenschaftlichen Profilverunterrichts vorgestellt und umfassende Forschungsbefunde zu den mit den profilspezifischen Unterrichtskonzepten verknüpften pädagogischen Potentialen skizziert. Das fünfte Kapitel endet mit der Beschreibung der bisher vorliegenden empirischen Befunde zur Entwicklung des Fähigkeitsselbstkonzepts sowie des Interesses in Profilklassen.

Die in den ersten fünf Kapiteln dargestellten vielschichtigen theoretischen Annahmen sowie umfassenden empirischen Befunde werden in einem Zwischenfazit (Kapitel 6) zusammengefasst und die sich daraus ergebenden Forschungsdesiderate benannt. Daran anschließend können in Kapitel 7 die Forschungsfragen formuliert und die entsprechenden Hypothesen aufgestellt werden. Im achten Kapitel wird die Datengrundlage sowie das methodische Vorgehen beschrieben. Dazu wird zunächst das Projekt ProBiNi (*Profilbildung an Niedersächsischen Gymnasien und Integrierten Gesamtschulen*; Stubbe, Nonte, Haas & Krieg, 2022) vorgestellt, in dessen Rahmen die Daten für die vorliegende Arbeit erhoben wurden. Nachfolgend werden die eingesetzten Erhebungsinstrumente der zentralen Konstrukte (domänenspezifisches Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept) mit entsprechenden Kennwerten beschrieben. Da im Analyseprozess ein besonderer Fokus auf den adäquaten Umgang mit einer ordinalen Datenstruktur gelegt wird, wird der Einsatz entsprechender Analysemethoden reflektiert und im Anschluss die Analyseverfahren der vorliegenden Arbeit beschrieben. Neben einer allgemeinen Einführung in die Strukturgleichungsmodellierung beinhaltet dieses Kapitel auch die Darstellung des methodischen Vorgehens in konfirmatorischen Faktorenanalysen, der Messinvarianzprüfung sowie in der Spezifizierung latenter Cross-Lagged-Panel-Modelle unter Berücksichtigung einer ordinalen Datenstruktur. Im neunten Kapitel werden – der Logik der Forschungsfragen systematisch folgend – die Ergebnisse der empirischen Analysen schrittweise beschrieben. Zunächst werden die für diese Arbeit zentralen

Konstrukte hinsichtlich ihrer Faktorenstruktur sowie der Messinvarianz im Längsschnitt sowie im Multigruppenvergleich überprüft. Anschließend werden deskriptive Befunde zur Entwicklung des durchschnittlichen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts in den Domänen Musik und Naturwissenschaften anhand von Mittelwertvergleichen präsentiert. Die Ergebnisse der latenten Cross-Lagged-Panel-Modelle geben Aufschluss über die längsschnittliche Stabilität sowie Zusammenhänge des fachspezifischen Interesses und des Fähigkeitsselbstkonzepts zwischen der fünften und siebten Jahrgangsstufe. Die latenten Cross-Lagged-Panel-Modelle werden zunächst für die Gesamtstichprobe sowie nachfolgend für Profilklassen und Klassen ohne Profil getrennt modelliert, um den Effekt des Klassenprofils in der Zusammenhangsstruktur herauszustellen. Im zehnten Kapitel werden die Ergebnisse abschließend zusammenfassend diskutiert und praktische Implikationen für die Schulpraxis abgeleitet. Bisher offene gebliebene Forschungsaspekte werden in einem Ausblick formuliert und mögliche Anknüpfungspunkte für zukünftige Untersuchungen benannt.

2 Interesse

Das Interesse gehört zu den zentralen Konstrukten der pädagogisch-psychologischen Forschung und wird als einer der einflussreichsten Bedingungsfaktoren des schulischen Lernens bezeichnet (Krapp, 1998). Im schulischen Kontext sind durch Interessen angeregte Lernprozesse vor allem auch deshalb besonders relevant, da diese mit einer hohen subjektiven Bedeutung sowie der Wissensaneignung über den Interessengegenstand einhergehen (Krapp, 2018). Die Förderung von Interessen wird daher als entscheidendes Ziel schulischer Bildung formuliert (Krapp, 1998; Willems, 2018). Wiederholt verwiesen empirische Forschungsergebnisse jedoch auf einen erheblichen Interessenrückgang im Laufe der Schulzeit (vgl. u.a. Daniels, 2008; Fredricks & Eccles, 2002; Gottfried et al., 2001; Häußler & Hoffmann, 1995; Watt, 2004; Wigfield et al., 1997). Folglich werden Erklärungsansätze für die Interessenverringern sowie Möglichkeiten zur Förderung des Interesses von Schülerinnen und Schülern im Schulkontext dringend gesucht (vgl. Daniels, 2008; Ehrmann et al., 2019; Harackiewicz & Knogler, 2017; Krieg & Stubbe, 2018; Ryan & Powelson, 1991; Todt & Schreiber, 1998).

Dieses Kapitel widmet sich unterschiedlichen theoretischen Zugängen zur Beschreibung von Interessen, wobei ein Fokus auf pädagogisch-psychologisch sowie didaktische Interessentheorien gelegt wird. Wie aus der kurzen historischen Abhandlung zum Interessenkonstrukt in Kapitel 2.1 deutlich wird, verweisen moderne

Interesstheorien auf die Gegenstandsspezifität von Interessen, die aus der Interaktion zwischen einer Person und ihrer Umwelt entstehen (Person-Gegenstands-Relation; Kapitel 2.2). Die Frage nach den Entstehungsbedingungen situationsspezifischer (situationaler) Interessen und deren Transformation in längerfristig stabile (individuelle) Interessen wird in Kapitel 2.3 behandelt. Dazu wird auf zwei Modelle der Interessengenese zurückgegriffen und neben der Person-Gegenstands-Theorie der sogenannten Münchener Gruppe (vgl. Krapp, 1998) auf das parallel dazu im us-amerikanisch-kanadischen Raum entwickelte Vier-Phasen-Modell der Interessenentwicklung von Hidi und Renninger (2006) eingegangen, die große Überschneidungen in der Konzeptualisierung des Interessenkonstrukts sowie dessen Entwicklung aufweisen (vgl. Hidi et al., 2004). Eine eher didaktisch geprägte Erweiterung bisheriger Interessentheorien stellt die Konzeption der sogenannten Kieler Gruppe (vgl. u.a. Häußler, 1987; Hoffmann & Lehrke, 1986) dar, die verschiedene Dimensionen von Interessengegenständen untersucht und zwischen dem Interesse am Unterrichtsfach (Fachinteresse) und Interessen an einem Themenbereich (Sachinteresse) unterscheidet (Kapitel 2.4). Aufgrund der vielfach berichteten empirischen Befunde zur Interessenabnahme im Schulverlauf (s.o.) werden in Kapitel 2.5 zunächst theoretische Erklärungsansätze zur Interessenabnahme beziehungsweise -differenzierung erläutert und anschließend nationale ebenso wie internationale Forschungsergebnisse zur Interessenentwicklung zusammengefasst, wobei ein Fokus auf die Sekundarstufe I gelegt wird (Kapitel 2.6).

2.1 Historische Grundlagen pädagogisch-psychologischer Interessentheorien

Prenzel, Lankes und Minsel (2000) bezeichnen das Interesse als „wichtiges, vielleicht sogar grundlegendes Forschungsgebiet der Pädagogik bzw. Pädagogischen Psychologie“ mit einer bereits über 200 Jahren andauernden Geschichte, in der jedoch „insgesamt wenig kontinuierliche[r] Theorieentwicklung betrieben wurde“ (Prenzel et al., 2000, S. 11). So finden sich historische Vorläufer heutiger interesstheoretischer Annahmen bereits in den Arbeiten von Herbart (1806) und Dewey (1913). Mit der in den 1930er Jahren aufkommenden behavioristischen Denkweise kam die pädagogisch-psychologische Interessenforschung fast gänzlich zum Erliegen. Stattdessen wurden Interessen nun vornehmlich Untersuchungsgegenstand der differenzialpsychologischen Berufsinteressenforschung (vgl. Holland, 1973), die Interessen als motivationales, dispositionales Persönlichkeitsmerkmal – im Sinne einer längerfristig stabilen Präferenz für bestimmte Handlungs- oder Wissensgebiete – verstand (Daniels, 2008; Krapp, 1998; Prenzel & Krapp, 1992; Schiefele, 1996).

In den 1960er und 1970er Jahren waren es vor allem Motivationstheorien, die unter einem eher leistungsbezogenen Fokus versuchten, Lernprozesse zu erklären (Prenzel et al., 2000; Rheinberg & Vollmeyer, 2000; vgl. u.a. Atkinson² & Lens, 1980). Hervorzuheben sind hier die sogenannten Erwartungs-Wert Modelle (vgl. Eccles et al., 1983), die Lernmotivation aus einer individuellen Bedeutsamkeit einer guten Leistung (Wert) und der Erfolgserwartung, dass diese gute Leistung auch erreichbar ist, erklären (für eine detaillierte Beschreibung s. Kapitel 4). Mit den Arbeiten von Todt (1978), der erneut unterrichts- und entwicklungspsychologische Aspekte in die Interessenforschung aufnahm, sowie insbesondere durch die zunächst theoretisch angelegten Arbeiten von Schiefele (1974) nahm die Interessenforschung in den 1970er Jahren eine erneute Wende und wurde unter einer pädagogischen Perspektive neu ausgerichtet. Schiefele (1974) war bestrebt, Auslöser und Bedingungen des Lernens in den pädagogischen Kontexten zu suchen und zu erklären (vgl. Prenzel et al., 2000). Damit wurde versucht, der Überbetonung des Leistungsmoments der Motivationsforschung entgegenzutreten, die implizierte, dass sich die Motivation von Schülerinnen und Schülern allein draus speist, entkoppelt vom jeweiligen Unterrichtsfach gute Leistungen zu erreichen (vgl. Heckhausen, 1989). Außerdem lautete die Kritik an bisherigen Interessenkonzeptionen, dass durch die einseitige Annahme des statischen, persönlichkeitsorientierten Charakters von Interessen über die Dynamik von Interessen hinweggesehen werde (vgl. Krapp, 2002) und die Berücksichtigung einer Inhalts- oder Gegenstandsspezifität der Interessen fehle (vgl. Rheinberg & Vollmeyer, 2000). Ausgehend von den Arbeiten der sogenannten *Münchener Gruppe* um Ulrich Schiefele und Andreas Krapp (vgl. u.a. Krapp, 1992a; Prenzel & Krapp, 1992; Schiefele, 2000, 2001) bildete sich im Diskurs mit anderen internationalen Forschungsgruppen (vgl. u.a. Hidi et al., 2004; Renninger, Hidi & Krapp, 1992) ein gemeinsames konzeptuelles Verständnis des Interessenkonstrukts aus, das die Grundlage moderner Interessentheorien darstellt. Im Zentrum der Münchener Interessentheorie sowie aller neueren Interessenkonzepte steht die sogenannte *Person-Gegenstands-Relation* (vgl. Krapp, 1992a; 1998, 1999, 2018; s. Kapitel 2.2) sowie die Annahme, Interessen seien sowohl als langfristiges als auch kurzfristiges, situationsspezifisches Konstrukt zu verstehen (vgl. Kapitel 2.3). Diese zentralen Annahmen der modernen Interessentheorie werden im Folgenden detailliert vorgestellt.

² Für eine detaillierte Beschreibung der Annahmen von Atkinson (1957) zur Erklärung von leistungsbezogenem Verhalten unter Berücksichtigung von Leistungsmotivation, Erfolgserwartung und individuellen Werten siehe Kapitel 4.1.

2.2 Das Interessenkonstrukt in der Person-Gegenstands-Relation

Der Münchener Interessentheorie folgend, ist als zentrales Gestaltungsmerkmal von Interessen ihre Gegenstandspezifität zu nennen (Krapp, 1992a, 1999, 2018). Allgemein kommen als Interessengegenstände spezifische Aufgaben, konkrete Objekte, Tätigkeitsbereiche oder Themengebiete in Frage. Im schulischen Kontext sind damit vor allem Lerngegenstände im Sinne von Unterrichtsinhalten oder Schulfächern gemeint (Krapp, 1999, 2001). Prenzel et al. (2000) betonen, dass Interessengegenstände von den jeweiligen Personen konstruiert werden und die Gegenstands konstruktion von den sozialen oder kulturellen Kontexten abhängig ist, in der sich eine Person befindet. Die Beziehung zwischen einer Person und entsprechenden Interessengegenständen ist durch drei zentrale Eigenschaften gekennzeichnet. Unterschieden werden die emotionale, wertbezogene und kognitive Merkmalskomponente des Interesses (Krapp, 1992a, 1999; Schiefele, 1996). Die *emotionale* Komponente beschreibt die Begleitung der Interessenhandlung durch positive emotionale Wahrnehmungen. Während der Beschäftigung mit einem Interessengegenstand empfindet eine Person folglich Freude, Kompetenzerleben, Selbstbestimmungsgefühl, Aktiv- oder Beteiligtsein und wird nicht durch Angst oder äußeren Druck behindert beziehungsweise geleitet (Krapp, 1998, 1999; Schiefele 1996). Die emotionale Merkmalskomponente wird auch als *gefühlsbetonte Valenz* (vgl. Pekrun, 1988) des Interesses bezeichnet und findet in angelehnten Konzepten ebenfalls unter den Begriffen *Lernfreude* (Helmke, 1993), *Flow* (Csikszentmihalyi, 2010) oder der Befriedigung der *Basic Needs* (Deci & Ryan, 1985, 1993) Anwendung.

Die *wertbezogene* Merkmalskomponente (auch wertbezogene Valenz) bezieht sich auf die subjektive Bedeutung, die eine Person einem Interessengegenstand beimisst. Die Bedeutungszuschreibung ist mit einer erhöhten Bereitschaft verbunden, sich intensiv mit dem Interessengegenstand zu beschäftigen und geht mit dem Bestreben einher, mehr über den Gegenstand zu erfahren („epistemische Orientierung“, Krapp, 2018, S. 287, vgl. Krapp, 1999; Prenzel, 1988). Die aus Interesse vollzogene Person-Gegenstands-Interaktion beinhaltet folglich auch eine *kognitive* Komponente. Durch die wiederholte Beschäftigung mit dem Interessengebiet erlangt die Person gegenstands spezifische Kompetenzen und erweitert beziehungsweise differenziert ihr Wissen (Krapp, 1998, 1999; Prenzel, 1988). Im schulischen Kontext sind aus Interesse vollzogene Lernhandlungen folglich von besonderer Relevanz, da sie die Schülerinnen und Schüler dazu veranlassen sich neues Wissen über den Gegenstand anzueignen (Krapp, 1998). Da das gegenstandsbezogene Wissen aber gleichzeitig sowohl Voraussetzung als auch Folge einer interessengeleiteten Handlung sein kann, wird die kognitive Komponente in späteren Arbeiten zur Münchener Interessentheorie „als eine von Interesse konzeptionell unabhängige Größe“ gesehen (Schiefele, 1996, S. 77).

Nach Schiefele (1996) sind interessengeleitete Handlungen zudem selbstbestimmt (selbstintentional), das heißt intrinsisch motiviert³. Für die Beschäftigung mit dem Interessengegenstand bedarf es folglich keiner äußeren Anreize (Belohnung oder auch Tadel), vielmehr liegen die Gründe für die Beschäftigung in der Handlung selbst. Die Selbstintentionalität bezieht Schiefele (1996) dabei sowohl auf die emotionale als auch wertbezogene Valenz. Dennoch ist die Selbstintentionalität kein weiteres Konstruktionsmerkmal des Interesses, sondern dient vielmehr der präziseren Definition der emotionalen und wertbezogenen Merkmalskomponente (Schiefele, 1996). Die intrinsisch emotionale Merkmalskomponente meint dabei, dass „die auf dem Interesse beruhende Handlung mit positiven Erlebensqualitäten bzw. Gefühlen verknüpft ist“ (Krapp, 1999, S. 400). Die intrinsisch wertbezogene Komponente beschreibt hingegen die Identifikation der Person mit dem Interessengegenstand (vgl. Krapp, 1999). Je höher die subjektive Bedeutsamkeit eines Interessengegenstandes, desto wichtiger ist der Gegenstand auch für das eigene Selbstbild und damit langfristig für die eigene Identität (Krapp, 1992a; 2007). Krapp (1999, S. 400) schreibt diesbezüglich: „Die mit dem Interessengegenstand verbundenen Intentionen (Handlungsziele) sind mit dem das „Selbst“ [Hervorhebung im Original] konstituierenden Einstellungen, Erwartungen und Werten kompatibel“. Demnach verweist die intrinsisch wertbezogene Komponente darauf, dass „der Gegenstand des Interesses vorübergehend oder dauerhaft in die zentralen Regionen des Selbstkonzepts einer Person integriert worden ist“ (Krapp, 1999, S. 400).

Darüber hinaus beschäftigte sich unter anderem die Münchener Interessentheorie mit der Frage nach den Entstehungsbedingungen situationsspezifischer (situationaler) Interessen sowie der Frage nach der Struktur zeitlich stabiler (individueller) Interessen (Krapp, 2018). Beide Konzepte werden im folgenden Kapitel näher erläutert.

2.3 Situationalles und individuelles Interesse

Moderne Interessenkonzeptionen unterscheiden zwischen dem situationalen und individuellen Interesse und verweisen damit auf den sowohl temporären als auch längerfristig stabilen Charakter von Interessen (vgl. Krapp, 1992a, 1999, 2018). Neben der Münchener Interessentheorie (vgl. Krapp, 1998) nimmt auch das Vier-Phasen-Modell zur Interessenentwicklung von Hidi und Renninger (2006) eine Unterscheidung zwischen dem situationalen und individuellen Interesse vor, wobei das situationale Interesse die erste Entwicklungsstufe hin zu einem längerfristig verankerten Interesse bildet (vgl. Hidi et al., 2004).

Das *situationale Interesse* wird als temporärer, situationsspezifischer motivationaler Zustand beschrieben, der primär durch Anregungen aus der (Lern-)Umwelt ausgelöst wird (Hidi & Renninger, 2006; Hidi et al., 2004; Krapp, 2018; Schiefele, 2000).

³ Zum Konzept der intrinsischen Motivation siehe Brandstätter, Schüler, Puca und Lozo (2013).

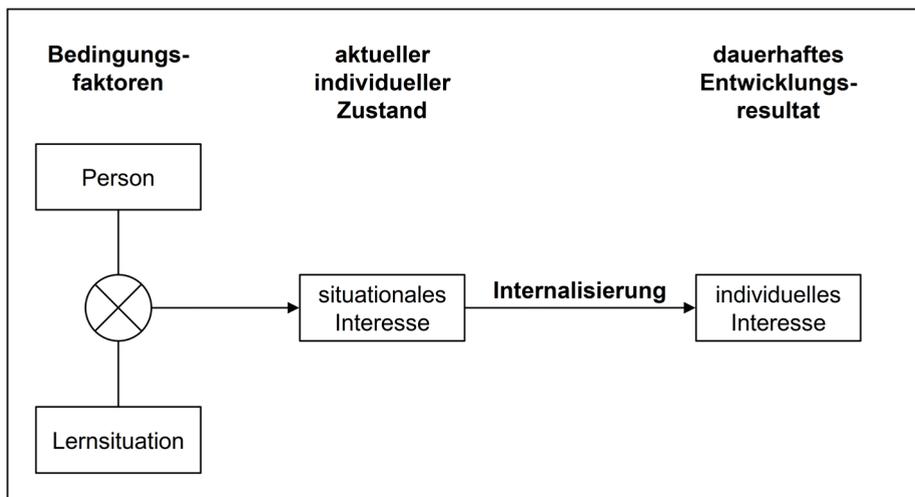
Schiefele (2000, S. 229) bezeichnet das situationale Interesse auch als „initiale Zuwendung zu einem Sachverhalt, der dann Inhalt einer Handlung oder eines Lernprozesses wird“ und betont damit, dass es für situationales Interesse keinem vorhandenen beziehungsweise zu aktivierenden Interesse bedarf. Das situationale Interesse wird von Beginn an durch eine epistemische Orientierung und gefühlsbetonte Valenz begleitet, während sich die wertbezogene Valenz und die Selbstintentionalität erst später entwickeln (Schiefele, 2000; s. Kapitel 2.2). Da es für die Entstehung von situationalem Interesse äußerer Anreize bedarf, kommt der Umwelt und der Qualität der darin enthaltenen externen Anreize eine entscheidende Rolle zu. Im schulischen Kontext bedeutet dies, dass eine didaktisch gelungene Aufbereitung der Unterrichtsinhalte das situationale Interesse der Schülerinnen und Schüler wecken und für eine kurze Zeit aufrechterhalten kann (Hidi et al., 2004; Krapp, 1998). Auch die soziale Lernumgebung (z.B. mono-edukativer Physikunterricht für Mädchen; vgl. Häußler & Hoffmann, 1998), aktive Lehr-Lern-Formen (vgl. Palmer, 2009) oder außerschulische Lernorte (z.B. Science Museum, Schülerlabor; vgl. Simon, 2019) können situationales Interesse fördern (Hidi & Harackiewicz, 2000; Hidi et al., 2004). Unter bestimmten Voraussetzungen entwickelt sich aus dem temporär auftretenden situationalen Interesse ein zeitlich überdauerndes individuelles Interesse (Krapp, 1992a, 1998, 2018).

Das *individuelle* (persönliche) *Interesse* ist „als generalisierte[n] „habituelle“ [Hervorhebung im Original] Tendenz oder einer persönlichkeitspezifischen Disposition“ konzeptualisiert (Krapp, 1992a, S. 316). Das individuelle Interesse wird folglich als relativ stabile Persönlichkeitseigenschaft definiert und meint eine längerfristige Präferenz einer Person für einen bestimmten Interessengegenstand, „die – aufgrund situativer Gegebenheiten – ein gewisses Maß an Veränderlichkeit zulässt“ (Daniels, 2008, S. 28). Je ausgeprägter das individuelle Interesse einer Person ist, desto weniger Bedeutung kommt den externen Anreizen in der Interessenentwicklung zu, da die Person selbstintentional und aktiv Situationen aufsucht, die ihren Interessen entsprechen (Krapp, 1992a). Interessenshandlungen sind mit positiven Emotionen verbunden und weisen außerdem eine hohe subjektive Bedeutung auf, was die Herausbildung des individuellen Interesses begünstigt (vgl. Kapitel 2.2; Krapp, 1992a). Die subjektive Bedeutsamkeit geht einher mit der persönlichen Identifikation mit dem entsprechenden Interessengegenstand, sodass Interessen das individuelle Selbstbild einer Person mit definieren (Daniels, 2008; Krapp, 2007). Krapp (1992a) verweist zudem darauf, dass mittels des Konzepts des individuellen Interesses über die Beschreibung der Persönlichkeitsstrukturen auch Rückschlüsse auf das Lernverhalten sowie Leistungsunterschiede von Schülerinnen und Schülern möglich sind.

Nach der Person-Gegenstands-Relation wird davon ausgegangen, dass sich im Entwicklungsprozess einer Person Präferenzen für bestimmte Wissens- oder Handlungsgebiete herausbilden, die die Person dazu veranlassen, sich wiederholt und ohne externe Anreize mit dem Interessengegenstand zu beschäftigen (Krapp, 1992b). Abbildung 2.01 zeigt schematisch den Prozess der Interessengenese vom

kurzfristigen situationalen Interesse hin zu einem langfristigen individuellen Interesse. Krapp (1998) beschreibt diese Transformation als mehrstufigen Prozess (vgl. Hidi et al., 2004). Dafür muss zunächst die durch externe Anreize entwickelte Neugier an einem Interessengegenstand in eine länger andauernde Beschäftigung mit dem Gegenstand überführt werden. Anschließend wird versucht, das Interesse am Gegenstand von den äußeren Anreizbedingungen zu entkoppeln und in ein längerfristiges individuelles Interesse zu überführen, das sich aus internen Faktoren speist (Krapp, 1998; vgl. Harackiewicz & Knogler, 2017).

Abbildung 2.01: Rahmenmodell der Interessengenese



Krapp, 1998, Abbildung 1, modifiziert

Ähnliche Annahmen finden sich auch im sogenannten *Vier-Phasen-Modell* der Interessenentwicklung von Hidi und Renninger (2006) sowie Hidi et al. (2004), das die Interessengenese als eine sich ständig weiterentwickelnden Person-Objekt-Beziehung in vier aufeinanderfolgenden Phasen beschreibt (vgl. Harackiewicz & Knogler, 2017). Das in der ersten Phase durch externe Faktoren angeregte situationale Interesse (*triggered situational interest*) kann in Phase zwei längerfristig aufrechterhalten werden (*maintained situational interest*), sofern die Aufgaben oder Inhalte als sinnvoll und subjektiv bedeutsam wahrgenommen werden. Phase drei kennzeichnet das beginnende individuelle Interesse (*emerging individual interest*), das entsteht, wenn das Interesse über die jeweiligen Situationen hinaus anhält und mit der Akkumulation von Wissen verbunden ist. Nehmen das Wissen über den Interessengegenstand sowie der damit verbundene Wert weiter zu, so kann sich das Interesse in Phase vier schließlich zu einem differenzierten individuellen Interesse entwickeln (*well-developed individual interest*; Hidi & Renninger, 2006; Hidi et al., 2004; vgl. Harackiewicz & Knogler, 2017). Hidi und Renninger (2006) betonen die Rolle der Umwelt für die

Bereitstellung entsprechender Gelegenheitsstrukturen im gesamten Prozess der Interessengenese. Selbst in Phase vier, in der das individuelle Interesse bereits gut entwickelt ist, sind demnach entsprechende interessenfördernde Gelegenheiten notwendig, um eine Verringerung oder gar das Verschwinden des Interesses zu vermeiden (Hidi & Renninger, 2006).

Im schulischen Kontext kann sich individuelles Interesse als Interesse an spezifischen Unterrichtsinhalten (Sachinteresse) oder Unterrichtsfächern (Fachinteresse) entwickeln und manifestieren (Hoffmann, 2002; Krapp, 1998). Aufbauend auf den Annahmen der Münchener Interessentheorie entstand das sogenannte *Kieler Interessenkonstrukt*, das einen theoretischen Rahmen zur Unterscheidung von Fach- und Sachinteresse bietet. Dieses wird im folgenden Kapitel näher erläutert.

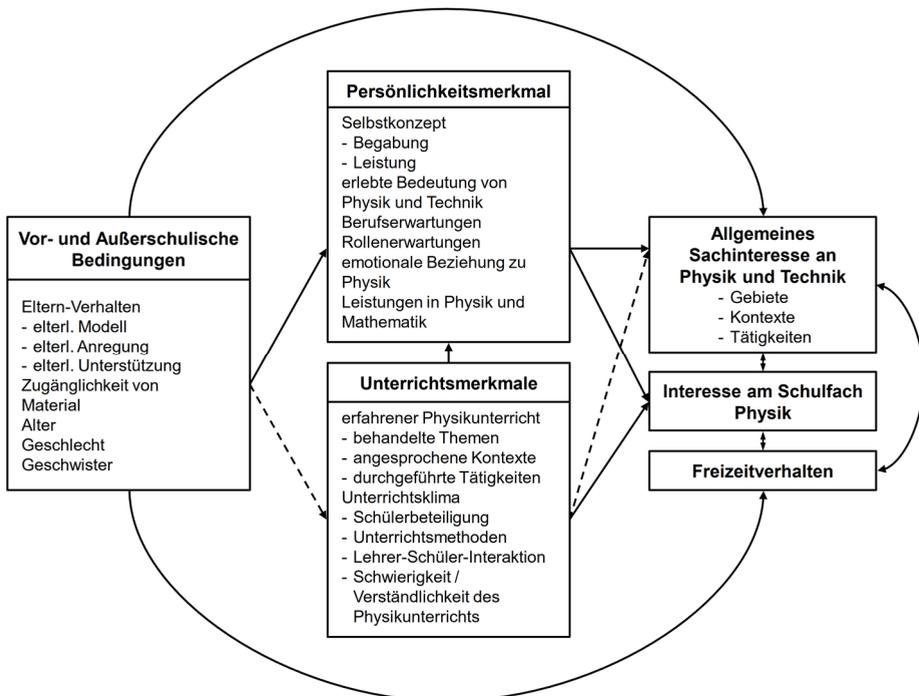
2.4 Fach- und Sachinteresse

Im Gegensatz zur eher psychologisch geprägten Interessenforschung der Münchener Gruppe, die innerhalb der Person-Gegenstands-Relation einen Fokus auf die Person legt, verfolgt die sogenannte Kieler Gruppe um die Forschenden Lore Hoffmann und Peter Häußler einen pädagogisch-didaktischen Ansatz und untersucht verschiedene Dimensionen der Interessengegenstände (vgl. Häußler, 1987; Häußler & Hoffmann, 1995, 1998; Hoffmann, 2002; Hoffmann & Lehrke, 1986). Die Kieler Interessenforschung greift zunächst einige der bereits beschriebenen Merkmale des Münchener Interessenkonstrukts auf. Dies betrifft die Person-Gegenstands-Relation, die annimmt, dass die Beziehung zwischen einer Person und einem Interessengegenstand durch verschiedene affektive und sowie kognitive Merkmale bestimmt und durch die Interaktion zwischen der Person und der Umwelt gebildet wird (vgl. Krapp, 1992a). Ebenfalls wird die Unterscheidung zwischen temporären (situationalen) Interessen sowie zeitlich relativ stabilen (individuellen) Interessen adaptiert (Häußler & Hoffmann, 1995; Krapp, 1992a). Eine Erweiterung bisheriger Thesen findet sich jedoch in der inhaltlichen Konzeptualisierung unterschiedlicher Interessengegenstände im Sinne des Interesses an einem Unterrichtsfach (Fachinteresse) oder als Interesse an einem Themenbereich (Sachinteresse; Hoffmann, 2002). In Anlehnung an Krapp, Hidi und Renninger (1992) beinhaltet nach Hoffmann (2002) das physikalische Fachinteresse sowohl ein generelles (individuelles) Interesse für Physik als auch ein kurzfristiges (situationales) Interesse an bestimmten Themen, das von der Gestaltung des konkreten Physikunterrichts anhängig ist. Je nachdem, in welchen thematischen Kontext die physikalischen Unterrichtsinhalte eingebettet werden oder welche Aktivitäten im Physikunterricht stattfinden, beeinflusst somit auch das Fachinteresse. Hoffmann (2002) betont demnach die Bedeutsamkeit einer interessanten Lernumgebung für die Entstehung von situationalem Interesse, das sich wiederum zu individuellem Interesse entwickeln kann (vgl. Krapp, 1992a). Das physikalische Sachinteresse wird als individuelles Interesse im

Sinne einer längerfristigen (aber veränderbaren) Präferenz (Disposition) für bestimmte physikalische Themen interpretiert (Hoffmann, 2002). Da das Sachinteresse aber nicht nur durch den thematischen Gegenstand an sich, sondern auch durch den Kontext, in den das Thema eingebettet ist und die Art der mit dem Thema verbundenen Aktivitäten bestimmt wird, entwickelten Häußler, Frey, Hoffmann, Rost und Spada (1988) ein dreidimensionales Konstrukt des Sachinteresses.

Abbildung 2.02 verdeutlicht die im Kieler Interessenmodell angenommene Beziehungsstruktur zwischen Fach- und Sachinteresse sowie weitere für die Interessenstruktur relevante Faktoren. Dazu gehören einerseits Merkmale der Unterrichtsgestaltung, wie thematische Inhalte, Tätigkeiten oder die Beteiligung der Schülerinnen und Schüler, sowie Persönlichkeitsmerkmale, wie die subjektive Bedeutung, emotionale Valenz oder das Selbstkonzept von Schülerinnen und Schülern. In der Abbildung ebenfalls aufgeführt, ist die Untergliederung des Sachinteresses in die drei Dimensionen des kontextuellen, inhaltlichen und tätigkeitsbezogenen Interesses.

Abbildung 2.02: Bedingungen des Fach- und Sachinteresses



Ausgangspunkt für die weitere Ausdifferenzierung des physikalischen Sachinteresses bildete die Veröffentlichung von Gardner (1985), in der unter anderem kritisiert wurde, Interessen wären in bisherigen Studien lediglich eindimensional erfasst worden und es sei unklar, welche Komponenten des Interessenkonstrukts gemessen würden (vgl. Häußler, 1987).

Die Grundlage des daraufhin entwickelten Kieler Interessenkonstrukts bildete die Studie *Physikalische Bildung für heute und morgen*, in der ein curriculares Modell physikalischer Bildung entstand (Häußler, Frey, Hoffmann, Rost & Spada, 1980; vgl. Daniels, 2008). Zur Vorbereitung der Operationalisierung des physikalischen Sachinteresses griffen die Forscherinnen und Forscher auf Ergebnisse der sogenannten curricularen *Delphi-Studie* zurück. Häußler et al. (1980) befragten dazu Expertinnen und Experten „nach der pädagogisch wünschbaren physikalischen Bildung“ (Häußler & Hoffmann, 1995, S. 109), das heißt nach Kontexten, Themen und Tätigkeiten, die für die physikalische Bildung als essentiell bewertet werden. Das aus den Aussagen von Häußler et al. (1980) entwickelte Modell des Sachinteresses an physikalischen Themen umfasst drei Dimensionen: (a) Interesse am Kontext, in dem Physik eine wichtige Rolle spielt, (b) Interesse an einem physikalischen Inhalt beziehungsweise inhaltlichen Gebiet, das in diesen Kontext eingebettet ist und (c) Interesse an mit diesen Themen verknüpften Tätigkeiten (vgl. Häußler & Hoffmann, 1995).

Durch das dreidimensionale Modell des Sachinteresses (kontextuelles, inhaltliches, tätigkeitsbezogenes Interesse) kann berücksichtigt werden, für welche Bereiche physikalischer Bildung Schülerinnen und Schüler Interesse aufbringen. Darüber hinaus „wird dem Umstand Rechnung getragen, dass die Reaktion einer Person auf die Frage, wie interessant für sie ein physikalisches Gebiet ist, davon beeinflusst wird, welche Tätigkeiten oder Kontexte mehr oder weniger zufällig von dieser Person assoziiert werden“ (Daniels, 2008, S. 28). Empirisch konnte gezeigt werden, dass das in dieser Weise operationalisierte physikalische Sachinteresse und das Fachinteresse für Physik nur eine mittlere Korrelation aufweist und dass das physikalische Sachinteresse lediglich 32 Prozent der Varianz des Fachinteresses erklären kann (Hoffmann, 2002). Daher muss angenommen werden, dass das Fach- und Sachinteresse getrennt voneinander zu betrachtende Konstrukte darstellen (Häußler & Hoffmann, 1995; Hoffmann, 2002; vgl. Heß, 2017b). Die Skala des dreidimensionalen physikalischen Sachinteresses wurde im Rahmen der sogenannten *IPN-Interessenstudie Physik*, die zwischen 1984 und 1989 am Leibniz-Institut Kiel durchgeführt wurde, eingesetzt (vgl. Häußler & Hoffmann, 1995; Hoffmann, Häußler & Lehrke, 1998; Hoffmann & Lehrke, 1986). Die IPN-Interessenstudie untersuchte die Interessenentwicklung von Schülerinnen und Schülern des fünften bis zehnten Jahrgangs im Fach Physik. Die Ergebnisse werden im Kapitel 2.6 im Anschluss an einige theoretische Überlegungen zur Interessenentwicklung beziehungsweise -differenzierung während der Sekundarstufe I (Kapitel 2.5) eingehend beschrieben.

2.5 Interessengenes und -differenzierung – theoretische Annahmen

Modelle der Interessengenes beschreiben, wie eine Person durch die fortlaufende Beschäftigung mit dem Interessengegenstand bereits vorhandene (situationale) Interessen vertiefen, um sie in stabilere (individuelle) Interessen umzuwandeln (s. Krapp, 1998, 1999, 2018 sowie Vier-Phasen-Modell der Interessenentwicklung n. Hidi & Renninger, 2006; Hidi et al., 2004). Dies wurde im Kapitel 2.3 detailliert dargelegt. Dabei wurde auch die Rolle der Umwelt und die darin enthaltenen externen Anreize sowie (bestenfalls) interessenfördernden Gelegenheiten (z.B. im Rahmen der Unterrichtsgestaltung) betont, um die Interessenentwicklung nicht zu unterbrechen oder gar eine Verringerung der Interessen zu bewirken (vgl. Harackiewicz & Knogler, 2017). Da nationale und internationalen Studien jedoch wiederholt auf einen Interessenrückgang im Laufe der Schulzeit verweisen (vgl. u.a. Anderhag et al., 2016; Daniels, 2008; Fredricks & Eccles, 2002; Gottfried et al., 2001; Häußler & Hoffmann, 1995; Potvin & Hasni, 2014; Wigfield et al., 1997), werden verschiedene Erklärungsansätze verfolgt. Diese nehmen Aspekte der Interessendifferenzierung sowie entwicklungsbedingte Faktoren in den Blick, um die Interessenabnahme zu erklären (vgl. zusammenfassend Daniels, 2008; Ferdinand, 2014).

Interessendifferenzierung als Grund für den Interessenrückgang

Da die Interessen einer Person in jungen Jahren noch sehr unspezifisch sind, spricht Travers (1978) von *universellen Interessen* im Vorschulalter. Kleinkinder richten ihr Interesse demnach auf unterschiedlichste Gegenstandsbereiche der sie umgebenden Umwelt und verwenden vielfältige Möglichkeiten für die Suche nach einer individuellen Interessenstruktur. Todt (1990) sowie Todt und Schreiber (1998) greifen das Konzept der universellen (*allgemeinen*) Interessen nach Travers (1978) auf und konstatieren mit zunehmendem Alter eine phasenweise Interessendifferenzierung hin zu *spezifischen Interessen*. Allgemeine Interessen definiert Todt (1990) als „Verhaltens- oder Handlungstendenzen (Dispositionen), die relativ überdauernd und relativ verallgemeinert sind“ (Todt, 1990, S. 225). Diese stehen in einem engen Zusammenhang mit der kognitiven Entwicklung und der Entwicklung eines Selbstbildes und wirken zudem als Filter, nach dem eine Person Erfahrungen einordnet und eine Auswahl von Tätigkeiten trifft mit denen sie sich beschäftigen möchte (Todt, 1990). Ausgelöst durch externe Anreize kommt es nach Todt (1985) anschließend zur wiederholten Beschäftigung mit einem bestimmten Interessengegenstand. Ist diese Beschäftigung mit positiven Effekten verbunden (z.B. Erfolg oder Anerkennung; vgl. Kapitel 2.2 emotionale Valenz; Krapp, 1998, 1999), so kann sich daraus ein recht stabiles spezifisches Interesse entwickeln. Da es zur Ausbildung spezifischer Interessen demnach externer Anregung bedarf, kommt der Umwelt eine besondere Bedeutung zu. Die zunehmende Interessendifferenzierung geht dabei mit einer gleich-

zeitigen Reduzierung potentieller Interessengebiete (bzw. Kompetenzgebiete) einher (Todt, 1990; Todt & Schreiber, 1998; vgl. Prenzel et al., 2000). Ab einem Alter von etwa sieben Jahren, das heißt mit der Einschulung, wird die Fähigkeitsselbst-einschätzung ein zunehmender Faktor der Interessengenese. Mit fortschreitendem Alter passen sich Interessen und Fähigkeiten (bzw. dessen Selbstwahrnehmung) stärker aneinander an (Todt & Schreiber, 1998). Entsprechend spezialisieren sich im Schulverlauf die Interessen der Schülerinnen und Schüler, da bestimmte Interessengegenstände nicht ins Repertoire der individuellen Interessenstruktur aufgenommen werden. Den Abschluss der Interessendifferenzierung datiert Todt (1985) auf das Ende der allgemeinbildenden Schule, sodass Personen im Erwachsenenalter eine spezifische Struktur individueller Interessen aufweisen.

Empirisch ergeben sich aus der zunehmenden Interessendifferenzierung entsprechend folgende Implikationen: Wird über einen längeren Zeitraum das durchschnittliche Interesse einer spezifischen Gruppe (z.B. einer Schulklasse) in verschiedenen Domänen analysiert, so ist die gemessene Stabilität der fachspezifischen Interessen davon abhängig, wie viele Schülerinnen und Schüler entsprechend spezielle Interessen in diesen Domänen entwickeln und wie viele Schülerinnen und Schüler sich von diesen Bereichen abwenden. Interessieren sich zunehmend mehr Schülerinnen und Schüler für ein bestimmtes Fach (d.h. sie bilden spezifische Interessen im Fach X aus), steigt das durchschnittliche Interesse der Untersuchungsgruppe in diesem Fach. Wenden sich hingegen mehr Schülerinnen und Schüler vom Fach X ab (und interessieren sich stattdessen für Fach Y), so sinkt das durchschnittliche Interesse der Untersuchungsgruppe in Fach X. Somit liefert der Ansatz der Interessendifferenzierung eine Erklärung für die durchschnittliche Interessenabnahme in bestimmten Fächern.

Entwicklungsbedingte Faktoren als Grund für den Interessenrückgang

Der entwicklungstheoretische Erklärungsansatz sucht die Faktoren des Interessenrückgangs hingegen in verschiedenen Entwicklungsaufgaben, vor denen Kinder und Jugendliche während des Heranwachsens und im Schulverlauf stehen. Dazu gehört zum einen die Ablösung von den Eltern bei einer gleichzeitigen vermehrten Hinwendung zu Gleichaltrigen (vgl. Lohaus & Vierhaus, 2019), durch die sich die Bezugsnorm des sozialen Vergleichs ändert und somit eine Interessenverschiebung bedingen kann (zur Rolle sozialer Beziehungen in der Interessenentwicklung s. u.a. Ryan & Powelson, 1991; vgl. Kapitel 3.3). Eine weitere zentrale Entwicklungsaufgabe stellt die Herausbildung einer eigenen Identität durch die Übernahme sozialer Rollen dar (vgl. Erikson, 1959). Damit eng verbunden ist die Geschlechtsrollenentwicklung sowie die Ausbildung geschlechtsstereotypischer Interessenprofile (vgl. Ehrtmann et al., 2019). Während sich Jungen stereotypisch eher für die mathematisch-naturwissenschaftlichen Fächer interessieren, sind künstlerisch-kreative Fächer eher mit dem Weiblichen assoziiert (Eccles, Wigfield, Harold & Blumenfeld, 1993; Ehrtmann et al., 2019; Hannover & Kessels, 2002; Kessels, 2005; Kessels &

Hannover, 2004). Vielfach in der Forschungsliteratur belegt ist zudem der enge Zusammenhang zwischen dem Rückgang (bzw. der Differenzierung) der Fähigkeits-selbsteinschätzungen und der Interessenentwicklung im Schulverlauf (vgl. u.a. Daniels, 2008; Fredricks & Eccles, 2002; Häußler & Hoffmann, 1995; Wigfield et al., 1997). Der Zusammenhang zwischen den Entwicklungsprozessen des Fähigkeits-selbstkonzepts sowie des Interesses von Schülerinnen und Schülern steht daher auch im Zentrum der vorliegenden Arbeit und wird in den folgenden Kapiteln 3 und 4 detailliert theoretisch beschrieben sowie durch umfassende Forschungsergebnisse empirisch belegt.

2.6 Empirische Befunde zur Interessenentwicklung in der Sekundarstufe I

Am Ende der Grundschulzeit weisen Schülerinnen und Schüler noch eine überwiegend positive schulische Einstellung (gemessen als subjektiver Wert/Interesse; vgl. Kapitel 4.2; vgl. Eccles & Wigfield, 2002) auf, wie aktuelle Ergebnisse aus TIMSS 2019 (*Trends in International Mathematics and Science Study*) beispielsweise für die Fächer Mathematik und Naturwissenschaften zeigen (Nonte, Steinmayr & Scholz, 2020). Nach dem Übergang in die Sekundarstufe I verringert sich das durchschnittliche Interesse der Schülerinnen und Schüler an schulischen Inhalten sowie am Unterricht jedoch erheblich, wie verschiedene Studien belegen (vgl. u.a. Anderhag et al., 2016; Daniels, 2008; Ferdinand, 2014; Fredricks & Eccles, 2002; Frenzel et al., 2010; Gottfried et al., 2001; Häußler & Hoffmann, 1995; Hoffmann et al., 1998; Hoffmann & Lehrke, 1986; Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002; Köller, Baumert & Schnabel, 2001; Potvin & Hasni, 2014; Watt, 2004; Wigfield et al., 1997). In Abhängigkeit des Fachgebiets sowie der analysierten Altersstufe ergeben sich diesbezüglich jedoch differenzielle Ergebnisse (vgl. u.a. Ferdinand, 2014; Hoffmann et al., 1998; Wigfield et al., 1997). Es gibt Hinweise darauf, dass insbesondere die ‚harten‘ mathematisch-naturwissenschaftlichen Fächer von der Interessenabnahme betroffen sind (Krapp, 1996).

Für das Fach Physik wurde im Rahmen der *IPN-Interessenstudie* (vgl. Häußler & Hoffmann, 1995; Hoffmann et al., 1998; Hoffmann & Lehrke, 1986) das Fach- sowie Sachinteresse erstmalig differenziert und längsschnittlich erhoben (vgl. Kapitel 2.4). Die Stichprobe umfasste 51 Schulklassen der Schulformen Hauptschule, Realschule, Gymnasium und Gesamtschule. Die Schülerinnen und Schüler wurden einmal jährlich von der fünften bis zur zehnten Jahrgangsstufe jeweils zum Ende des Schuljahres befragt (Hoffmann & Lehrke, 1986). Die zentralen Ergebnisse der IPN-Interessenstudie lassen sich wie folgt zusammenfassen: Im fünften Jahrgang befand sich das physikalische Sachinteresse der Schülerinnen und Schüler auf einem relativ hohen Niveau, nahm im Laufe der Sekundarstufe jedoch kontinuierlich ab. Dabei zeigte sich erstens, dass im siebten Jahrgang der vergleichsweise stärkste Rückgang des physikalischen Sachinteresses zu beobachten war und zweitens, dass Jungen

über ein signifikant höheres physikalisches Sachinteresse verfügten als Mädchen. Auch für das Fachinteresse ließen sich signifikante Unterschiede zwischen Jungen und Mädchen feststellen. Die Geschlechterdifferenzen im Fachinteresse waren dabei stärker ausgeprägt als die Unterschiede im Sachinteresse. Während das physikalische Fachinteresse der Jungen zwischen der fünften und zehnten Klassenstufe relativ stabil blieb, verringerte sich das Fachinteresse der Mädchen hingegen stetig, sodass die Unterschiede zwischen Jungen und Mädchen im Zeitverlauf zunahm (Häußler & Hoffmann, 1995). Insgesamt zeigte sich für das Fachinteresse ebenfalls in der siebten Jahrgangsstufe der stärkste Rückgang. In den höheren Jahrgängen flachte die negative Entwicklungstendenz dann etwas ab (Häußler & Hoffmann, 1995; Hoffmann et al., 1998).

Wie in Kapitel 2.4 beschrieben, lag ein Fokus der IPN-Interessenstudie auf der Analyse des physikalischen Sachinteresses als dreidimensionales Konstrukt (kontextuelles, inhaltliches, tätigkeitsbezogenes Sachinteresse). Insbesondere anwendungsbezogene Kontexte stellten sich dabei sowohl für Jungen als auch für Mädchen als interessenfördernd heraus. Wurden physikalische Inhalte in Beispiele eingebettet, die sich auf alltägliche Erfahrungen sowie die Umwelt der Schülerinnen und Schüler bezogen, zeigten sich positive Effekte auf das Sachinteresse. Zudem konnten Themen/Phänomene, „über die man staunen kann und die zu einem Aha-Erlebnis führen“ (Häußler & Hoffmann, 1995, S. 113) das Interesse der Schülerinnen und Schüler steigern (zu Effekten von anwendungsbezogenen Unterrichtsformaten auf die Interessenentwicklung s.a. Kapitel 5.6). Häußler und Hoffmann (1995) schließen daraus, dass „Schülerinnen und Schüler der Sekundarstufe I [...] also im allgemeinen keine „kleinen Forscher“ [Hervorhebung im Original] [sind], die physikalische Erkenntnisse um ihrer selbst willen [...] zu erwerben suchen. [...] Mit anderen Worten: Sie interessieren sich weniger für die Physik als wissenschaftliche Disziplin, sondern vielmehr für ihre Anwendung und ihren lebenspraktischen Nutzen“ (Häußler & Hoffmann, 1995, S. 113).

Um die Interessenabnahme während der Sekundarstufe I zu erklären, wurde der Zusammenhang zwischen dem Fach- beziehungsweise Sachinteresse und weiteren Variablen (Persönlichkeitsmerkmale, soziales Umfeld, für Fachinteresse: Unterrichtsmerkmale) untersucht (Häußler & Hoffmann, 1995). Als stärkste Prädiktoren für das Sachinteresse stellten sich die Bereitschaft zur Beschäftigung mit technischen Geräten sowie die Faszination von Naturphänomenen heraus. Zur Aufklärung des Fachinteresses trug weder der erwartete Nutzen von Physik für die eigene berufliche Zukunft, noch das Geschlecht oder das Sachinteresse bei. Häußler und Hoffmann (1995) schlussfolgerten, dass es der Physikunterricht in seiner damaligen Ausrichtung trotz vorhandenem Interesse an physikalischen Themen folglich nicht schaffe, die Schülerinnen und Schüler für das Fach Physik zu begeistern. Stattdessen erwies sich das Selbstvertrauen in die eigene Leistungsfähigkeit (Fähigkeitsselbstkonzept) als der mit Abstand stärkste beeinflussende Faktor des physikalischen Fachinteresses. Als unterrichtspraktische Konsequenzen leiteten Häußler und Hoffmann (1995, S. 116) aus den Ergebnissen ab, vor allem die „Interaktionsmuster

im Klassenzimmer“ müssten verändert werden und meinen damit (a) die Einbettung der Unterrichtsinhalte in alltagsnahe Kontexte sowie (b) Maßnahmen und Unterrichtsformate, die das Fähigkeitsselbstkonzept der Schülerinnen und Schüler stärken.

Um zu untersuchen, ob die Interessenentwicklung im Fach Physik tatsächlich gegenüber anderen Schulfächern besonders negativ ausfällt, wurde zusätzlich die Genese des Fachinteresses in Biologie, Chemie, Kunst, Mathematik, Deutsch und Englisch untersucht (Hoffmann et al., 1998). Auch im Fächervergleich zeigten sich deutliche Geschlechterunterschiede. Während die Jungen das vergleichsweise höchste Interesse für das Fach Physik aufbrachten, beurteilten die Mädchen Physik als das am wenigsten interessante Fach und verloren im Zeitverlauf weiter Interesse am Fach Physik. Die beliebtesten Fächer der Mädchen waren hingegen Biologie, Englisch und Kunst. Im Längsschnitt zeigte sich für die Mädchen im Fach Deutsch eine Interessensteigerung, während das Fachinteresse in Biologie, Englisch und Kunst auf einem ähnlich hohen Niveau verblieb. Die Jungen zeigten nach Physik am meisten Interesse für das Fach Mathematik, während Deutsch und Englisch hingegen die unbeliebtesten Fächer der Jungen waren. Das Fachinteresse der Jungen verringerte sich im Untersuchungszeitraum in Englisch und verblieb in Mathematik auf einem ähnlichen Niveau, während es in Deutsch leicht anstieg (Hoffmann et al., 1998).

Ausgehend von den Ergebnissen der IPN-Interessenstudie untersuchte Daniels (2008) anhand der BIJU-Daten (*Bildungsprozesse und psychosoziale Entwicklung im Jugendalter und jungen Erwachsenenalter*; vgl. Baumert, Gruehn, Heyn, Köller & Schnabel, 2003) die Interessenentwicklung von Schülerinnen und Schülern der siebten bis zehnten Jahrgangsstufe. Drei Messzeitpunkte fanden in der siebten Jahrgangsstufe statt, der vierte Messzeitpunkt in der zehnten Jahrgangsstufe. Untersucht wurde das Fach- sowie Sachinteresse in den Domänen Biologie, Physik, Mathematik und Englisch. Jedoch wurden die beiden Skalen nicht in allen Domänen zu allen vier Messzeitpunkten eingesetzt. Als zentralen Befund stellt Daniels (2008) die signifikante Verringerung des durchschnittlichen Interesses in allen untersuchten Domänen heraus. Innerhalb der siebten Jahrgangsstufe war dabei fächerübergreifend der Interessenverlust besonders ausgeprägt, schwächte sich dann jedoch zum Ende der Sekundarstufe I in der 10. Jahrgangsstufe ab. Am stärksten vom Interessenrückgang betroffen waren die Fächer Physik, Mathematik und Biologie, weniger stark betroffen die Fächer Deutsch und Englisch. Als mögliche Erklärung für den Interessenrückgang wurde der Zusammenhang der Interessenentwicklung mit veränderten domänenspezifischen Fähigkeitsselbstkonzepten untersucht. Während anhand der BIJU-Daten eine Verringerung der Fähigkeitsselbstkonzepte in den mathematisch-naturwissenschaftlichen Fächern zu beobachten war, entwickelten sich die Fähigkeitsselbstkonzepte in den sprachlichen Fächern im Zeitverlauf positiv. Die negative Entwicklung der mathematisch-naturwissenschaftlichen Fähigkeitsselbstkonzepte verlief somit parallel zur negativen Interessenentwicklung in diesen Fächern, während sich in den sprachlichen Fächern gegenläufige Tendenzen zeigten. Die

korrelativen Zusammenhänge der beiden Konstrukte waren in allen Fächern als moderat bis hoch zu bewerten (in absteigender Reihenfolge: Mathematik, Englisch, Deutsch, Physik, Biologie). Daniels (2008) erklärt die Unterschiede in korrelativen Zusammenhängen durch verschiedene Schwierigkeitsgrade der Fächer, der Häufigkeit der eingesetzten Leistungsrückmeldung (Klassenarbeiten, Tests) sowie der eingesetzten Skalen. So wurde in Biologie und Physik lediglich das Sachinteresse in Beziehung zum Fähigkeitsselbstkonzept gesetzt. In Regressionsanalysen stellte sich heraus, dass in Physik und Mathematik der Interessenrückgang signifikant mit der Verringerung des Fähigkeitsselbstkonzepts verknüpft war. Weniger stark ausgeprägt waren die Effekte zwischen dem Fähigkeitsselbstkonzept und der Interessenentwicklung in Englisch und Biologie. Kein signifikanter Effekt des Fähigkeitsselbstkonzepts ließ sich im Fach Deutsch feststellen. Daniels (2008) schließt daraus, dass die Interessenabnahme zum Teil durch die Verringerung des jeweiligen fachspezifischen Fähigkeitsselbstkonzepts erklärt werden kann.

Die Befunde von Daniels (2008) decken sich mit den bereits berichteten Ergebnissen der IPN-Interessenstudie (vgl. Häußler & Hoffmann, 1995; Hoffmann et al., 1998; Hoffmann & Lehrke, 1986) und weisen zudem Parallelen zur Studie von Gottfried et al. (2001) auf, die die Interessenentwicklung us-amerikanischer Schülerinnen und Schüler von der Grundschulzeit (im Alter von 9 Jahren) bis zum Ende der High-School (im Alter von 17 Jahren) in den Fächern Naturwissenschaften, Mathematik, Englisch und Gesellschaftswissenschaften/Geschichte untersuchten. Analysiert wurde die auf die jeweiligen Schulfächer bezogene intrinsische Motivation (im Sinne von Lernfreude, Neugier, Ausdauer), die dem Konstrukt des Fachinteresses recht nahekommt. Mit Ausnahme der intrinsischen Motivation für Gesellschaftswissenschaften/Geschichte konnten Gottfried et al. (2001) in allen übrigen Fächern eine Verringerung der fachspezifischen intrinsischen Motivation im Schulverlauf feststellen. Der stärkste Interessenverlust war in Mathematik zu verzeichnen. Interessant ist der Befund, dass sich ab einem Alter von 16 Jahren der Negativtrend umkehrte und eine leichte Steigerung der fachspezifischen intrinsischen Motivation zu beobachten war. Ähnliche Befunde bestätigen auch Frenzel et al. (2010), die ebenfalls für das Fach Mathematik von der fünften bis zur siebten Jahrgangsstufe eine Verringerung des durchschnittlichen Fachinteresses beobachteten, im achten und neunten Jahrgang hingegen eine Stabilisierung des mathematischen Fachinteresses auf einem ähnlichen Niveau verzeichneten.

Wie bereits deutlich wurde, lag der Fokus empirischer Studien zur Interessenentwicklung bisher vornehmlich auf den auf mathematisch-naturwissenschaftlichen Fächern sowie sprachlichen Fächern (Deutsch, Englisch). Dies liegt wahrscheinlich zum einen daran, dass diese Domänen zu den Kernfächern der Schule gehören. Zum anderen könnte die Fokussierung im engen Zusammenhang zwischen dem Interesse und dem jeweiligen fachspezifischen Fähigkeitsselbstkonzept begründet liegen, das – wie im folgenden Kapitel 3 detailliert dargelegt wird – zu zwei übergeordneten bereichsspezifischen Selbstkonzeptdimensionen (verbales vs. mathematisches Fähigkeitsselbstkonzept) zugeordnet werden kann.

Eine Studie, die die Interessenentwicklung in den Fächern Musik und Sport (neben den Fächern Mathematik und Lesen (Englisch)) analysiert, stammt von Wigfield et al. (1997). Wigfield et al. (1997) untersuchten Schülerinnen und Schüler der ersten bis sechsten Jahrgangsstufe (entspricht der Grundschule) in den USA. Das Fachinteresse in den verschiedenen Domänen wurde über den subjektiven Wert des jeweiligen Faches (gemessen als Interesse, Nützlichkeit, subjektive Bedeutsamkeit/Wichtigkeit; für eine detaillierte Beschreibung der Wertkomponente im Rahmen der EVT s. Kapitel 4.1; vgl. Eccles & Wigfield, 2002) erhoben. Für das mathematische Fachinteresse zeigten sich uneindeutige Ergebnisse. Während in einer Kohorte weder eine Zu- noch Abnahme des Fachinteresses zu verzeichnen war, stieg das mathematische Fachinteresse in Kohorte zwei an, sank jedoch in Kohorte drei. Für die Domäne Lesen (Englisch) war ein Rückgang des Interesses zu verzeichnen, der jedoch im vierten Jahrgang gestoppt werden konnte. Das Fachinteresse in Sport blieb über die Zeit konstant. Die größte Interessenabnahme stellten Wigfield et al. (1997) für das Fach Musik fest. Gleichzeitig war der Rückgang des musikalischen Fähigkeitsselbstkonzepts gegenüber den anderen untersuchten Fächern am stärksten ausgeprägt. Die beiden Konstrukte wiesen außerdem korrelative Zusammenhänge auf, die sich im Zeitverlauf intensivierten.

Zusammengefasst kann festgehalten werden, dass das Interesse einer Person theoretisch als multidimensionales Konstrukt beschrieben wird, dass neben affektiven auch kognitive Merkmale beinhaltet und sowohl situationsspezifisch temporär als auch längerfristig stabil auftreten kann (Krapp, 1992a, 2018). Zudem wurde die Gegenstandsspezifität als zentrales Merkmal des Interessenkonstrukts herausgestellt, wobei sich Interessengegenstände im schulischen Kontext auf Unterrichtsfächer sowie einzelne Unterrichtsinhalte beziehen können (Hoffmann, 2002; Krapp, 1999). Eine entsprechende Unterscheidung zwischen dem Fach- und Sachinteresse wurde im Rahmen der sogenannten Kieler Interessentheorie entwickelt und anschließend empirisch untersucht (vgl. Häußler & Hoffmann, 1995; Hoffmann, 2002; Hoffmann & Lehrke, 1986). Die in diesem Kapitel berichteten Forschungsergebnisse verwiesen auf einen Interessenrückgang im Laufe der Schulzeit. Dabei zeigten sich Unterschiede zwischen der Entwicklung des Fach- und Sachinteresses (vgl. Häußler & Hoffmann, 1995), zwischen den untersuchten Domänen sowie hinsichtlich der betrachteten Altersstufen. Besonders stark vom Interessenrückgang betroffen waren die mathematisch-naturwissenschaftlichen Fächer (vgl. Daniels, 2008; Frenzel et al., 2010; Gottfried et al., 2001; Häußler & Hoffmann, 1995) sowie das Fach Musik (Wigfield et al., 1997). Die negative Entwicklung des Interesses war zudem wiederholt mit einem ebenfalls negativen Trend des Fähigkeitsselbstkonzepts assoziiert, sodass die Selbsteinschätzung der eigenen Fähigkeiten als stärkster Prädiktor der Interessenentwicklung identifiziert werden konnte (vgl. Daniels, 2008; Häußler & Hoffmann, 1995). Im nächsten Kapitel 3 wird daher detailliert auf das Konstrukt des Fähigkeitsselbstkonzepts eingegangen.

3 Fähigkeitsselbstkonzept

Das schulische (akademische) Selbstkonzept, mit seinen fachspezifischen Subdimensionen, stellt eines der in der pädagogisch-psychologischen Forschung am meisten untersuchten Konstrukte dar. Da das Selbstkonzept als einer der zentralen Indikatoren für schulisches Lernverhalten gilt, wird dessen Förderung als wichtiges pädagogisches Ziel benannt. Im schulischen Kontext nimmt das akademische Selbstkonzept eine Mediatorrolle zwischen bisherigen Leistungserfahrungen und dem weiteren Lernverhalten ein (Möller & Trautwein, 2020). Insbesondere der positive wechselseitige Zusammenhang zwischen dem (Fähigkeits-)Selbstkonzept und der Leistung innerhalb einer Domäne konnte wiederholt bestätigt werden (vgl. u.a. Becker, Willems & Bos, 2014; Lauermaun, Meißner & Steinmayr, 2020; Marsh, Kong & Hau, 2000; Marsh, Lüdtke, Trautwein & Morin, 2009; Marsh et al., 2014; Möller, Pohlmann, Köller & Marsh, 2009; Nagengast & Marsh, 2012). Darüber hinaus steht das fachbezogene Selbstkonzept in einem engen Zusammenhang mit der Interessenstruktur von Schülerinnen und Schülern (vgl. u.a. Arens et al., 2019; Lauermaun, Tsai & Eccles, 2017; Spinath & Steinmayr, 2008; Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006; Watt, 2004; für eine detaillierte Beschreibung s. Kapitel 4.2).

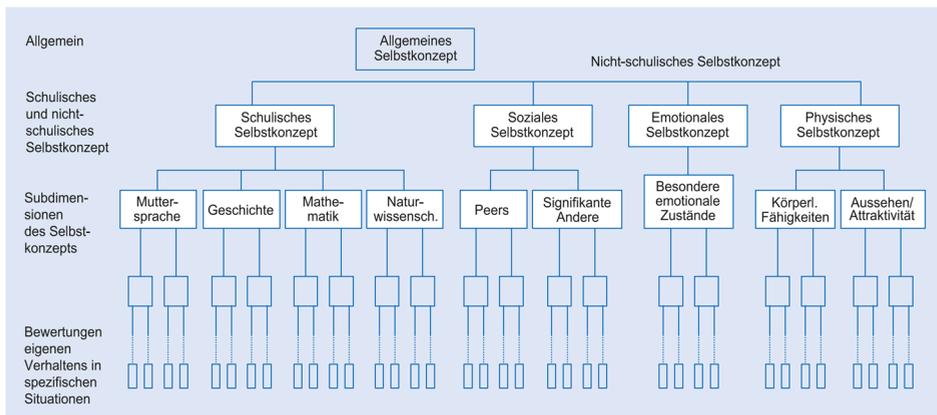
Im folgenden Kapitel wird das Konstrukt zunächst theoretisch definiert (Kapitel 3.1) und danach die multidimensionale Struktur des Selbstkonzepts beschrieben

(Kapitel 3.2). Anschließend werden mögliche Determinanten der Selbstkonzeptgenese aufgezeigt und dabei die besondere Rolle des internen und externen Vergleichs hervorgehoben (Kapitel 3.3). In Kapitel 3.4 werden zentrale nationale und internationale Forschungsergebnisse zu Entwicklungsprozessen des akademischen Selbstkonzepts im Schulverlauf zusammengefasst, wobei ein Fokus auf die in dieser Arbeit untersuchte Altersstufe (Beginn der Sekundarstufe I) gelegt wird.

3.1 Theoretische Konzeption

Ganz allgemein wird in der pädagogisch-psychologischen Forschung das Selbstkonzept einer Person als Selbstwahrnehmung, Selbstbild oder Selbstwert verstanden und bezeichnet folglich Einschätzungen, Einstellungen oder Bewertungen zur eigenen Person (Möller & Trautwein, 2020). Nach Shavelson, Hubner und Stanton (1976) umfasst das allgemeine Selbstkonzept verschiedene Subdimensionen, wie das akademische (schulische), soziale, physische und emotionale Selbstkonzept (s. Abbildung 3.01).

Abbildung 3.01: Multidimensionale und hierarchische Struktur des Selbstkonzepts (nach Shavelson et al., 1976)



Möller & Trautwein, 2020, Abbildung 8.2

Im schulischen Kontext ist vor allem das akademische Selbstkonzept von Bedeutung, das synonym auch unter den Begriffen ‚Fähigkeitsselbstkonzept‘ sowie ‚Selbstkonzept der Begabung‘ Anwendung findet. Gemeint ist damit die Bewertung der eigenen Fähigkeiten sowie des empfundenen Leistungsniveaus. Die aktuelle pädagogisch-psychologische Forschung spricht demnach vom akademischen Selbstkonzept (bzw. Fähigkeitsselbstkonzept) als „kognitive Repräsentation eigener Fähigkeiten und/oder Begabungen“ (Möller & Trautwein, 2020, S. 193). Da sich Fä-

higkeitsselbstkonzepte immer auf spezifische Bereiche (z.B. Unterrichtsfächer) beziehen, findet sich in der Forschungsliteratur häufig auch die Bezeichnungen ‚fachspezifisches Selbstkonzept‘ (Möller & Trautwein, 2020). In der vorliegenden Arbeit werden die Begriffe Fähigkeitsselbstkonzept und fachspezifisches Selbstkonzept als Subdimensionen des akademischen Selbstkonzepts synonym verwendet.

Aussagen wie „Ich gehöre im naturwissenschaftlichen Unterricht zu den guten Schülerinnen und Schülern.“ oder „Ich weiß im naturwissenschaftlichen Unterricht die Antwort auf eine Frage schneller als die Anderen.“ können als Beispiele für die Subdimension des naturwissenschaftlichen Selbstkonzepts (bzw. synonym des naturwissenschaftlichen Fähigkeitsselbstkonzepts) herangezogen werden (DISK-Gitter; Rost & Sparfeldt, 2002). Aus den Beispielitems wird zudem deutlich, dass sich das Selbstkonzept einer Person stets durch und mit dem sozialen Umfeld ausbildet. Historisch sind entsprechende Überlegungen zu Beginn des 20. Jahrhunderts im Kontext des symbolischen Interaktionismus zu verorten. So formulierte beispielsweise bereits Mead (1934) die These, dass sowohl Einzelpersonen aber auch soziale Gruppen mit ihren Einstellungen das Selbstkonzept einer Person formen (vgl. Möller & Trautwein, 2020). Fragen der Selbstkonzeptausbildung wurden ebenfalls in der Entwicklungspsychologie bearbeitet. Zu nennen sind hier die Arbeiten von Piaget (1960) sowie Harter (1998, 1999). Harter (1998, 1999) beschrieb, wie sich das in „der Kindheit stark positiv verzerrt[e] [Selbstkonzept] [...] durch die allmähliche Integration auch negativer Informationen über eigene Fähigkeiten und Eigenschaften“ vom Kindes- bis zum Jugendalter zunehmend realistischer gestaltet und differenziert (Möller & Trautwein, 2020, S. 192). Entsprechende Annahmen, der (a) multidimensionalen Struktur des Selbstkonzepts (vgl. Shavelson et al., 1976) und (b) der sozialen Umwelt als wichtige Einflussgröße der Selbstkonzeptgenese im symbolischen Interaktionismus (vgl. Mead, 1934) sowie (c) entwicklungspsychologische Annahmen der zunehmenden Ausdifferenzierung des Selbstkonzepts (vgl. Harter, 1998, 1999; Piaget, 1960) finden sich auch in der aktuellen Selbstkonzeptforschung wieder und werden im Folgenden näher erläutert.

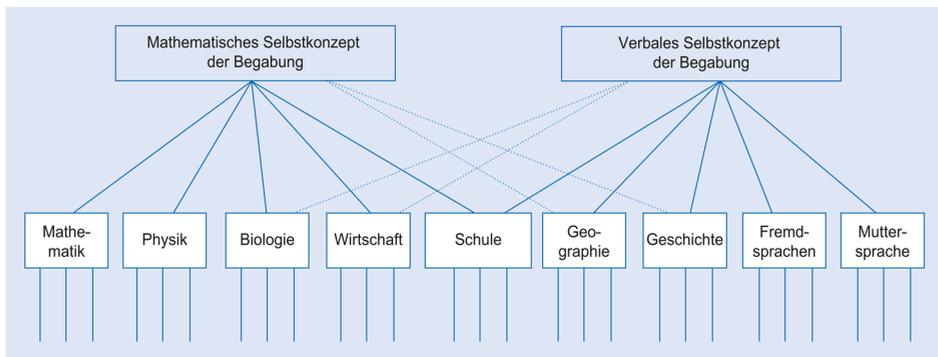
3.2 Struktur des Selbstkonzepts

Das von Shavelson et al. (1976) entwickelte Modell wird häufig als Grundlage heutiger pädagogisch-psychologischer Forschung zur Struktur des Selbstkonzepts herangezogen. Für die im sogenannten ‚Shavelson-Modell‘ postulierten Eigenschaften des Selbstkonzepts sammelten Shavelson et al. (1976) Ergebnisse bisheriger Forschungen und subsummierten diese zu zwei zentralen Hypothesen: Das Selbstkonzept einer Person ist *hierarchisch* sowie *multidimensional* aufgebaut. Hierarchisch meint hier die Anordnung verschiedener Selbstkonzeptfacetten unter einem generellen (schulischen) Selbstkonzept. Die verschiedenen Subdimensionen des Selbstkonzepts bilden sich erst im Laufe des Heranwachsens aus. Im Kleinkindalter ist das

Selbstkonzept eher global, undifferenziert und situationsspezifisch angelegt und differenziert sich dann mit steigendem Alter sowie zunehmender Erfahrungen weiter aus (Multidimensionalität; Shavelson et al., 1976).

Während die These des hierarchischen Aufbaus des Selbstkonzepts mittlerweile revidiert wurde, da sich empirisch kein einzelner Generalfaktor des schulischen Selbstkonzepts bestätigen ließ (Marsh, Bryne & Shavelson, 1988), ist die Annahme der Multidimensionalität bestehen geblieben. Auf globaler Ebene wird stattdessen heute von zwei übergeordneten (bereichsspezifischen) Selbstkonzepten ausgegangen unter denen sich die Subdimensionen anordnen (s. Abbildung 3.02): Das *verbale Selbstkonzept* beinhaltet die Fähigkeitsselbsteinschätzung zum sprachlichen Bereich (Muttersprache, Fremdsprache sowie ‚sprachlastige‘ Fächer wie Geschichte) und das *mathematische Selbstkonzept* subsumiert die mathematisch-naturwissenschaftlichen Fähigkeitsselbstkonzepte (Mathematik, Physik, Chemie, Biologie) sowie ‚zahlenlastige‘ Fächer (Wirtschaft; Marsh et al., 1988).

Abbildung 3.02: Struktur des schulischen Selbstkonzepts im revidierten Modell (nach Marsh et al., 1988)



Möller & Trautwein, 2020, Abbildung 8.3

Empirisch konnte die Annahme der zwei übergeordneten bereichsspezifischen Selbstkonzeptfacetten unter anderem im Rahmen der groß angelegten Metastudie von Möller et al. (2009) untermauert werden. In den von Möller et al. (2009) analysierten 69 Datensätzen fanden die Autoren hohe Korrelationen zwischen der verbalen und der mathematischen Leistung, während das verbale und das mathematische Selbstkonzept nahezu unkorreliert war. Darüber hinaus konnten in Pfadanalysen hohe Zusammenhänge zwischen der Leistung und dem Fähigkeitsselbstkonzept innerhalb einer Domäne gezeigt werden, während die domänenübergreifenden Effekte negativ ausfielen (Zusammenhang verbales Selbstkonzept und mathematische Leistung und umgekehrt). Entsprechend wird davon ausgegangen, dass sich für das schulische Selbstkonzept fachspezifische Facetten abbilden lassen (Marsh, Martin, Yeung & Craven, 2017). Bereits für Kinder im Grundschulalter (6 bis 11 Jahre)

konnten Marsh, Barnes, Cairns und Tidman (1984) verschiedene Selbstkonzeptdimensionen empirisch bestätigen, wiesen aber auch darauf hin, dass sich die Facetten mit zunehmendem Alter weiter ausdifferenzieren. Für die Jahrgangsstufen fünf und sechs konnte Marsh (1990) anhand von Faktorenanalysen 13 fachspezifische Selbstkonzeptdimensionen unterscheiden und betont daher explizit, dass „researchers interested in self-concepts in particular subjects are advised to use self-concept scales specific to those subject areas [...]“ (Marsh, 1990, S. 623).

3.3 Die Rolle des Vergleichs für die Selbstkonzeptgenese

Die Entwicklung des Selbstkonzepts einer Person erfolgt insbesondere über den Vergleich, entweder mit dem sozialen Umfeld als Bezugsnorm (Eltern, Lehrkräfte, Freunde sowie Gruppen, z.B. die Schulklasse), aber auch über den Abgleich mit eigenen vorangegangenen Erfahrungen („Internal/External-Frame-of-Reference-Modell“ (I/E-Modell); Marsh, 1986). Unterschieden werden diesbezüglich (externe) soziale sowie (interne) temporale, dimensionale und kriteriale Vergleiche (bzw. Vergleichsinformationen). Vergleichen Schülerinnen und Schüler beispielsweise die eigene Note in einer Klassenarbeit mit den Noten ihrer Klassenkameradinnen und -kameraden, so findet ein (externer) sozialer Vergleich statt (Marsh et al., 2017). Von temporalen Vergleichen wird gesprochen, wenn die aktuelle Leistung mit eigenen vorangegangenen Leistungen abgeglichen wird (Möller & Trautwein, 2020). Dimensionale Vergleiche meinen den Vergleich der eigenen Leistung zwischen verschiedenen Schulfächern (Möller & Köller, 2004) und in kriterialen Vergleichen werden extern festgelegte Kriterien für den Abgleich der eigenen Leistung herangezogen (z.B. die Note eins steht für eine sehr gute Leistung; Möller & Trautwein, 2020).

Ob durch die Vergleiche das jeweilige fachspezifische Selbstkonzept gestärkt oder geschwächt wird, hängt vor allem davon ab, ob die gute/schlechte Leistung subjektiv mit internen oder externen Faktoren begründet wird und wie veränderbar die Ursachen wahrgenommen werden. Die Effekte (positiv wie negativ) auf das Fähigkeitsselbstkonzept sind stärker, wenn eine gute/schlechte Leistung aus der eigenen Begabung abgeleitet wird. Veränderbare Kausalattributionen (bspw. „ich habe nicht genug gelernt/mich nicht genug angestrengt“) können hingegen gerade bei schlechter Leistung das fachspezifische Selbstkonzept vor einem allzu starken Einbruch bewahren, da hier Handlungsspielraum für zukünftige Leistungserfahrungen besteht („beim nächsten Mal fange ich früher an zu lernen“; Möller & Trautwein, 2020).

Zwei der bekanntesten Effekte des (externen) sozialen Vergleichs auf die Selbstkonzeptgenese stellen der *Big-Fish-Little-Pond-Effekt* (Marsh, Trautwein, Lüdtke & Köller, 2008) sowie der *Reflected-Glory-Effekt* (Marsh et al., 2000) dar. Die zwei Effekte werden im Folgenden kurz vorgestellt.

Big-Fish-Little-Pond-Effekt

Als Big-Fish-Little-Pond-Effekt (BFLPE; dt. Bezugsgruppeneffekt; Marsh et al., 2008) wird der Effekt des sozialen Vergleichs mit der Bezugsgruppe (z.B. Schulklasse) auf das individuelle fachspezifische Selbstkonzept beschrieben (externer Bezugsrahmen). Besuchen Schülerinnen und Schüler eine Klasse mit durchschnittlich hohen Leistungen, so sind negative soziale Vergleiche wahrscheinlicher (die eigene Leistung ist schlechter als die Durchschnittsleistung der Klasse), worunter das eigene fachspezifische Selbstkonzept leidet. Besuchen Schülerinnen und Schüler hingegen eine Klasse mit niedrigem Leistungsniveau („little pond“), so ist die Wahrscheinlichkeit für einen positiven sozialen Vergleich höher, wovon das eigene fachspezifische Selbstkonzept profitiert („big fish“) (Marsh et al., 2008; vgl. Arens, 2017). Der BFLPE liefert folglich Antworten auf die Frage, warum sich Schülerinnen und Schüler bei gleicher Leistung hinsichtlich ihres Fähigkeitsselbstkonzepts unterscheiden können (Marsh et al., 2017).

Empirisch kann der BFLPE mittels Mehrebenenanalysen untersucht werden. Auf Schülerebene hängen die fachspezifische Leistung und das entsprechende fachspezifische Selbstkonzept positiv zusammen, während auf der Klassenebene das individuelle Fähigkeitsselbstkonzept in Abhängigkeit eines hohen durchschnittlichen Leistungsniveaus der Klasse niedriger ist (Arens, 2017). Empirische Belege des BFLPE finden sich in zahlreichen Forschungsarbeiten für unterschiedliche Länder und kulturelle Kontexte (Nagengast & Marsh, 2012; Seaton, Marsh & Craven, 2009), verschiedene Domänen (z.B. Mathematik: Köller, Trautwein, Lüdtke & Baumert, 2006; Trautwein et al., 2006; Seaton et al., 2009; Naturwissenschaften: Nagengast & Marsh, 2012) sowie in Zusammenhang mit Geschlechtereffekten (Marsh, Trautwein, Lüdtke, Baumert & Köller, 2007). Trautwein et al. (2006) konnten darüber hinaus anhand von PISA-Daten für Neuntklässlerinnen und Neuntklässler sowie Köller et al. (2006) anhand von BIJU-Daten (vgl. Baumert et al., 2003) im zehnten und zwölften Jahrgang zeigen, dass sich Bezugsgruppeneffekte nicht nur für das Fähigkeitsselbstkonzept in Mathematik, sondern auch für das mathematische Fachinteresse beobachten lassen. Unter Kontrolle der individuellen Testleistung zeigten Schülerinnen und Schüler, die eine Klasse mit einem hohen Leistungsniveau besuchten, ein niedrigeres mathematisches Interesse als Schülerinnen und Schüler in leistungsschwachen Klassen (Köller et al., 2006; Trautwein et al., 2006). Trautwein et al. (2006) schlussfolgern daher, dass dem Fähigkeitsselbstkonzept eine Mediatorwirkung zukommt: Das niedrige Leistungsniveau der Klasse steht mit einem höheren Fähigkeitsselbstkonzept in Verbindung, was entsprechend einen positiven Effekt auf das mathematische Fachinteresse nach sich zieht.

Reflected-Glory-Effekt

Demgegenüber nimmt der Reflected-Glory-Effekt (dt. Assimilationseffekt; Marsh et al., 2000) an, dass Schülerinnen und Schüler in akademisch positiv-selektiven Gruppen, in denen ein hohes Leistungsniveau besteht, über ein höheres akademisches Selbstkonzept verfügen, da sie sich durch die eigene Teilnahme privilegiert fühlen („e.g. if I am good enough to be in this selective school with all these other very smart students, than I must be very smart“; Marsh et al., 2000, S. 338). Der Reflected-Glory-Effekt kann insbesondere nach Prozessen der Leistungsdifferenzierung beobachtet werden (z.B. am Übergang in die weiterführende Schule; Arens, 2017). So konnten Arens und Watermann (2015) für besonders leistungsstarke Schülerinnen und Schüler, die für eine Hochbegabtschule ausgewählt wurden, feststellen, dass der Reflected-Glory-Effekt höher war als der negative BFLPE. Marsh et al. (2000) sowie Köller et al. (2006) merken diesbezüglich an, dass sowohl der Bezugsgruppeneffekt (BFLPE) als auch der Assimilationseffekt (Reflected-Glory-Effekt) gleichzeitig auftreten (können) und sich zum Teil gegenseitig ausgleichen. Der empirisch häufiger zu beobachtende negative BFLPE ist somit als ‚Nettoeffekt‘ der beiden entgegengesetzt wirksamen Prozesse zu interpretieren (Marsh et al., 2000).

3.4 Empirische Befunde zur Entwicklung des Selbstkonzepts in der Sekundarstufe I

Die Entwicklung der schulischen Selbstkonzeptfacetten stand und steht im Fokus zahlreicher nationaler und internationaler Studien. Weitgehend übereinstimmend wird dabei eine Verringerung der mittleren fachspezifischen Fähigkeitsselbstkonzepte im Schulverlauf berichtet und zwar unabhängig von der untersuchten Altersstufe (vgl. u.a. Dickhäuser et al., 2017; Jacobs et al., 2002; Lazarides et al., 2020; Marsh & Ayotte, 2003; Marsh et al., 1984; Scherrer & Preckel, 2019; Spinath & Spinath, 2005; Watt, 2004). So untermauern beispielsweise Scherrer und Preckel (2019) in einer groß angelegten, länderübergreifenden Metaanalyse entsprechende Befunde negativer Entwicklungen des mathematischen und verbalen Selbstkonzepts zwischen erster und dreizehnter Jahrgangsstufe. Da sich am Übergang zwischen der Primar- und Sekundarstufe die Lernumgebung und damit auch die sozialen Vergleichsmaßstäbe als Faktor der Selbstkonzeptgenese ändern (Arens, 2017; s. Kapitel 3.3), stellt der Übergang in die weiterführende Schule ein besonders spannendes Forschungsfeld dar. Im Vergleich zur Grundschule gelten in der Sekundarstufe strengere Bewertungskriterien und das Leistungsniveau wird angehoben. Gleichzeitig ist der Schulübertritt durch verstärkte Kontroll- und Disziplinierungsmaßnahmen sowie weniger persönlichen Beziehungen zwischen den Lehrkräften und den Schülerinnen und Schülern geprägt (Eccles, Midgley et al., 1993; vgl. Rieger

et al., 2019). Entsprechend berichteten Wigfield und Eccles (1994) für us-amerikanische Schülerinnen und Schüler beim Übergang in die Junior High School einen Einbruch des Fähigkeitsselbstkonzepts in Mathematik, Englisch und Sport. Anders als in Deutschland findet in den USA der Übergang in die weiterführende Schule erst zwischen sechster und siebter Jahrgangsstufe statt, sodass sich der Schulwechsel (mit den beschriebenen externen Veränderungen) und das Einsetzen der Pubertät (und damit verbundenen entwicklungsbedingten Veränderungen; s. Kapitel 2.5) zeitlich überschneiden. Arens, Yeung, Craven, Watermann und Hasselhorn (2013) griffen diesen Forschungsansatz auf und analysierten die Effekte des Schulübergangs auf das fachspezifische Fähigkeitsselbstkonzept für deutsche Schülerinnen und Schüler. Um den negativen BFLPE möglichst zu kontrollieren, fokussierten sich Arens et al. (2013) auf Schülerinnen und Schüler, die auf eine Gesamtschule wechseln, da angenommen wurde, dass dort heterogenere Lernumgebungen zu finden sind als beispielsweise auf Gymnasien und somit Schulformeffekte weniger stark ins Gewicht fallen. Arens et al. (2013) konnten auch für einen früheren Übergang nach der vierten Klasse negative Entwicklungen des Fähigkeitsselbstkonzepts in den Domänen Deutsch und Mathematik bestätigen.

4 Erwartungs-Wert Theorie

Die Expectancy-Value Theory (EVT; dt. Erwartungs-Wert Theorie; Eccles et al., 1983; Eccles & Wigfield, 2002; Wigfield & Eccles, 2000) ist in der pädagogisch-psychologischen Forschung aktuell eines der am weitesten verbreiteten Modelle (Arens, 2021; Gaspard, Hasselhorn et al., 2019) und wurde bereits vielfach angewendet, um Lernverhalten, Lernerfolg, Motivation oder schulische Wahlprozesse zu untersuchen (vgl. u.a. Arens, 2021; Gaspard, Laueremann et al., 2019; Guo et al., 2017; Köller et al., 2019; Lowe & Coy, 2016; McCormick & McPherson, 2007; Nagy et al., 2006; Nonte et al., 2021). In der EVT werden zwei zentrale Komponenten unterschieden: Die Erfolgserwartung (Erwartungskomponente) und die Wertüberzeugung (Wertkomponente) werden als separate aber positiv miteinander verknüpfte Faktoren betrachtet (Eccles & Wigfield, 2002).

Nach einer kurzen historischen Einordnung werden im Folgenden zunächst die theoretischen Annahmen der EVT zusammengefasst und die Erwartungs- und Wertkomponente inhaltlich definiert (Kapitel 4.1). Anschließend werden Forschungsbefunde zum Zusammenhang der beiden Komponenten beschrieben, die vor allem auch die Frage nach der Veränderbarkeit der Komponenten am Übergang in die Sekundarstufe I adressieren (Kapitel 4.2).

4.1 Das Modell – theoretische Annahmen

Historisch hat die Verwendung der Konstrukte der Erwartungs- und Wertdimension in der (pädagogisch-)psychologischen Forschung bereits eine lange Tradition. Die ersten theoretischen Definitionen der Erwartungs- und Wertkomponente finden sich bereits bei Lewin (1938) sowie Tolman (1932) (vgl. Wigfield et al., 2016). Das erste Erwartungs-Wert Modell geht auf Atkinson (1957) zurück und wurde zur Erklärung von leistungsbezogenem Verhalten verwendet (z.B. Streben nach Erfolg, aufgabenbezogenes Wahlverhalten, Beharrlichkeit/Ausdauer). Bereits Atkinson (1957) formulierte die These, dass leistungsbezogenes Verhalten auf Leistungsmotivation, Erfolgserwartung und individuellen Werten beruht. Allerdings definierte Atkinson (1957) die Erwartungs- und Wertkomponente vor allem situations- und aufgabenspezifisch. Zudem wurde angenommen, dass sich die beiden Komponenten invers zueinander verhalten. Aufgaben mit einem hohen persönlichen Wert würden demnach als besonders schwierig wahrgenommen (vgl. Wigfield et al., 2016).

Die moderne Erwartungs-Wert Theorie (Eccles et al., 1983; Eccles & Wigfield, 2002; Wigfield & Eccles, 2000; s. Abbildung 4.02) basiert auf der Grundlage von Atkinson (1957), wurde aber durch eine erweiterte Beschreibung der Erwartungs- und Wertkomponente sowie durch vielfältige psychologische, soziale und kulturelle Faktoren im Beziehungsgefüge der Komponenten ergänzt. Zudem stand die Frage nach der Entwicklung der Erwartungs- und Wertkomponente im Fokus der Forschungsarbeiten (Wigfield et al., 2016). Im modernen Erwartungs-Wert Modell wird angenommen, dass die Art und Weise, wie Kinder das sie umgebende kulturelle Milieu, die Einstellungen und das Verhalten der Sozialisationsinstanzen (v.a. der Eltern und Lehrkräfte) sowie individuelle leistungsbezogene Erfahrungen wahrnehmen, die Erwartungs- und Wertkomponente formen (Wigfield, Rosenzweig & Eccles, 2017; Wigfield et al., 2016; s. Abbildung 4.02).

Erwartungskomponente

Eccles und Wigfield (2002, S. 119) definieren die Erwartungskomponente als „expectancies for success as individuals' beliefs about how well they will do on upcoming tasks either in the immediate or longer-term future“. Aus der Definition wird deutlich, dass damit konzeptionell sowohl die Selbsteinschätzung der eigenen Fähigkeiten als auch die Erfolgserwartung, eine bevorstehenden Aufgabe gut zu meistern, gemeint ist (vgl. Trautwein et al., 2012).

Empirisch zeigten sich hohe korrelative Zusammenhänge zwischen der Kompetenz- und Erfolgsüberzeugung, sodass Eccles und Wigfield (2002) schlussfolgern, „children and adolescents do not distinguish between these two different levels of beliefs. [...] in real-world achievement situations they are highly related and empirically indistinguishable“ (Eccles & Wigfield, 2002, S. 119).

Infolgedessen hat sich in der empirischen Forschung die Operationalisierung der Erwartungskomponente über ein gemeinsames Konstrukt – das Fähigkeits-selbstkonzept – etabliert (vgl. u.a. Arens et al., 2019; Guo et al., 2017; Trautwein et al., 2012).

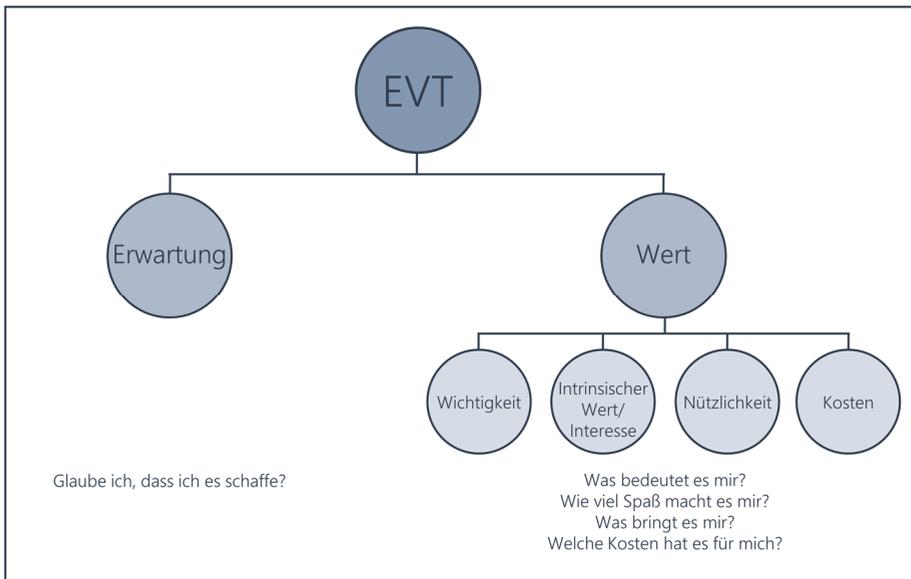
Wertkomponente

Die Wertkomponente beinhaltet nach der EVT vier Subfacetten. Unterschieden werden (a) die subjektive Bedeutsamkeit/Wichtigkeit (*attainment value*), (b) der subjektive Wert beziehungsweise das Interesse (*intrinsic value*), (c) die Nützlichkeit (*utility value*) sowie (d) eventuelle Kosten, verstanden als erforderlicher Aufwand oder Opportunitätskosten (*relative costs*; Eccles & Wigfield, 2002; s. zur Veranschaulichung Abbildung 4.01). Insbesondere von den beiden ersten Subfacetten (Wichtigkeit und Interesse) wird angenommen, dass diese in einem engen Zusammenhang stehen (Trautwein et al. (2012) berichten beispielsweise eine korrelative Beziehung von $r = .97$ zwischen den Facetten). Wigfield und Cambria (2010) weisen zudem darauf hin, dass zwischen der Dimension *attainment value*, vor allem aber der Dimension *intrinsic value* inhaltliche und theoretische Überschneidungen zum Interessenkonstrukt bestehen (Wigfield et al., 2016; vgl. Schiefele, 2001). Im Sinne der Person-Gegenstands-Relation (Krapp, 1992a) kann der intrinsische Wert der Wertkomponente mit der *gefühlsbezogenen Valenz* des Interesses verglichen werden (vgl. Kapitel 2.2; Krapp, 1998, 1999; Pekrun, 1988; Schiefele, 1996). Die Ausübung einer Tätigkeit mit hohem intrinsischen Wert wird folglich durch positive, emotionale Erfahrungen (Freude, Spaß, Aktiv- oder Beteiligtsein) begleitet. Die Wichtigkeit (*attainment value*) weist hingegen Überschneidungen zur *wertbezogenen Valenz*, das heißt der subjektiven Bedeutsamkeit eines Interessengegenstandes, auf (Krapp, 1999; Schiefele, 1996; vgl. Gaspard, Nagengast & Trautwein, 2019).

Wie bereits in Kapitel 2.2 beschrieben, können sich nach der Person-Gegenstands-Relation Interessengegenstände im Schulkontext sowohl auf Schulfächer als auch spezifische Themen beziehen (Krapp, 1999). Auch im Rahmen EVT-basierter Forschung wurde die Frage, ob der subjektive Wert das Interesse einer Person an einem Schulfach (Fachinteresse) oder das Interesse an einer Thematik allgemein (Sachinteresse) meint, bisher uneinheitlich beantwortet (vgl. Kapitel 2.4). So finden sich sowohl Forschungsarbeiten, die (unter Rückgriff auf die EVT) die Dimension *intrinsic value* über Skalen zum Fachinteresse operationalisieren (vgl. u.a. Arens et al., 2019; Gaspard, Lauer mann et al., 2019; Trautwein et al., 2012) aber auch Studien, die den subjektiven Wert über Skalen zum Sachinteresse messen (vgl. u.a. Fritzsche et al., 2011; Fritzsche, Pfeiffer & Kröner, 2015; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005; Wigfield et al., 1997). Beide Lesarten sind nach der Definition von Eccles und Wigfield (2002) durchaus legitim, denn dort heißt es: „Intrinsic value is the enjoyment the individual gets from performing the activity or the subjective interest the individual has in the subject“ (Eccles & Wigfield, 2002, S. 120). Allerdings sind je nach angesprochenem Bereich (Interesse am Unterrichtsfach vs.

Interesse am Themenbereich) unterschiedliche Ergebnisse zu erwarten, da Fach- und Sachinteressen nicht notwendigerweise gleichzeitig auftreten oder sich bedingen (vgl. Häußler & Hoffmann, 1995; Heß, 2017b). Da bisherige Forschungsarbeiten jedoch entweder die eine oder die andere Interessenfacette verwendeten, fehlt es aktuell an einer Gegenüberstellung möglicher Ergebnisunterschiede. Diesem Forschungsdesiderat wird in der vorliegenden Arbeit Rechnung getragen.

Abbildung 4.01: Schaubild der Erwartungs- und Wertkomponente (nach Eccles & Wigfield, 2002)



Eigene Darstellung; zur Beschreibung der Komponenten s. Gaspard, Hasselhorn et al. (2019).

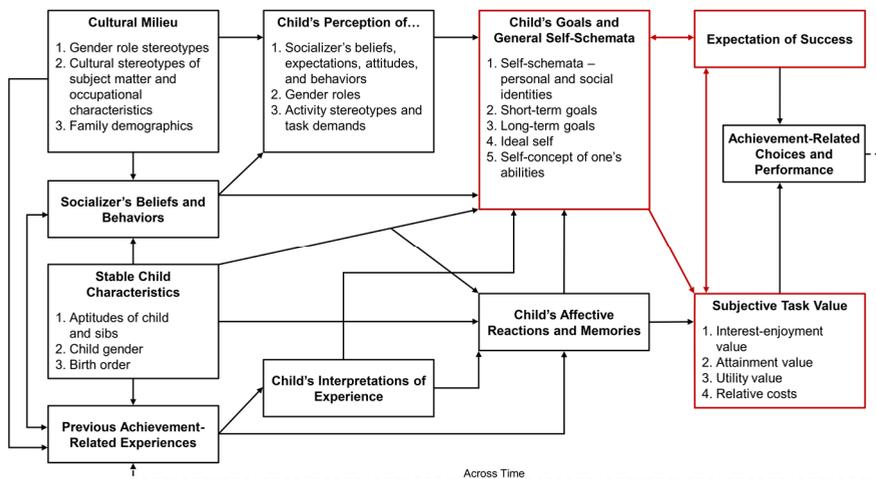
Aus der Definition zum *intrinsic value* (vgl. Eccles & Wigfield, 2002) wird zudem deutlich, dass die Wertkomponente explizit fach- beziehungsweise themenspezifisch konzeptualisiert ist. Theoriekonform konnte empirisch gezeigt werden, dass Wertüberzeugungen verschiedener Fächer nur marginal beziehungsweise sogar negativ miteinander in Beziehung stehen und demnach voneinander zu unterscheiden und entsprechend getrennt zu operationalisieren sind (Bong, 2001; Gaspard, Nagengast et al., 2019; Trautwein et al., 2012). Dies trifft ebenfalls für das Fähigkeitsselbstkonzept (Erwartungskomponente) in unterschiedlichen Domänen zu (vgl. Marsh, 1990; Trautwein et al., 2012; s.a. Kapitel 3.2).

4.2 Empirische Befunde zum Zusammenhang der Wert- und Erwartungskomponente

Wie im Schema des EVT-Modells deutlich wird (s. Abbildung 4.02), stehen die Erfolgserwartung und die Wertkomponente theoretisch in einem bidirektionalen Zusammenhang zueinander – sie beeinflussen sich also gegenseitig positiv („expectancies and values are assumed to be positively related to each other“; Eccles & Wigfield, 2002, S. 118). Die Selbsteinschätzung der eigenen Fähigkeiten (Fähigkeitsselbstkonzept) ist dagegen als gerichtete Beziehung zur Wertkomponente modelliert. Da sich aber die Kompetenzüberzeugung und die Erfolgserwartung empirisch nicht trennen lassen (s. Kapitel 4.1; vgl. Eccles & Wigfield, 2002), ergibt sich hieraus die Frage nach der tatsächlichen Beziehung zwischen der Erwartungs- und Wertkomponente.

Wiederholt konnte der theoretisch angenommene positive Zusammenhang zwischen der Erwartungs- und Wertkomponente in querschnittlichen Analysen bestätigt werden (vgl. u.a. Dinkelmann & Buff, 2016; Denissen, Zarrett & Eccles, 2007; Fredricks & Eccles, 2002; Trautwein et al., 2012; Wigfield et al., 1997). So fanden Dinkelmann und Buff (2016) in einem schweizer Sample von Schülerinnen und Schülern der sechsten Jahrgangsstufe hohe, signifikant positive Korrelationen zwischen dem Fähigkeitsselbstkonzept und dem subjektiven Wert im Fach Mathematik. Am Ende der Sekundarstufe II zeigten sich für die Domänen Mathematik und Englisch ebenfalls hohe korrelative Zusammenhänge zwischen dem Fähigkeitsselbstkonzept und der Bedeutsamkeit/Wichtigkeit (als zweite Subdimension der Wertkomponente; vgl. Trautwein et al., 2012).

Abbildung 4.02: Erwartungs-Wert Modell (nach Eccles & Wigfield, 2002)



Eccles & Wigfield, 2002, Abbildung 1, modifiziert

Fredricks und Eccles (2002) konnten darüber hinaus in einer breit angelegten Studie mit us-amerikanischen Schülerinnen und Schülern des ersten bis zwölften Jahrgangs zeigen, dass der Zusammenhang zwischen dem Fähigkeitsselbstkonzept in Mathematik sowie Sport und der Wertkomponente (gemessen als Interesse sowie Wichtigkeit) in Mathematik und Sport im Verlauf der Schulzeit stärker wird. Aber auch hier wurden die Korrelationsanalysen nach Jahrgängen getrennt (d.h. querschnittlich) berechnet. Ähnliche Ergebnisse zeigten sich auch bei Wigfield et al. (1997), die ebenfalls zunehmend positive Zusammenhänge der Erwartungs- und Wertkomponente zwischen der zweiten und sechsten Jahrgangsstufe in den Domänen Mathematik, Lesen, Musik und Sport beobachteten. Lediglich in der ersten Jahrgangsstufe fanden Wigfield et al. (1997) keinen wechselseitig signifikanten Zusammenhang zwischen dem Fähigkeitsselbstkonzept und Wertkomponente (gemessen als Kombination aus Wichtigkeit und Sachinteresse), woraus die Autoren schließen, „it appears to be the second-grade year when competence beliefs and values become more synchronous for many children“ (Wigfield et al., 1997, S. 465). Auch aktuellere Studien (vgl. Denissen et al., 2007; Fredricks & Eccles, 2002) verweisen darauf, dass eine Intensivierung der Wert-Erwartungs-Beziehung nach der Grundschulzeit stattfindet. Dieser Befund wird auch als „Spezialisierungsprozess“ beschrieben, das heißt Kinder und Jugendliche entwickeln in bestimmten Bereichen, für die sie sich interessieren, auch ein entsprechend fachspezifisches Fähigkeitsselbstkonzept und umgekehrt (Arens et al., 2019, S. 666; vgl. Kapitel 2.5). Diese intraindividuelle Differenzierung führt zu einem Rückgang der durchschnittlichen Ausprägung der Erwartungs- und Wertkomponente (bspw. einer Klasse), da Schülerinnen und Schüler nur in wenigen Bereichen ein hohes Fähigkeitsselbstkonzept beziehungsweise Interesse ausbilden (Gaspard, Lauer mann et al., 2019). Nach Wigfield (1994) lassen sich bereits im Grundschulalter die Erwartungs- und Wertkomponente als separate Faktoren empirisch nachweisen. Aber wie steht es nun um die längsschnittliche Beziehung der beiden Komponenten zueinander?

Wie bereits angemerkt, wird die Frage nach der Direktionalität des Zusammenhangs zwischen der Erwartungs- und Wertkomponente ursprünglich theoretisch nicht eindeutig definiert. Aktuellere Abhandlungen zum EVT-Modell geben hingegen Hinweise darauf, dass die Erwartungskomponente der Wertkomponente zeitlich vorgelagert ist. So merken Wigfield et al. (2016) diesbezüglich an: „Thus children’s task values, expectancy, and competence beliefs increasingly are positively related, suggesting that children come to value what they are good at“ (Wigfield et al., 2016, S. 61). Hingegen unterstreichen eher interessenorientierte Ansätze die Rolle des subjektiven Wertes, um zunächst die Hinwendung einer Person zu einem bestimmten Themenbereich zu erklären, wovon später auch die eigene Fähigkeitsselbsteinschätzung profitiert (Wigfield et al., 2017; vgl. Eccles, 2005). So schreiben Harackiewicz und Knogler (2017, S. 345): „When students do perceive value in course topics, however, they develop greater interest in the course, work harder, perform better, persist longer, and are more likely to take additional courses and

complete their degree programs“. Entsprechende empirische Evidenzen zeichnen ein ebenso heterogenes Bild zur Beziehung der Erwartungs- und Wertkomponente.

In dem von Lauermaun et al. (2017) analysierten Sample us-amerikanischer Schülerinnen und Schüler des neunten sowie zwölften Jahrgangs fanden die Autorinnen für die Domäne Mathematik neben den bereits berichteten signifikant reziproken Beziehungen zwischen der Erwartungs- und Wertkomponente ebenfalls einen signifikant positiven Effekt des mathematischen Fähigkeitsselbstkonzepts in Jahrgang neun auf das spätere mathematische Fachinteresse in Jahrgang zwölf. Der umgekehrte Effekt vom Fachinteresse (9. Jg.) auf das spätere Fähigkeitsselbstkonzept (12. Jg.) war hingegen nicht signifikant.

Marsh et al. (2005) nutzten für ihre Analysen eine große repräsentative Stichprobe aus dem deutschen Projekt BIJU (vgl. Baumert et al., 2003), in dessen Rahmen Schülerinnen und Schüler des siebten Jahrgangs zwei Mal innerhalb des Schuljahres zum mathematischen Fähigkeitsselbstkonzept und zum Fachinteresse in Mathematik befragt wurden. Auch Marsh et al. (2005) bestätigten signifikante wechselseitige Beziehungen zwischen dem mathematischen Fähigkeitsselbstkonzept und dem Fachinteresse für Mathematik. Anhand längsschnittlicher Strukturgleichungsmodellierung fanden die Autoren zudem einen signifikanten Effekt zwischen dem mathematischen Fähigkeitsselbstkonzept des ersten Messzeitpunkts und dem mathematischen Fachinteresse des zweiten Messzeitpunkts. Hinweise für den signifikanten Einfluss des früheren Interesses auf das spätere Fähigkeitsselbstkonzept ergaben sich hingegen nicht. Analoge Analysen führten Marsh et al. (2005) in einem zweiten Sample deutscher Schülerinnen und Schüler des siebten und achten Jahrgangs durch. Dabei wurden zwischen den beiden Messzeitpunkten sowohl signifikante Effekte zwischen dem mathematischen Fähigkeitsselbstkonzept (7. Jg.) und dem mathematischen Fachinteresse (8. Jg.) als auch umgekehrt zwischen mathematischen Fachinteresse (7. Jg.) und dem mathematischen Fähigkeitsselbstkonzept (8. Jg.) festgestellt. Der Effekt zwischen dem mathematischen Fähigkeitsselbstkonzept (7. Jg.) und dem späteren mathematischen Fachinteresse (8. Jg.) war allerdings etwas stärker ausgeprägt.

Für die Primarstufe und die Domänen Mathematik und Deutsch konnten Spinath und Steinmayr (2008) anhand latenter Cross-Lagged-Panel-Modelle zwar ebenfalls mittlere bis hohe korrelative Zusammenhänge zwischen dem Fähigkeitsselbstkonzept und dem jeweiligen Interesse an fachspezifischen Aktivitäten (Sachinteresse) feststellen, von den analysierten kreuzversetzten Pfaden⁴ waren aber nur drei der insgesamt 24 Pfade überhaupt signifikant und zudem zwischen den Domänen und unterschiedlichen Messzeitpunkten willkürlich verteilt. Somit konnte nicht auf eine längsschnittliche Beziehung zwischen der Erwartungs- und Wertkomponente geschlossen werden (vgl. dazu auch Spinath & Spinath, 2005). Die fehlenden kreuzversetzten Pfade begründen Spinath und Steinmayr (2008) auch mit der von ihnen gefundenen Stabilität der Konstrukte über die Zeit. Denn „the more

⁴ Für eine Erklärung des Analyseverfahrens latenter Cross-Lagged-Panel-Modelle siehe Kapitel 8.4.5.

stable a construct, the lower the possibility of finding other variables that potentially explain variance“ (Spinath & Steinmayr, 2008, S. 1566). Allerdings fanden die vier Messzeitpunkte alle innerhalb eines Jahres (Ende 3. Jg. bis Ende 4. Jg.) statt, sodass vermutet werden kann, dass nicht nur die fehlenden kreuzversetzten Effekte, sondern auch die postulierte Stabilität der Konstrukte ein Artefakt der kurzen Messintervalle sind.

Für die Untersuchung längsschnittlicher Beziehungsmuster der Erwartungs- und Wertkomponente zu Beginn der Sekundarstufe I empfehlen Arens et al. (2019) daher einen Abstand von einem Schuljahr zwischen den Messzeitpunkten. Die Autorinnen nutzten ebenfalls latente Cross-Lagged-Panel-Modelle für die Analyse der längsschnittlichen Zusammenhgangsstruktur des Fähigkeitsselbstkonzepts sowie des Fachinteresses und der Wichtigkeit/Bedeutsamkeit in den Domänen Mathematik, Deutsch und Englisch der Jahrgänge fünf bis neun. Die Datenbasis bildete das Projekt BiKS⁵, in dem die erste Befragung der Schülerinnen und Schüler im fünften Jahrgang im Jahr 2008 stattfand. Hinsichtlich der querschnittlichen Zusammenhgangsstruktur zwischen der Erwartungs- und Wertkomponenten replizierten Arens et al. (2019) bisherige Forschungsergebnisse und zwar unabhängig von der betrachteten Domäne. Die Stärke der Korrelationen blieb im Zeitverlauf jedoch auf einem ähnlich hohen Niveau. Bezüglich der längsschnittlichen Zusammenhgangsstruktur fanden Arens et al. (2019) für die Domänen Mathematik und Englisch signifikante kreuzversetzte Pfade zwischen dem früheren Fähigkeitsselbstkonzept auf das spätere Fachinteresse sowie auf die von den Schülerinnen und Schülern benannte Bedeutsamkeit (*attainment value*). Inhaltlich bedeutet dies, dass eine hohe Einschätzung der eigenen Fähigkeiten im jeweiligen Bereich das spätere domänenspezifische Fachinteresse sowie die subjektive Bedeutsamkeit des Unterrichtsfachs stärkt. Hinsichtlich der Domäne Deutsch zeigten sich jedoch umgekehrte Ergebnisse: Das Fachinteresse und auch die Bedeutsamkeit erwiesen sich hier förderlich für ein höheres Fähigkeitsselbstkonzept zu einem späteren Zeitpunkt. Arens et al. (2019) schließen daraus, dass ein hohes Maß an subjektiver Relevanz (Interesse, Bedeutsamkeit) auch das Engagement und die Anstrengung der Schülerinnen und Schüler in diesem Bereich erhöht, wovon entsprechend das Fähigkeitsselbstkonzept profitiert. Hinsichtlich der Stabilität der Konstrukte über die Zeit bewegte sich das Fähigkeitsselbstkonzept in den drei Domänen im mittleren Bereich ($\beta = .43$ bis $\beta = .55$), während sich das jeweilige Fachinteresse im Längsschnitt als etwas weniger stabil erwies ($\beta = .28$ bis $\beta = .40$).

Zusammenfassend kann festgehalten werden, dass sich insbesondere die inhaltliche Definition und Operationalisierung der Wertkomponente zwischen den Forschungsarbeiten erheblich unterscheidet, was die Interpretation und Vergleichbarkeit der Befunde erschwert. Vornehmlich werden die beiden Subfacetten Bedeutsamkeit/Wichtigkeit und Interesse getrennt oder kombiniert verwendet (vgl. u.a.

⁵ *Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter* (BiKS; vgl. Artelt, Blossfeld, Faust, Roßbach & Weinert, 2013)

Arens et al., 2019; Fredricks & Eccles, 2002; Trautwein et al., 2012). Das Interesse kann sich dabei sowohl auf ein Schulfach als auch themenspezifische Aktivitäten beziehen (Marsh et al., 2005). Die Erwartungskomponente hingegen wird einheitlich als Selbsteinschätzung der eigenen Fähigkeiten (Fähigkeitsselbstkonzept) erhoben.

Wiederholt bestätigte sich in den berichteten Forschungsarbeiten vor allem der korrelative Zusammenhang zwischen der Erwartungs- und Wertkomponente (vgl. u.a. Dinkelmann & Buff, 2016; Lauer mann et al., 2017; Spinath & Steinmayr, 2008), wobei sich mit zunehmendem Alter der Schülerinnen und Schüler offenbar ein stärkerer Zusammenhang beider Komponenten einstellt, der auch als Spezialisierungsprozess interpretiert wird (Denissen et al., 2007; Fredricks & Eccles, 2002; Wigfield et al., 1997). Hinsichtlich der längsschnittlichen Beziehungsstruktur der Erwartungs- und Wertkomponente zeigten sich häufiger Effekte zwischen dem vorherigen Fähigkeitsselbstkonzept und dem späteren Interesse als umgekehrt (vgl. Arens et al., 2019; Lauer mann et al., 2017; Marsh et al., 2005). Allerdings ergaben sich hier domänenspezifische Unterschiede (Arens et al., 2019). Vornehmlich fokussierten sich die berichteten Ergebnisse dabei auf die Domäne Mathematik, gefolgt von Englisch und Deutsch (vgl. Arens et al., 2019; Lauer mann et al., 2017; Marsh et al., 2005). Die Fächer Naturwissenschaften und Musik wurden jedoch bisher in den längsschnittlichen Analysen ausgespart. Darüber hinaus wurde die Entwicklung der Erwartungs- und Wertkomponente in der Sekundarstufe bisher vor allem anhand von zwei Messzeitpunkten sowie in höheren Jahrgängen der Sekundarstufe betrachtet (Lauer mann et al., 2017; Marsh et al., 2005). Daher war es bisher nicht möglich, Veränderungen hinsichtlich der Stabilität der Konstrukte an sich und auch die Entwicklung der Beziehungen zwischen dem Fähigkeitsselbstkonzept und dem Interesse zu betrachten und entsprechende Effekte über einen längeren Zeitraum der Sekundarstufe I zu untersuchen. Arens et al. (2019) haben dieses Forschungsdesiderat aufgegriffen und untersuchten Schülerinnen und Schüler der fünften bis neunten Jahrgangsstufe. Allerdings fanden sie zwischen den untersuchten Domänen Differenzen in den längsschnittlichen Beziehungsmustern der Wert- und Erwartungskomponente und betonen daher den weiteren Forschungsbedarf.

5 Schul- und Klassenprofilierung

Im Folgenden wird das Thema Profilierung auf Schul- und Klassenebene zunächst allgemein eingeführt und dazu im Rahmen der Schulautonomiedebatte sowie im Kontext verstärkter Wettbewerbsorientierung verortet. Zudem werden die Begriffe Schul- und Klassenprofil inhaltlich definiert und die Profilbildung als nationales ebenso wie internationales Phänomen beschrieben (Kapitel 5.1). Anschließend werden Forschungsbefunde zur Verbreitung von Schul- und Klassenprofilen präsentiert (Kapitel 5.2). Der Frage nach den Bedingungsfaktoren der Wahl von Profilangeboten widmet sich Kapitel 5.3. Dabei wird einerseits thematisiert, welche Schülerklientel bevorzugt eine profilierte Schule wählt und andererseits, welche Wahlkriterien der nachfragenden Schülerinnen und Schüler ermöglichen, durch die Schulen festgelegten, Aufnahmekriterien für eine Profilklassse gegenüberstehen. Da im Zentrum dieser Arbeit die Untersuchung niedersächsischer Profilklassen steht, werden in Kapitel 5.4 entsprechende Rahmenbedingungen der niedersächsischen Schulgesetzgebung zur Gestaltung des Profilverrichts erläutert. Nachfolgend werden in Kapitel 5.5 konkrete organisatorische und didaktische Gestaltungsmerkmale des musikalischen sowie naturwissenschaftlichen Profilverrichts beschrieben und Forschungsbefunde zur wahrgenommenen Unterrichtsqualität in Profilklassen berichtet. Die mit den profilspezifischen Unterrichts Konzepten verknüpften pädagogischen Potentiale werden in Kapitel 5.6 vorgestellt, wobei ein Fokus auf mögliche Effekte auf das Interesse sowie das Fähigkeitsselbstkonzept der Schülerinnen und

Schüler in den Domänen Musik und Naturwissenschaften gelegt wird. Abschließend werden in Kapitel 5.7 die wenigen bisher vorliegenden empirischen Befunde zur Entwicklung des Fähigkeitsselbstkonzepts sowie des Interesses in Profilklassen präsentiert.

5.1 Profilbildung im Kontext verstärkter Schulautonomie und Wettbewerbsorientierung

In den 1990er Jahren begann in zahlreichen europäischen Staaten (vgl. u.a. Koivuhovi, 2021) darunter auch in Deutschland eine „neue Phase der Schulmodernisierung“ (Altrichter, Rürup & Schuchart, 2016, S. 107), deren schulpolitische Reformen insbesondere im Zeichen der Förderung von schulischer Autonomie standen (vgl. Altrichter & Heinrich, 2007). Die Ausgestaltung von Schulautonomie divergiert international zum Teil erheblich hinsichtlich der Entscheidungskompetenzen, die den Schulen jeweils eröffnet werden (Eurydice, 2007; Juranek, 2019). Während in anderen europäischen Ländern Schulautonomie beispielsweise die Verfügungsgewalt über finanzielle und personelle Ressourcen beinhaltet (z.B. Investitionen oder Besetzung von Lehrpersonal), haben Schulen in Deutschland insbesondere hinsichtlich der Organisation des Schulalltags sowie der Entwicklung von schulinternen Curricula mehr Mitspracherecht erhalten (Altrichter et al., 2016; Nonte, 2013). In der (deutschen) Autonomiedebatte wird der Begriff Schulautonomie folglich durchaus kontrovers diskutiert (Rürup, 2019; vgl. z.B. Avenarius, 1994), da damit keineswegs eine vollständige Autonomie – im Sinne einer Entkopplung der Einzelschule von der staatlichen Verantwortung für das Bildungswesen – verstanden wird, sondern vielmehr eine „relative Autonomie“ oder „Teilautonomie“ (Altrichter et al., 2016, S. 108), die mit der Steigerung der Handlungs- und Gestaltungsspielräume der Einzelschule⁶ verbunden ist (vgl. Juranek, 2019). Da Bildung in Deutschland Ländersache ist, überrascht es nicht, dass sich auch die Bundesländer im Grad der Schulautonomie unterscheiden. Dies zeigt eine Untersuchung des Aktionsrates Bildung von 2010 (Blossfeld et al., 2010), in der die Bereiche Finanzwesen, Personalwesen, Organisation und Verwaltung, pädagogische Aufgaben und Schulqualität bezüglich ihres Autonomiegrades vergleichend analysiert wurden. Das höchste Maß an Schulautonomie stellten die Autorinnen und Autoren in Nordrhein-Westfalen fest. In den Bundesländern Berlin, Bremen, Hamburg, Hessen, Niedersachsen, Rheinland-Pfalz, Sachsen-Anhalt und Schleswig-Holstein wurden die Schulgesetzte „in drei bis vier Regelungsbereichen als autonomiebegünstigend eingestuft“ (Blossfeld et al., 2010, S. 25). Diese Bundesländer befinden sich folglich in einem Entwicklungsprozess hin zu mehr Schulautonomie. In Bayern, dem Saarland und Thüringen war hingegen vergleichsweise wenig schulische Eigenverantwortung im

⁶ Für eine detaillierte Aufstellung schulischer Entscheidungsfreiräume sowie einen Rechtsvergleich zwischen Deutschland (Bayern, Hessen), Italien und Österreich siehe Juranek (2019).

Schulgesetz feststellbar. Einen relativ hohen Autonomiegrad fanden Blossfeld et al. (2010) bundeslandübergreifend im Bereich ‚pädagogische Aufgaben‘, zu dem eine Flexibilisierung der Lehrplanung, Unterrichtsorganisation, Stundentafel und Stundendauer (vgl. Buddeberg & Hußmann, 2021) sowie erweiterte Gestaltungsspielräume in der Unterrichts- und Erziehungsarbeit gezählt werden (Blossfeld et al., 2010). Konkret sieht beispielsweise das niedersächsische Kultusministerium für Gymnasien und integrierte/kooperative Gesamtschulen des Landes folgende Handlungsspielräume vor, über die der Schulvorstand nach § 38 a Abs. 3 Nr. 1 NSchG entscheiden kann: Um Schülerinnen und Schüler besonders zu fördern, spezielle Lehr-Lern-Formen zu entwickeln oder fächerübergreifendes Unterrichten zu ermöglichen, kann die Schule von den in den Stundentafeln vorgegebenen Fachstunden abweichen. Darüber hinaus ist in den Jahrgängen fünf bis zehn Epochenunterricht sowie Projektunterricht vorgesehen, der fachübergreifend oder fächerverbindend sowie schuljahrgangs- oder auch schul- und schulformübergreifend gestaltet werden kann. Schülerinnen und Schüler können zudem an wahlfreiem Unterricht teilnehmen, dessen Angebot ebenfalls klassen- und schuljahrgangsübergreifendes sowie schul- und schulformübergreifendes Unterrichten möglich macht (Niedersächsisches Kultusministerium, 2014, 2015a, 2015b).

Diese neu gewonnene „pädagogische Autonomie“ (Altrichter & Rürup, 2010, S. 114) bietet der Einzelschule die Möglichkeit, inhaltliche Schwerpunkte zu setzen und diese im Rahmen von Profilierungsprozessen auf Schul- oder Klassenebene zu verankern (Altrichter, Heinrich & Soukup-Altrichter, 2011; Blossfeld et al., 2010). Die Implementation von schulischen Profilen ist dabei als prominenteste Form der praktischen Umsetzung zunehmender Schulautonomie zu nennen (Heinrich, Altrichter & Soukup-Altrichter, 2011; Zunker & Neumann, 2020) und meint allgemein die Gestaltung von spezifischen und von außen identifizierbaren Schulmerkmalen. Ein schulisches Profil kann sich per Definition sowohl auf eine von der Einzelschule intendierte Außenwirkung (Image, Ruf, Ansehen) beziehen oder auch das pädagogische Selbstverständnis einer Schule beinhalten, das sich beispielsweise aus einer für die Schule charakteristischen didaktischen Gestaltung des Unterrichts speist. Darüber hinaus können im Rahmen von Profilierungsprozessen besondere curriculare oder extra-curriculare Angebote entwickelt und implementiert werden (Altrichter et al., 2011). Profilierungsmaßnahmen werden daher häufig auch im Zusammenhang mit Schulentwicklung genannt und als „intentionale[r] Prozess [bezeichnet], mit dem die Akteure einer Schule versuchen, das bestehende Profil dieser Schule – vor allem ihre unterrichtlichen und außerunterrichtlichen Angebote und ihre pädagogische Identität – weiterzuentwickeln“ (Altrichter et al., 2011, S. 12). Aus der Definition wird deutlich, dass Profilierung – in einem umfassenden Verständnis – folglich auf Schulebene als übergeordnetes pädagogisches Selbstverständnis/Konzept (im Folgenden als Schulprofil bezeichnet) zu verstehen ist, sich aber auch auf Klassenebene als besondere Form der Unterrichtsgestaltung und -organisation zeigen kann, die sich an schulinternen Lehrplänen orientiert (im Folgenden als Klassenprofil bezeichnet). Auf Schulebene finden sich inhaltliche

Schwerpunktsetzungen in Form von Profilen beispielsweise unter den Schlagworten *Europaschule* (vgl. Sonnenburg, 2020), *Bewegte Schule* (vgl. Thiel & Teubert, 2018) oder *MINT-EC-Schule*⁷. Profilklassen (d.h. Profilierungen auf Klassenebene) an weiterführenden Schulen haben sich unter anderem im Sport, im Fremdsprachenbereich (bilinguale Klassen), in Musik (Bläser-, Orchester- oder Bandklassen) oder den naturwissenschaftlichen Disziplinen (MINT-, Entdecker- oder Forscherklassen) etabliert (Göllner, 2017; Nonte, 2013; Seidel et al., 2016). Die Einrichtung von Profilklassen ist dabei keineswegs ein rein nationales Phänomen. Auch international finden sich Klassen mit einer spezifischen inhaltlichen Schwerpunktsetzung. So werden beispielsweise in Finnland ebenfalls naturwissenschaftliche, musikalische oder mathematische Profilklassen sowie profilierte Klassen im Bereich Sprachen oder Sport ab der ersten, dritten oder siebten Jahrgangsstufe angeboten (Koivuhovi, 2021). National ebenso wie international werden Profilierungsmaßnahmen auch genutzt, um im Wettbewerb gegenüber anderen Schulen zu bestehen (Altrichter et al., 2011; Kosunen, Bernelius, Seppänen & Porkka, 2020). Dabei werden von Schülerinnen und Schülern häufig angewählte Profilklassen verstetigt und unattraktive Profile hingegen eingestellt, um die Attraktivität einer Schule zu erhöhen (vgl. Altrichter & Rürup, 2010).

Profilierungsmaßnahmen sind folglich nicht selten durch einen erhöhten Wettbewerbsdruck gegenüber anderen Schulen in der Region motiviert und stellen gleichzeitig ein zentrales Werkzeug einer wettbewerbsorientierten Schullandschaft dar (Heinrich et al., 2011; Zunker, Neumann & Maaz, 2019). Gewirtz, Ball und Bowe (1995) unterscheiden diesbezüglich zwischen Wettbewerb erster und zweiter Ordnung. *Wettbewerb erster Ordnung* besteht zwischen Schulen hinsichtlich der Anmeldungen von Schülerinnen und Schülern. Eine Schule (und damit auch die Profilbildung) ist demnach erfolgreich, wenn sie (rein nominell) genügend Anmeldungen von Schülerinnen und Schülern erhält. Eine ausreichend große Schülerschaft sichert der Schule deren Fortbestand sowie die Zuweisung entsprechender (z.B. personeller) Ressourcen (Heinrich et al., 2011). Von *Wettbewerb zweiter Ordnung* ist hingegen die Rede, wenn Schulen nicht nur um die Anzahl an Schülerinnen und Schülern konkurrieren, sondern auch um eine bestimmte Schülerklientel (Gewirtz et al., 1995). Gibt es mehr Anmeldungen als zur Verfügung stehende Plätze, können anhand schulinterner Kriterien folglich besonders ‚passende‘ Schülerinnen und Schüler ausgewählt werden (s.a. Kapitel 5.3; vgl. Clausen, 2007; Heinrich et al., 2011; Helsper, Dreier, Gibson & Niemann, 2015; Krüger, Keßler & Winter, 2015; Varjo & Kalalahti, 2019). Auswahlkriterien beziehen sich beispielsweise auf Leistungen, Vorerfahrungen oder Interessen der Schülerinnen und Schüler (Hase, 2015; Helsper, Dreier, Gibson, Kotzyba & Niemann, 2015). In Abhängigkeit des Profils und dem damit verbundenen Selbstverständnis der Schule wird daraus ein Bild des idealen Schülers beziehungsweise der idealen Schülerin konstruiert, das sich aus leis-

⁷ <https://www.mint-ec.de/>

tungsbezogenen Merkmalen (z.B. Schulnoten), familiären Merkmalen (z.B. sozio-ökonomischer Hintergrund) sowie weiteren Merkmalen (z.B. Einstellungen, Bezug zur Schule) speisen kann (vgl. Helsper, Kramer & Thiersch, 2014; Helsper & Krüger, 2015; Krüger, 2015). Wettbewerbsprozesse und Profilierungsmaßnahmen bergen demnach auch die Gefahr einer verstärkten sozialen Segregation (vgl. u.a. Bellmann & Weiß, 2009; Morris-Lange, Wendt & Wohlfahrt, 2013; Specht, 2011; Zunker & Neumann, 2020). So stellten Weber, Moosbrugger, Hasengruber, Altrichter und Schrodtt (2019) beispielsweise im Nationalen Bildungsbericht Österreich 2018 fest, dass „*innerschulische Schwerpunkte als Ergebnis von Schulprofilierung* [Hervorhebung im Original] („Klassenprofile“ als Kriterium der Klassenzusammensetzung) sowohl in der Volksschule als auch der Sekundarstufe [...] mit einem höheren Ausmaß an innerschulischer Segregation einhergehen“ (Weber et al., 2019, S. 172). Auch weitere europäische Forschungsbefunde weisen darauf hin, dass Profilklassen insbesondere von leistungsstarken Schülerinnen und Schülern aus sozioökonomisch gut gestellten Familien besucht werden, sodass Fragen der Chancengleichheit im Zusammenhang mit Profilierungsmaßnahmen diskutiert werden (Koivuhovi, 2021; Kosunen et al., 2020).

Doch der Wettbewerbsdruck betrifft nicht alle Schulen gleichermaßen. Am Beispiel der Schullandschaft in Hamburg konnten Leist und Pietsch (2017) mittels Daten der Schulstatistik und stochastischer Netzwerkmodellierung zeigen, dass sich innerhalb der Stadt marktähnliche Strukturen mit verschiedenen lokalen Clustern schulischer Märkte (vgl. Weiß, 2001) bilden, die nicht zwingend den Grenzen der Stadtbezirke entsprechen, sondern vielmehr durch natürliche Grenzen (z.B. Flüsse oder Bahnlinien) entstehen. Zwischen den lokalen Clustern fanden die Autoren Unterschiede hinsichtlich des Herfindahl-Index (vgl. Belfield & Levin, 2002), der als Maß für die Stärke des schulischen Wettbewerbsdrucks um bestimmte Schülerinnen und Schüler verwendet wurde (Leist & Pietsch, 2017).

Welche Rolle Wettbewerb für die schulische Profilierung spielt, untersuchte Unger (2015) anhand vergleichender Fallanalysen und kommt zu dem Schluss, dass Schulen mit traditioneller Steuerung vergleichsweise weniger profiliert sind als Schulen mit Wettbewerbssteuerung und freie Schulen einen höheren Profilierungsgrad aufweisen als staatliche Schulen. Weber et al. (2021) zeigten zudem regionale Unterschiede in der schulischen Wettbewerbsorientierung zwischen städtischen und ländlichen Gebieten auf. Schulen im städtischen Raum nutzen demnach vermehrt Profilierungsmaßnahmen, um im Wettbewerb mit anderen Schulen zu bestehen (vgl. Nonte & Naacke, 2010), während Schulen im ländlichen Raum weniger Profilierungen aufweisen, da der „Zwang zur Profilierung“ offenbar in geringerem Maße wirksam ist (Heinrich et al., 2011, S. 276). Das Phänomen der Profilcluster innerhalb und im näheren Umfeld von größeren Städten kann auch im europäischen Ausland beobachtet werden (Koivuhovi, 2021).

5.2 Verbreitung von Schul- und Klassenprofilen

Unter Rückgriff auf die von Altrichter et al. (2011) relativ weit gefasste Definition des Profilierungsbegriffs weist Nonte (2022) darauf hin, dass heute nahezu alle allgemeinbildenden Schulen in Deutschland über ein Profil – im Sinne einer intendierten Außenwirkung – verfügen, das in der Öffentlichkeit (z.B. auf der Schulhomepage oder bei Informationsveranstaltungen) präsentiert wird (vgl. Budde, Kansteiner & Bossen, 2016). Hinsichtlich der Verbreitung von Schul- und Klassenprofilen im engeren Sinne, das heißt als pädagogisches Konzept oder spezifische Form der Unterrichtsgestaltung und -organisation, liegen hingegen nur sehr begrenzt aktuelle Zahlen für Deutschland im Allgemeinen und die 16 Bundesländer im Speziellen vor. Trotz der zunehmenden Verbreitung von schulischen Profilierungsmaßnahmen (vgl. Zunker & Neumann, 2020) stellen die zuständigen Schulbehörden auch aktuell keine offiziellen Angaben zur Verbreitung von Schul- und Klassenprofilen bereit, was die systematische Betrachtung des Angebots von Profilierungsmaßnahmen erheblich erschwert. Im Rahmen von Forschungsarbeiten, die das Themenfeld Profilierung adressieren, müssen daher Angaben zur Verbreitung von Schul- beziehungsweise Klassenprofilierungsmaßnahmen für das jeweils spezifische Untersuchungsfeld aufwendig zusammengestellt werden, wobei öffentliche Statistiken (z.B. Schullisten der Schulverwaltungen) durch eigene Recherchen zum Profilierungsangebot zu ergänzen sind.

Für die Schullandschaft in Berlin berichten Jurczok (2019) sowie Zunker, Neumann und Maaz (2018) entsprechende Zahlen zum schulischen Profilangebot. Auf Basis der Berliner Schulstatistik aus dem Jahr 2011 untersuchte Jurczok (2019) $n = 207$ Berliner Schulen in staatlicher Trägerschaft (Gymnasien, integrierte Sekundarschulen mit/ohne Oberstufe), wobei die Anteile der Schulformen zwischen den einzelnen Stadtbezirken stark variierten. Am häufigsten waren unter den analysierten Schulen Sprachprofile (45 %) sowie Profile im Bereich Kunst, Musik und Sport (zusammen 11 %) zu finden. Sieben Prozent der Schulen verfügten über ein MINT-Profil (Naturwissenschaften, Mathematik oder Informatik) und nur zwei Prozent der Schulen boten mehrere Profile gleichzeitig an. 45 Prozent der analysierten Schulen wiesen keine spezielle Profilierung auf.

Auch innerhalb der einzelnen Stadtbezirke ergaben sich Differenzen hinsichtlich des Profilangebots. So verfügten im Stadtbezirk Mitte 71 Prozent der Schulen über ein Profil. Die Stadtteile Steglitz-Zehlendorf (64 %) sowie Charlottenburg-Wilmersdorf und Neukölln (jeweils 61 %) wiesen etwas weniger Schulen mit Profil auf. Getrennt nach thematischen Schwerpunkten ergab sich in den drei Bezirken mit dem größten Profilangebot ebenfalls ein heterogenes Bild: Während im Bezirk Mitte am häufigsten Sprachprofile angeboten wurden (35 %), waren in Steglitz-Zehlendorf MINT-Profile vorherrschend (45 %) und im Bezirk Neukölln sprachliche sowie musisch-kulturell-sportliche Profile mit jeweils 22 Prozent gleichauf. Die Autorin schließt daraus, dass sich das Profilangebot an den Bedürfnissen der nachfragenden Schülerschaft orientiert. So war in Bezirken mit einem hohen Anteil von

Schülerinnen und Schülern mit nichtdeutscher Herkunftssprache (z.B. Bezirk Mitte oder Neukölln) der Bedarf an sprachlichen Profilen höher, während beispielsweise im Bezirk Steglitz-Zehlendorf eine hohe Leistungsorientierung der Schülerinnen und Schüler (gemessen an der durchschnittlichen Abiturnote im Vergleich zu den anderen Bezirken) die hohe Nachfrage nach MINT-Profilen erklären konnte. Folglich unterschieden sich auch die Intentionen für die Implementation von Profilen zwischen den Stadtbezirken (Jurczok, 2019).

Zunker et al. (2018) analysierten ebenfalls das Profilangebot an Berliner Schulen, beschränkten ihre Untersuchung jedoch auf integrierte Sekundarschulen mit beziehungsweise ohne gymnasialer Oberstufe und nutzten Daten aus dem Schuljahr 2014/2015. Als ‚Profilierung‘ definierten die Autoren sowohl thematische Schwerpunkte auf Schul- als auch auf Klassenebene und unterschieden zwischen naturwissenschaftlichen, sprachlichen, musisch-künstlerischen sowie beruflichen Profilen. Etwas mehr als die Hälfte der untersuchten Schulen bot mindestens ein Profil an und ein Fünftel der Schulen war sogar mehrfach profiliert. Insgesamt waren MINT-Profile in der Stichprobe am häufigsten vertreten, gefolgt von Sprachprofilen, musisch-künstlerischen Profilen sowie beruflichen Profilen. Die getrennte Betrachtung der integrierten Sekundarschulen mit gymnasialer Oberstufe und der integrierten Sekundarschulen ohne gymnasiale Oberstufe trug ein differenziertes Bild des Profilangebotes zu Tage: Integrierte Sekundarschulen mit gymnasialer Oberstufe boten demnach weitaus häufiger sprachliche, musisch-künstlerische sowie naturwissenschaftliche Profile an als integrierte Sekundarschulen ohne gymnasiale Oberstufe. Berufliche Profile waren hingegen verstärkt an integrierten Sekundarschulen ohne gymnasiale Oberstufe zu finden. Diese schulformspezifischen Unterschiede stützen folglich die These, dass die Schulen ihr Profilangebot ihrer Schülerklientel und damit verbundenen Interessen, Bedarfen oder Leistungsorientierungen der Schülerinnen und Schüler anpassen.

Weber et al. (2021) untersuchten ebenfalls Unterschiede in den Profilangeboten zwischen Schulformen (Gymnasien und integrierte/kooperative Gesamtschulen), legten darüber hinaus einen besonderen Fokus auf Unterschiede zwischen städtischen und ländlichen Regionen und differenzierten zudem zwischen Profilierungen auf Schul- und Klassenebene. Auf Grundlage einer systematischen Analyse von Schulhomepages untersuchten die Autorinnen und der Autor für Hamburg und den Regierungsbezirk Braunschweig (Niedersachsen) die Verbreitung von Schul- und Klassenprofilen und nutzten dafür amtliche Daten von 2013. Von den 306 analysierten Schulhomepages berichteten 37.9 Prozent der Schulen (entspricht 116 Schulen) öffentlich über Profilierungsmaßnahmen und inhaltliche Schwerpunktsetzungen. Von diesen 116 profilierten Schulen verfügte ein Drittel der Schulen über ein Schulprofil und zwei Drittel der Schulen über ein oder mehrere Klassenprofil(e). Keine der analysierten Schulen gab jedoch beide Profilierungsansätze an. Getrennt nach Standorten zeigte sich, dass in der Großstadt Hamburg 16.2 Prozent der Schulen ein Schulprofil präsentierten, während im eher ländlichen Regierungsbezirk Braunschweig nur 10.6 Prozent der Schulen angaben, ein Schulprofil zu haben.

Auch hinsichtlich der Schulform zeigten sich Unterschiede in der Verbreitung der Schulprofile. In Hamburg lag der Anteil der Gymnasien, die ein Schulprofil angeben, mit 16.4 Prozent etwas höher als der Anteil der Stadtteilschulen mit einem Schulprofil (14.5 %). Im Regierungsbezirk Braunschweig präsentierten 23.1 Prozent der Oberschulen, 17.4 Prozent der kooperativen und integrierten Gesamtschulen sowie 7.8 Prozent der Gymnasien ein Schulprofil auf ihrer Homepage. Zu den am häufigsten genannten inhaltlichen Schwerpunkten gehören in beiden Regionen sprachliche Profile, Sport-Profile sowie musikalische und naturwissenschaftliche Profile.

Deutlich häufiger als Schulprofile benannten die untersuchten Schulen hingegen Profilierungsmaßnahmen auf Klassenebene. In Hamburg gaben 47 von 117 Schulen (40.2 %) an, Klassenprofile anzubieten. Im Regierungsbezirk Braunschweig war der Anteil an Schulen mit Klassenprofilen deutlich geringer (30 von 189 Schulen; entspricht 16.0 %), wobei Gymnasien im Regierungsbezirk Braunschweig signifikant häufiger Klassenprofile einrichteten als andere Schulformen. Bei den Klassenprofilen konnten in beiden Bezirken die inhaltlichen Schwerpunkte Musik (61.0 %) und Naturwissenschaften/Technik (42.9 %) als die beliebtesten Profile herausgestellt werden.

Einen Beleg für die zunehmende Beliebtheit (vgl. Heß, 2017a) sowie die flächendeckende Verbreitung von musikalischen Profilklassen bietet zudem der vom Verband deutscher Musikschulen (VdM) herausgegebene statistische Jahresbericht, in dem Klassenmusizierprojekte bundesweit erfasst werden (für eine nähere Erklärung des Klassenmusizierbegriffs und den verschiedenen Konzepten von Musik-Profilklassen s. Kapitel 5.5). Da in Musik-Profilklassen der Instrumentalunterricht meist von externen Musikschulen übernommen wird (s. Kapitel 5.5), entsprechen die beim VdM erfassten Schülerinnen und Schüler in Klassenmusizierprojekten näherungsweise auch den Teilnehmerinnen und Teilnehmern in Musik-Profilklassen. Demnach nahmen im Jahr 2020 über 250 000 Schülerinnen und Schüler an Klassenmusizierprojekten (d.h. an Musik-Profilklassen) teil. Der weitaus größte Anteil der teilnehmenden Schülerinnen und Schüler entfiel dabei auf Bläserklassen. Im Vergleich zum Jahr 2013 hat sich im Jahr 2020 zudem der Anteil der an Klassenmusizierprojekten teilnehmenden Schülerinnen und Schüler sowohl in Bläser- als auch Streicher- oder Chorklassen verdoppelt (Verband deutscher Musikschulen, 2021).

Für naturwissenschaftliche Profilklassen liefert das Netzwerk der MINT-EC-Schulen einen Anhaltspunkt zur deutschlandweiten Verbreitung. Dem im Jahr 2000 gegründeten naturwissenschaftlichen Exzellenz-Netzwerk können sich „vorbildliche Schulen mit hervorragendem MINT-Profil“ anschließen (Arbeitsgruppe MINT-EC-Zertifikat, 2020, S. 7). Das sogenannte MINT-EC-Zertifikat wird an besonders leistungsstarke Schülerinnen und Schüler im MINT-Bereich vergeben. Für den Erwerb des Zertifikates können Schülerinnen und Schüler unter anderem durch den Besuch einer MINT-Profilklassen oder der Teilnahme an naturwissenschaftli-

chen Wettbewerben (s. dazu Kapitel 5.5) Punkte sammeln. Deutschlandweit gehören aktuell 339 Schulen sowie circa 360 000 Schülerinnen und Schüler dem Netzwerk an. Die meisten MINT-EC-Schulen finden sich in Nordrhein-Westfalen (102), gefolgt von Bayern (61) und Hessen (38). In Niedersachsen sind 34 Schulen Teil des MINT-EC-Netzwerkes (MINT-EC, 2021).

Zusammenfassend lässt sich festhalten, dass das Angebot von Profilen auf Schul- oder Klassenebene weiterhin als Alleinstellungsmerkmal einer Schule zu werten ist, da nur ein Teil der untersuchten Schulen überhaupt Profile anbietet. Weiterhin zeigen sich sowohl regional als auch zwischen den Schulformen Unterschiede hinsichtlich der Verbreitung von Profilierungsmaßnahmen, wobei in städtischen Regionen sowie an Schulformen, die zum Abitur führen, Profile tendenziell weiterverbreitet waren. Bezüglich der inhaltlichen Schwerpunkte traten Sprach-, MINT- und musisch-künstlerische Profile am häufigsten auf. Ein kleinerer Teil der Schulen bietet darüber hinaus mehrere inhaltliche Schwerpunkte gleichzeitig an, die jedoch eher auf Klassenebene (Profilklassen) als auf Schulebene zu verorten sind (Weber et al., 2021; Zunker et al., 2018). Für Niedersachsen konnten MINT- und Musik-Klassen als die am häufigsten vorkommenden Profile auf Klassenebene identifiziert werden (Weber et al., 2021). Daraus leitet sich der inhaltliche Fokus dieser Arbeit auf Profilklassen in den Domänen Musik und Naturwissenschaften ab.

5.3 Wahl von Profilangeboten

Seit dem Schuljahr 2015/2016 wird in Niedersachsen für den Übergang von der Grundschule auf die weiterführende Schule keine Schullaufbahempfehlung mehr ausgesprochen. Stattdessen können sich die Schülerinnen und Schüler gemeinsam mit ihren Eltern von den Grundschullehrkräften bezüglich der Lernentwicklung und des Leistungsstandes beraten lassen, um für den Übergang in die Sekundarstufe I eine geeignete weiterführende Schule beziehungsweise Schulform auszuwählen (Niedersächsisches Kultusministerium, 2020a, 2020b). Darüber hinaus finden Informationsveranstaltungen der Grundschulen sowie der weiterführenden Schulen zum Schulübergang statt (Niedersächsisches Kultusministerium, 2020a). Bei der Wahl der weiterführenden Schule gilt folglich der „freie Elternwille“ (ebd., S. 6), wobei explizit betont wird, dass „die Schülerin oder der Schüler [...] in die Beratungen einzubeziehen [ist]“ (Niedersächsisches Kultusministerium, 2020b, S. 359). Dennoch können nur jene Schulen gewählt werden, die im jeweiligen Schulbezirk des Wohnortes liegen. Freie Schulwahl besteht hingegen, sofern vom Schulträger keine Schulbezirke eingerichtet wurden sowie für die Anmeldung an einer Ganztagschule (Niedersächsisches Kultusministerium, 2020a). Auch regional können Unterschiede bezüglich der Schulwahlmöglichkeiten bestehen. Während in städtischen Bereichen ein vielfältiges Angebot an Schulen im Sekundarbereich vor-

herrscht, kann eine ‚echte‘ Wahlentscheidung in ländlichen Bereichen häufig aufgrund eines Mangels an Alternativen eingeschränkt sein (vgl. Clausen, 2006; Weiß, 2001).

Zur Wahl der weiterführenden Schule (bzw. der Schulform) finden sich bereits umfangreiche Forschungsarbeiten (vgl. u.a. Clausen, 2006, 2007; Jurczok, 2019; Wohlkinger, 2014; Wohlkinger & Ditton, 2012; Zunker et al., 2018, 2019). Als Entscheidungskriterien der Einzelschulwahl konnten dabei die Nähe zum Wohnort, der ‚Ruf‘ der Schule (d.h. die Außendarstellung), schulische Rahmenbedingungen (z.B. Schulform, soziales Umfeld, Ausstattung), die Schulqualität (z.B. Schulklima, Leistungsanforderungen bzw. -standards) sowie der Wunsch des Kindes herausgestellt werden (Clausen, 2006; Jurczok, 2019). Darüber hinaus betont Clausen (2006), dass insbesondere bei der Wahl eines Gymnasiums das Schulprofil, im Sinne eines pädagogischen Konzeptes mit bestimmten thematischen Schwerpunkten (z.B. im Wahlpflichtbereich) oder Zusatzangeboten (z.B. Arbeitsgemeinschaften, Hausaufgabenbetreuung), eine entscheidende Rolle spielt. Auch Jurczok (2019) konnte in ihren Analysen zur Berliner Schullandschaft feststellen, dass das Angebot eines Profils oder mehrerer Profile ein essentieller Faktor im Schulwahlprozess ist. Schulen, die mehrere Profile gleichzeitig anbieten, werden demnach besonders häufig gewählt. Ebenfalls für Berlin zeigten Zunker et al. (2018, 2019), dass die Implementation eines naturwissenschaftlichen, sprachlichen oder musisch-künstlerischen Profils positiv mit der Nachfrage einer Schule (gemessen an den eingegangenen Erstwunschanmeldungen zur 7. Klasse) an integrierten Sekundarstufen in Zusammenhang steht. Darüber hinaus stellten die Autoren fest, dass es an den profilierten Schulen mehr Erstwunschanmeldungen gab, als tatsächlich Plätze zur Verfügung standen und es daher zu einer „potentiellen Diskrepanz zwischen Angebot und Nachfrage“ kommen kann (Zunker et al., 2018, S. 603).

Aber welche Schülerklientel wählt eine profilierte Schule? Wiederholt konnte belegt werden, dass die Entscheidung für eine profilierte Schule mit dem familiären Hintergrund zusammenhängt (vgl. Clausen, 2006; Jurczok, 2019). Demnach hat das Kriterium ‚Profilangebot‘ für Familien mit einem höheren sozialen Status mehr Bedeutung bei der Schulwahl als dies für Familien mit niedrigem sozialen Status der Fall ist. Zunker und Neumann (2020) stellten diesbezüglich fest, dass die Wahrscheinlichkeit, eine integrierte Sekundarstufe ohne Profil zu wählen, für Schülerinnen und Schüler, deren Eltern maximal einen Hauptschulabschluss besitzen, signifikant höher ist als für Schülerinnen und Schüler aus Elternhäusern mit höheren Bildungsabschlüssen. Hingegen wurde das ‚akademische Profil‘ (d.h. Schulen mit sprachlichem, naturwissenschaftlichem oder musisch-künstlerischem Profil) signifikant seltener von Schülerinnen und Schülern aus Haushalten mit niedrigem Bildungsabschluss (max. Hauptschulabschluss) gewählt.

Für die Schulwahl scheinen sich folglich die theoretisch formulierten und von den Schulen intendierten Effekte durch die Profilierungen (Steigerung der schulischen Attraktivität verbunden mit höheren Anmeldezahlen sowie der Beeinflussung der Schülerklientel (vgl. Kapitel 5.1; Clausen, 2007; Heinrich et al., 2011; Helsper, Dreier, Gibson & Niemann, 2015) auch in der Praxis zu bestätigen.

Wahl von Profilklassen

Für die Wahl von Profilangeboten auf Klassenebene (Profilklassen) liegen hingegen bisher nur wenige Forschungsergebnisse vor. Göllner (2017) stellte auf Basis einer qualitativen Interviewstudie zu Bläserklassen in der Sekundarstufe I (als spezifische Form der Musik-Profilklassen; s. Kapitel 5.5) heraus, dass die Schülerinnen und Schüler die Wahl einer solchen musikalischen Profilklassse überwiegend „als einen freiwilligen und wohl überlegten Entschluss [beschreiben]“, wobei die unterrichtenden Lehrkräfte das besondere thematische Interesse der Schülerinnen und Schüler an den Unterrichtsinhalten der Bläserklassen betonen (Göllner, 2017, S. 225). Als Zielgruppe für eine Musik-Klasse benennt auch Pabst-Krueger (2020, S. 160) „musikalisch besonders interessierte[n] Schüler[n]“. MINT-Klassen richten sich ebenfalls an naturwissenschaftlich interessierte Schülerinnen und Schüler, die ihre Kompetenzen im MINT-Bereich erweitern möchten (Seidel et al., 2016). Um diese Zielklientel zu erreichen, ist in erster Linie entscheidend, dass die Nachfragenden, das heißt die Schülerinnen und Schüler sowie deren Eltern, das Profilanangebot überhaupt kennen. Dafür eignen sich insbesondere Informationsveranstaltungen, Broschüren sowie Informationen von Bekannten, Eltern oder Geschwistern, wie Holmeier (2017) für die Wahl von schweizer MINT-Klassen herausstellte.

Koivuhovi, Vainikainen und Kalalahti (2020) konnten auf Basis eines finnischen Samples von sprachlichen, naturwissenschaftlichen sowie musisch-künstlerischen Profilklassen des siebten bis neunten Jahrgangs feststellen, dass Musik- und MINT-Klassen vergleichsweise leistungsstarke Schülerinnen und Schüler anziehen. Darüber hinaus fanden die Autorinnen geschlechtsspezifische Unterschiede innerhalb der verschiedenen Klassentypen. Vor allem in den MINT-Klassen und den musisch-künstlerischen Klassen war die Ungleichverteilung zwischen Jungen und Mädchen am stärksten ausgeprägt. MINT-Klassen wurden demnach von überproportional vielen Jungen besucht, während in musisch-künstlerischen Profilklassen der Anteil der Mädchen höher war. Ähnliche Ergebnisse finden sich auch bei Eder (2011). In den im Schuljahr 2005/2006 untersuchten österreichischen Musik-Profilklassen der siebten und achten Jahrgangsstufe ($n = 20$) beobachtete Eder (2011) eine deutliche Ungleichverteilung zugunsten der Mädchen. Nur ein Drittel der Kinder in Musik-Klassen waren demnach männlich.

Nonte et al. (2021) sowie Krieg, Nonte, Haas und Stubbe (2022) griffen diese Forschungsergebnisse auf und untersuchten für musikalische und naturwissenschaftliche Profilklassen in Niedersachsen, welche konkreten Faktoren im Zusammenhang mit der Wahl einer Musik- beziehungsweise MINT-Profilklasse stehen.

Da gerade musisch-künstlerische sowie naturwissenschaftliche Fachdomänen wiederholt als besonders geschlechtskonnotierte Bereich beschrieben wurden (vgl. Kapitel 2.5; s.a. Bennetts, 2013; Eccles, Wigfield et al., 1993; Hall, 2005; Hannover & Kessels, 2002; Heß, 2017b; Lazarides & Lauermaun, 2019; Wolter & Hannover, 2016), legen die Autorinnen und der Autor einen Fokus auf die Rolle des Geschlechtsrollen-Selbstkonzepts bei der Wahl einer Musik- oder MINT-Profilklasse. Es wurde angenommen, dass Kinder, die sich selbst in einem höheren Maße weiblich konnotierte Eigenschaften zuschreiben, eher eine Musik-Klasse wählen, als Kinder die angeben über eher männlich konnotierte Eigenschaften zu verfügen (Nonte et al., 2021). Andererseits wurde untersucht, ob Kinder mit selbstattestierten männlich konnotierten Eigenschaften eher naturwissenschaftliche Profilklassen wählen (Krieg et al., 2022). Neben dem Geschlechtsrollen-Selbstkonzept wurden in beiden Forschungsarbeiten weitere Merkmale, wie das fachspezifische Fähigkeits-selbstkonzept sowie Sachinteresse der Schülerinnen und Schüler, der von ihnen wahrgenommene Peerdruck sich geschlechtsrollenkonform zu verhalten und die elterliche Einstellung gegenüber Musik beziehungsweise Naturwissenschaften in die Analysen einbezogen und dabei auch der Effekt des sozioökonomischen Hintergrundes kontrolliert. Weder für die untersuchten Musik-Klassen noch für die MINT-Klassen konnten signifikante Effekte zwischen dem sozioökonomischen Status der Schülerinnen und Schüler (bzw. ihrer Eltern) und der Wahl einer Profilklasse identifiziert werden, was darauf schließen lässt, dass Tendenzen sozialer Segregation bei der Wahl einer Profilklasse – zumindest im analysierten Sample beziehungsweise Modell – weniger Bedeutung zu haben scheinen (Krieg et al., 2022; Nonte et al., 2021). Für die Wahl einer Musik-Klasse zeigten die Ergebnisse, dass nicht das Geschlechtsrollen-Selbstkonzept, sondern das Interesse an Musik den stärksten Effekt auf die Wahl einer musikalischen Profilklasse hatte. Schülerinnen und Schüler, die angaben, an Musik interessiert zu sein, hatten demnach eine drei Mal höhere Chance (*odds ratio* = 3.2) eine Musik-Klasse zu wählen als Kinder mit einem niedrigen musikalischen Interesse (Nonte et al., 2021). Für die Wahl einer MINT-Klasse ließen sich hingegen keine signifikanten Effekte der untersuchten Merkmale feststellen. Es wird daher geschlussfolgert, dass im sonst eher geschlechtskonnotierten Bereich der Naturwissenschaften zu beobachtende geschlechtsspezifische Effekte bei der Wahl einer MINT-Profilklasse weniger zum Tragen kommen (Krieg et al., 2022).

Im Anschluss an die von Nonte et al. (2021) für die Musik-Klassen herausgestellte Bedeutsamkeit des musikalischen Interesses für die Wahl einer Musik-Klasse betont auch Holmeier (2017) in ihrer Evaluationsstudie eines MINT-Klassen-Konzepts in der Schweiz die entscheidende Rolle des Interesses an naturwissenschaftlichen Themen für die Wahl einer naturwissenschaftlichen Profilklasse. Als weiteren wichtigen Faktor der Wahlentscheidung für eine naturwissenschaftliche Profilklasse benennt Holmeier (2017) die den Schülerinnen und Schülern in MINT-Klassen gebotene Möglichkeit, sich praktisch mit MINT-Themen auseinanderzusetzen zu kön-

nen (vgl. Kapitel 5.6). Erstaunlicherweise stellte Holmeier (2017) fest, dass sich besonders interessierte Schülerinnen und Schüler auch dann für eine MINT-Klasse anmeldeten, wenn keine expliziten Aufnahmekriterien bestehen. Gleiches gilt für das naturwissenschaftliche Fähigkeitsselbstkonzept, woraus die Autorin schließt, dass die Schülerinnen und Schüler „selbst einschätzen [können], ob sie für das Angebot geeignet sind oder nicht“ (Holmeier, 2017, S. 397). Dennoch können Schulen Aufnahmekriterien für eine Profilklassse festlegen, die im Folgenden zusammengefasst werden.

Aufnahmekriterien für Profilklassen

Der Besuch einer Profilklassse ist zum Teil an Aufnahmekriterien gebunden, die sich auf die Leistungen oder Interessen der Schülerinnen und Schüler beziehen können (Biedermann, Weber, Herzog-Punzenberger & Nagel, 2016). Entsprechende Aufnahmebedingungen oder Beschreibungen der für das jeweilige Profilagebot angesprochenen Zielgruppe werden schulintern beschlossen und unterscheiden sich daher zum Teil zwischen den Einzelschulen. Informationen zu den Aufnahmekriterien beziehungsweise Zielgruppenbeschreibungen finden sich in den Schulprogrammen sowie auf den Schulhomepages. Die Informationen sind frei zugänglich und werden meist prominent in den Selbstdarstellungen der Schulen platziert, um das Profilagebot der Schule vorzustellen. Exemplarisch wird im Folgenden auf einige Informationsquellen von Schulen verwiesen, an denen Daten für die vorliegende Arbeit erhoben wurden, um so einen Einblick in die Zugangsvoraussetzungen der Schulen des vorliegenden Samples zu erhalten.

Für die Aufnahme in eine Musik-Klasse ist es an einigen Schulen erforderlich, dass die Schülerinnen und Schüler bereits ein Instrument beherrschen beziehungsweise privaten Instrumental- und/oder Gesangsunterricht in bestimmtem Umfang vorweisen können (s. u.a. Otto Hahn Gymnasium Göttingen, 2021). Weiterhin sind Aufnahmetests für Musik-Klassen möglich (Göllner, 2017; Pabst-Krueger, 2020). Eine Ausnahme – und damit ein voraussetzungsfreies Angebot – stellt das Konzept der Bläserklassen dar (s. Kapitel 5.5), in denen die Schülerinnen und Schüler mit Beginn der fünften Klasse ein für sie neues Blasinstrument lernen (Göllner, 2017). Bei der Wahl einer Musik-Klasse (auch einer Bläserklasse) muss jedoch beachtet werden, dass monatliche Kosten von circa 40 Euro für den Instrumentalunterricht (meist in einer externen Musikschule oder unter Anleitung externer Musiklehrkräfte), Notenmaterial, die Ausleihe und Versicherung der Instrumente sowie notwendige Reparaturen der Instrumente anfallen (s. u.a. Gymnasium Martino Katharineum, 2017; Gymnasium Stolzenau, 2020a; IGS Roderbruch, 2020).

Für die Aufnahme in eine MINT-Klasse wird ein „verstärktes Interesse an naturwissenschaftlichen Fragestellungen“, „Beharrlichkeit und Frustrationstoleranz“ sowie „Teamfähigkeit“ gefordert, um den Anforderungen der MINT-Klasse gerecht zu werden (Gymnasium Stolzenau, 2020b, S. 1). Der Besuch einer MINT-Klasse ist zudem unter Umständen damit verknüpft, dass die Schülerinnen und

Schüler an einem naturwissenschaftlichen Wettbewerb (vgl. Kapitel 5.6) teilnehmen, wobei für die Wettbewerbsteilnahme ebenfalls Kosten (Fahrtkosten, Teilnahmegebühren) anfallen (Gymnasium Stolzenau, 2020b). Auch für stattfindende Exkursionen sind zusätzliche Kosten einzuplanen (Gymnasium unter den Eichen, 2018). Teilweise ist außerdem ein internetfähiger Computer notwendig, der privat zur Verfügung gestellt werden muss (Gymnasium Stolzenau, 2020b) oder durch die Schule für die Profilkasse neu angeschafft wird. Die Kosten dafür sind von den Eltern zu tragen (Gymnasium Alfeld, 2021).

5.4 Rechtliche Rahmenbedingungen zur Unterrichtsgestaltung in Profilklassen

An niedersächsischen Gymnasien obliegt es dem Schulvorstand über die Entwicklung und Implementierung von Profilen beziehungsweise Profilverunterricht zu entscheiden (Niedersächsisches Kultusministerium, 2015b). Dabei sind folgende inhaltliche Schwerpunktsetzungen der Profile möglich: Alte und neue Sprachen, Musik sowie Mathematik/Naturwissenschaften. Der Profilverunterricht findet normalerweise im Klassenverband statt. Die Verteilung der Fachstunden wird nach der Stundentafel 2 vorgenommen, die – im Gegensatz zur Stundentafel 1 – Profilverunterricht explizit aufführt. Verfügt eine Schule über ein musikalisches Profil, so wird das Fach Musik wöchentlich vier Stunden (statt der vorgesehenen zwei Unterrichtsstunden in den Jahrgängen fünf bis sieben) unterrichtet. Bietet ein Gymnasium ein Profil im Bereich Mathematik/Naturwissenschaft an, so werden die mathematisch-naturwissenschaftlichen Fächer zusätzlich durch die in der Stundentafel vorgesehenen Stunden für den Profilverunterricht (drei bis vier Unterrichtsstunden pro Woche) aufgestockt. Alternativ zur Einrichtung von Profilverunterricht kann in der Sekundarstufe Wahlunterricht in den Bereichen Fremdsprachen, Musik/Kunst, Gesellschaftswissenschaften (Geschichte, Politik, Erdkunde u.a.), Naturwissenschaften oder Sport angeboten werden. Falls der Schule eine entsprechende Unterrichtsgenehmigung vorliegt, ist zudem Wahlunterricht unter anderem in den Fächern Pädagogik, Philosophie oder Wirtschaftslehre möglich. Wahlunterricht kann bereits ab der fünften Jahrgangsstufe erteilt und ebenfalls mit zusätzlichen Unterrichtsstunden verknüpft werden. Sowohl die Stundentafel 1 als auch die Stundentafel 2 führt den Wahlunterricht auf, jedoch überwiegend ohne konkrete Maßgabe für die zu unterrichtenden wöchentlichen Stunden. Einzig für die Wahlfremdsprache und den Wahlunterricht mit musikalischem Schwerpunkt sind vier Stunden pro Woche für die Schülerinnen und Schüler nach der Stundentafel vorgeschrieben. Der Schulvorstand kann zudem über eine von den Stundentafeln abweichende Verteilung der Unterrichtsstunden für bestimmte Fächer entscheiden, so lang die Vorgaben zur erlaubten Gesamtwochenstunden der Sekundarstufe I eingehalten werden beziehungsweise pro Woche nicht mehr oder weniger als eine Unterrichtsstunde von der Stundentafel abgewichen wird (Niedersächsisches Kultusministerium, 2015b). Somit sind inhaltliche

Schwerpunktsetzungen und zusätzliche Unterrichtsstunden im Rahmen des Wahlunterrichts auch in anderen (als den für den Profilverricht vorgesehenen) Fächern möglich.

In den Rechtsvorschriften des niedersächsischen Kultusministeriums für die Jahrgänge fünf bis zehn an integrierten und kooperativen Gesamtschulen wird Profilverricht nicht explizit aufgeführt. Jedoch wird auch an Gesamtschulen das Angebot von Wahl(pflicht-)unterricht genutzt, um Schülerinnen und Schülern eine inhaltliche Schwerpunktsetzung zu ermöglichen. Der Wahl(pflicht-)unterricht wird als Differenzierungsmaßnahme geführt und soll die Förderung der Interessen der Schülerinnen und Schüler ermöglichen, die durch die Wahldifferenzierung individuelle „Lernschwerpunkte entwickeln können“ (Niedersächsisches Kultusministerium, 2014, S. 445). Wahl(pflicht-)unterricht kann fachübergreifend sowie – nach Genehmigung der Schulleitung – auch als Blockunterricht erteilt werden. Die jeweiligen Schwerpunkte werden von den Schülerinnen und Schülern für ein Schulhalbjahr gewählt, wobei Abweichungen in der Dauer auch hier möglich sind. Im siebten und achten Jahrgang sind pro Woche vier zusätzliche Schulstunden im Wahlpflichtbereich vorgesehen, die in einem Fach mit vier oder in zwei Fächern mit je zwei zusätzlichen Stunden erteilt werden können. Für ein sprachliches Wahlpflichtfach sind vier Stunden vorgeschrieben. Wahlpflichtfächer in den Naturwissenschaften, der Gesellschaftslehre oder der musisch-kulturellen Bildung können zwei- oder vierstündig erteilt werden. Wahlunterricht kann hingegen bereits ab der fünften Jahrgangsstufe angeboten werden. Konkrete Vorgaben zum wöchentlichen Stundenumfang liegen jedoch nicht vor. Eine mögliche Abweichung von der in der Stundentafel vorgesehenen Verteilung der Fachstunden bietet jedoch auch den Gesamtschulen zusätzliche Flexibilität hinsichtlich der Gestaltung der Stundenpläne und ermöglicht es inhaltliche Schwerpunkte zu setzen (Niedersächsisches Kultusministerium, 2014; 2015a). Bezüglich der Gesamtstundenzahl pro Woche wird in der Stundentafel explizit darauf hingewiesen, dass auch die Gesamtschulen ein „Stundenkontingent zur schuleigenen Schwerpunktsetzung [erhalten]“ (Niedersächsisches Kultusministerium, 2014, S. 11; Niedersächsisches Kultusministerium, 2015a, S. 449), um inhaltliche Differenzierungsmaßnahmen (z.B. im Wahl(pflicht-)unterricht) umzusetzen. Die Zuteilung des eigenverantwortlich zu verwaltenden Stundenkontingentes ist im sogenannten Erlass ‚Klassenbildung und Lehrkräfte-stundenzuweisung an den allgemein bildenden Schulen‘ (Niedersächsisches Kultusministerium, 2019) geregelt. Der über den Grundbedarf hinausgehende Stundenpool ist nicht generell zu bestimmen, da dieser durch spezifische Faktoren der Einzelschule beeinflusst wird. Dazu gehören unter anderem Ganztagsunterricht, Klassengrößen, das Angebot besonderer Fördermaßnahmen oder die Lehrkräfte-Soll-Stunden (ebd.).

In der Stundentafel der kooperativen Gesamtschulen findet sich zudem die Vorgabe des fachübergreifenden und fächerverbindenden Unterrichts in den naturwissenschaftlichen Fächern (Physik, Chemie, Biologie) im fünften und sechsten Jahr-

gang mit vier Unterrichtsstunden pro Woche (Niedersächsisches Kultusministerium, 2015a). Darüber hinaus kann die Schulleitung Kooperationen zu anderen Schulen sowie außerschulischen Kooperationspartnern einrichten, um spezielle Lehr-Lern-Angebote zu schaffen. Die Erprobung neuer organisatorischer und pädagogischer Konzepte ist im Rahmen von Schulversuchen umsetzbar (vgl. Blossfeld et al., 2010).

Die Schulgesetzgebung bietet folglich den organisatorischen Rahmen, um die Entwicklung und Implementation von Profilunterricht in der Sekundarstufe I zu ermöglichen. Doch wie werden die im Schulgesetz vorgegebenen Richtlinien (z.B. erhöhte Stundenanzahl, fächerübergreifendes oder fächerverbindendes Unterrichten) sowie die Erwartungen an Profilunterricht (z.B. Wahldifferenzierung individueller Lernschwerpunkte) in der Unterrichtspraxis konkret umgesetzt? Damit beschäftigt sich das folgende Kapitel.

5.5 Musik- und MINT-Profilklassen – Unterrichtsgestaltung und Unterrichtsqualität

Wie aus der Beschreibung der rechtlichen Rahmenbedingungen zur Implementation von Profilklassen deutlich wurde, besteht für die weiterführenden Schulen in Niedersachsen relativ viel Spielraum bezüglich der konkreten unterrichtlichen Ausgestaltung des Profilangebots. Da die praktische Umsetzung des Profilunterrichts schulinternen Curricula folgt und die Unterrichtskonzepte Teil der spezifischen Profilierungsstrategie der Einzelschulen sind (vgl. Altrichter et al., 2016; Nonte, 2013), lassen sich die Aussagen zur konkreten Gestaltung des Profilunterrichts schwer generalisieren, auch wenn einige grundsätzliche Handlungs- und Gestaltungsmechanismen herauszustellen sind. Im Folgenden werden verschiedene Gestaltungsmerkmale sowie Unterrichtskonzepte der Musik- und MINT-Klassen vorgestellt und dabei Spezifika sowie Gemeinsamkeiten herausgearbeitet. Die Begriffe Musik- und MINT-(Profil)klassen werden dabei als übergeordnete Bezeichnungen für die vielfältigen Konzepte einer musikalischen beziehungsweise naturwissenschaftlichen Schwerpunktsetzung verwendet.

Gestaltung des musikalischen Profilunterrichts

Musik-Profilklassen werden allgemein als „Angebote [bezeichnet], in denen alle SchülerInnen einer Schulklasse ein Instrument lernen und als Klassenensemble zusammenspielen“ (Göllner, 2107, S. 15; vgl. Heß, 2005). Die Angebote werden im Rahmen des Pflichtunterrichts an allgemeinbildenden Schulen implementiert und beziehen sich folglich nicht auf freiwillige Zusatzangebote, wie Musik-AGs oder Schülerbands (Pabst-Krueger, 2020). Zentrales Gestaltungsmerkmal von Musik-Klassen ist das sogenannte *Klassenmusizieren*, das nach Bähr (2005) als „didaktisch-

methodisch geplante, gemeinsame Ausübung mit Gesang, Instrumentalspiel, Bewegung und Szene – einzeln bzw. in Kombinationen“ beschrieben wird (Bähr, 2005, S. 160). Gemeint sind damit vielfältige Formen und Aktivitäten des praktischen Musizierens, die im Klassenverband gemeinsam ausgeführt werden (Bähr, 2005; Pabst-Krueger, 2020). Klassenmusizieren kann im Kontext des sogenannten *erweiterten Musikunterrichts* implementiert werden. Von erweitertem Musikunterricht wird gesprochen, wenn das Fach Musik mit einem „über das reguläre Maß hinausgehende Stundenkontingent“ unterrichtet wird (Heß, 2017a, S. 27) – also mehr Unterrichtsstunden pro Woche im Fach Musik zur Verfügung stehen (s. Kapitel 5.4). In Musik-Profilklassen werden diese beiden Gestaltungsmerkmale vereint und mittels verschiedener organisatorischer Konzepte praktisch umgesetzt (Pabst-Krueger, 2020). Im Bereich des Musikprofilunterrichts in der Sekundarstufe haben sich verschiedene Konzepte etabliert: Neben dem Angebot der sogenannten Bläserklassen existieren auch Streicher-, Orchester- oder Bandklassen (Göllner, 2017; Heß, 2005; Pabst-Krueger, 2020).

Wie aus den rechtlichen Rahmenbedingungen (s. Kapitel 5.4) deutlich wurde, wird in Musik-Klassen das Unterrichtsfach Musik mit einem höheren Stundenkontingent von circa zwei zusätzlichen Stunden pro Woche (insgesamt vier Unterrichtsstunden) unterrichtet. Für niedersächsische Schülerinnen und Schüler in musikalischen Profilklassen berichten Haas, Nonte, Krieg und Stubbe (2019) eine durchschnittliche Unterrichtszeit in Musik von drei Stunden pro Woche und eine Stunde zusätzlichen Instrumentalunterricht. Alternativ ist eine Dreiteilung aus zwei Stunden Musikunterricht im Klassenverband (geleitet durch die Musiklehrkraft), einer Stunde Instrumentalunterricht (geleitet durch externe Musiklehrkräfte der Musikschulen) und einer Stunde Ensemblespiel (geleitet durch die Musiklehrkraft der Schule oder/und externe Musiklehrkraft) möglich (Bähr, Jank & Schwab, 2018).

Der Unterricht wird folglich häufig gemeinsam von Musiklehrkräften der eigenen Schule in Zusammenarbeit mit externen Musikschulen oder Musikpädagoginnen beziehungsweise Musikpädagogen gestaltet, die den Instrumentalunterricht übernehmen (Bähr & Schwab, 2002; Bähr et al., 2018; Pabst-Krueger, 2020; vgl. Lehmann-Wermser, Naacke, Nonte & Ritter, 2010). Der mit dem Konzept des Klassenmusizierens assoziierte hohe Praxisanteil in den Musik-Klassen kann „neben dem praktischen Musizieren [...] auch viele andere musikbezogene Aktionsformen [...], z.B. Hören, Analysieren, Beschreiben, Reflektieren, Malen, Tanzen, Szenisches Spiel uvm.“ beinhalten (Pabst-Krueger, 2020, S. 160).

Die einzelnen Konzepte des Klassenmusizierens und damit auch des Musik-Profilunterrichts unterscheiden sich jedoch hinsichtlich ihrer didaktischen Gestaltung sowie des Anteils des Instrumentalunterrichts, sodass „kaum von *der Praxis* als solcher [Hervorhebung im Original]“ in musikalischen Profilklassen gesprochen werden kann (Siedenburg & Harbig, 2018a, S. 2; vgl. Heß, 2005). Göllner (2017) stellt daher fest, dass es „insbesondere an Kenntnissen über die Umsetzung von Musikklassenangeboten [mangelt]“ (Göllner, 2017, S. 15). Im Sekundarstufenbereich ist das Konzept der Bläserklassen bundesweit am weitesten verbreitet (Heß, 2017a;

Pabst-Krueger, 2020; Verband deutscher Musikschulen, 2021) und daher auch vergleichsweise gut wissenschaftlich erforscht (vgl. u.a. Buchborn, 2011; Göllner, 2017; Heß, 2017a; Klose, 2018; Siedenburg & Harbig, 2018b). Da das Konzept der Bläserklassen zudem im überwiegenden Teil der in dieser Arbeit untersuchten Musik-Klassen Anwendung findet, wird das Konzept im Folgenden stellvertretend für weitere Musik-Profilangebote näher vorgestellt.

Mit Beginn der Bläserklasse in Klassenstufe fünf wählen alle Schülerinnen und Schüler ein Blasinstrument aus, das sie erlernen möchten – Vorkenntnisse sind dabei nicht nötig. Gemeinsam bilden die Schülerinnen und Schüler der Musik-Klasse ein Klassenensemble (Blasorchester). Der ergänzend zu den regulären Musikstunden angebotene Instrumentalunterricht findet im sogenannten Registerunterricht, das heißt nach Instrumenten getrennt, in Kleingruppen statt und wird meist von externen Musikschullehrkräften geleitet (Göllner, 2017; Pabst-Krueger, 2020). Die Unterrichtsinhalte des Instrumental- und Musikunterrichts sind durch ein gemeinsames Teilcurriculum inhaltlich verknüpft (Bähr et al., 2018). Zwischen den verschiedenen externen Musikschullehrkräften und den Musiklehrkräften der Schulen bedarf es hierbei einer engen Zusammenarbeit. Dabei kommen meist Teamteaching-Formate zum Einsatz (Göllner & Niessen, 2015), wobei die Musiklehrkraft der Schule die Leitung des Ensembles übernimmt und für die Lerninhalte verantwortlich ist (Yamaha, 2016), was nicht selten zu Spannungen zwischen den Lehrenden führt (Göllner, 2017). Kulin, Schwippert und Rieckmann (2015) stellten diesbezüglich im Rahmen des Projektes SIGrun (*Studie zum Instrumentalunterricht in Grundschulen*) heraus, dass ausreichend personelle Ressourcen, eine längerfristige Zusammenarbeit sowie funktionierende Kommunikation zwischen den beteiligten Lehrkräften (der Schule und der außerschulischen Kooperationspartner) besonders förderlich für die Zusammenarbeit sind.

Die in den Musik-Profilklassen verwendeten Unterrichtsmaterialien stammten zunächst vornehmlich aus den USA, wo das Konzept des Klassenmusizierens (auch vor der Einführung in Deutschland) bereits eine lange Tradition hat, und wurden anschließend sukzessive weiterentwickelt (Buchborn, 2011; Nonte & Schwippert, 2012). Im Zentrum der unterrichtlichen Gestaltung stehen meist Musikstücke, die im Instrumentalunterricht praktisch erarbeitet und im allgemeinbildenden Unterricht vielfältig (auch theoretisch) bearbeitet werden. Zum Teil kann auf vorgefertigte Unterrichtseinheiten zurückgegriffen werden, die hinsichtlich der Musik- und Materialauswahl sowie der didaktischen Gestaltung den staatlichen Vorgaben entsprechen (Bähr et al., 2018). Bähr et al. (2018) betonen jedoch den zeitlichen Aufwand, der mit der Erstellung geeigneter Unterrichtsmaterialien einhergeht. In den Bläserklassen finden aktuell unter anderem die von der Firma Yamaha vertriebenen Unterrichtsmaterialien (Lehrwerk ‚Essential Elements‘) verbreitet Anwendung (vgl. Yamaha, 2016, 2019). Die im Rahmen des Instrumental- und Ensembleunterrichts eingeübten Musikstücke können „aufführungsreif gestaltet werden“ (Bähr et al., 2018, S. 139), sodass als weiteres zentrales Gestaltungselement der Musik-Klassen die Teilnahme an mehrmals pro Schuljahr stattfindenden Auftritten benannt wird

(z.B. im Rahmen von Schulfesten, Tag der offenen Tür, Weihnachtsfeiern u.a.). Durch die Auftritte wird das Schuljahr gewissermaßen in kleinere Projektabschnitte unterteilt und die Schülerinnen und Schüler arbeiten im Unterricht auf das Projektziel (den Auftritt) hin (Göllner, 2017).

Gestaltung des naturwissenschaftlichen Profilunterrichts

Naturwissenschaftliche Profilklassen sind auch unter den Begriffen Entdecker-, Forscher- oder MINT-Klassen bekannt (vgl. Krieg et al., 2022), wobei die Bezeichnung MINT-Klasse im Folgenden als übergeordnete Bezeichnung für ebensolche Schulklassen mit spezifischen Schwerpunkten im naturwissenschaftlichen Bereich verwendet wird. Parallel zu den musikalischen Profilklassen stehen auch in den MINT-Klassen wöchentlich (meist zwei) zusätzliche Unterrichtsstunden zur Verfügung (s. Kapitel 5.4). Zudem wird fächerübergreifender Unterricht ermöglicht, der je nach schulspezifischer Konzeption Inhalte aus den Fächern Biologie, Chemie, Physik, Technik, Mathematik und Informatik in unterschiedlicher Schwerpunktsetzung thematisiert (Seidel et al., 2016). Im fächerübergreifenden Unterricht werden Fragestellungen mit „problemlösendem Charakter“ bearbeitet, die einen möglichst hohen Bezug zur Alltagswelt der Schülerinnen und Schüler aufweisen (Seidel et al., 2016, S. 11). Die Bearbeitung erfolgt multiperspektivisch und nutzt dabei unterschiedliche methodische Herangehensweisen der MINT-Fächer. Für die Konzeption und Durchführung von fächerübergreifendem Unterricht arbeiten die jeweiligen Fachlehrkräfte der verschiedenen Disziplinen idealerweise eng zusammen (ebd.; s. dazu auch das Konzept des naturwissenschaftlichen Anfangsunterrichts; Stäudel, 2018).

Darüber hinaus ist der Unterricht in MINT-Klassen durch einen hohen Praxisbezug gekennzeichnet: Er kann „[...] Raum zum Experimentieren bzw. Forschendem Lernen bieten, kooperative Lernformen nutzen und dadurch abwechslungsreich gestaltet werden, digitale Medien ebenso wie alternative Prüfungsformate einsetzen und dadurch den durch die Schülerinnen und Schüler oft in Frage gestellten Lebensweltbezug näherbringen“ (Seidel et al., 2016, S. 10). Das Lehrkonzept *Forschendes Lernen* ist vor allem durch ein „Learning by Doing“ gekennzeichnet (Seidel et al., 2016, S. 12), wobei die Lernenden „[...] den Prozess eines Forschungsvorhabens, [...] von der Entwicklung der Fragen und Hypothesen über die Wahl und Ausführung der Methoden bis zur Prüfung und Darstellung der Ergebnisse in selbstständiger Arbeit oder in aktiver Mitarbeit in einem übergreifenden Projekt (mit)gestalten, erfahren und reflektieren“ (Huber, 2014, S. 25; vgl. Huber, 2009). Beim Forschenden Lernen steht folglich vor allem Selbstständigkeit und Eigenverantwortung sowie Teamfähigkeit und Problemlösekompetenz der Schülerinnen und Schüler bei der Bearbeitung und Reflektion von Unterrichtsinhalten im Vordergrund (Seidel et al., 2016; s.a. Kapitel 5.3 zu den Zugangsvoraussetzungen in MINT-Klassen). Kooperative Lernformate (z.B. Gruppenarbeit in kleinen Teams) können nach verschiedenen Kriterien, wie Leistungshomogenität beziehungsweise

-heterogenität oder thematischen Schwerpunkten, umgesetzt werden und bieten so unter anderem die Möglichkeit der internen Leistungsdifferenzierung sowie der Wahl interessengeleiteter Lernschwerpunkte. Als Lernorte außerhalb des Klassenzimmers, zum Beispiel im Rahmen von Exkursionen oder Projekttagen beziehungsweise -wochen, können Museen und Science Center sowie Schülerlabore oder Schülerforschungszentren genutzt werden (Lorke, 2018; Schwarzer & Parchmann, 2018; Seidel et al., 2016; Wilde, Bätz, Kovaleva & Urhahne, 2009). Zudem bestehen zwischen MINT-Profilschulen häufig längerfristige Kooperationen zu Einrichtungen des tertiären Bildungsbereichs, wie Universitäten oder Fachhochschulen, sowie zu externen Forschungseinrichtungen (Seidel et al., 2016). Ein weiteres zentrales Element der MINT-Klassen bildet die Teilnahme an naturwissenschaftlichen Schülerwettbewerben (z.B. Jugend forscht, Robotik-Wettbewerbe, Biologie-, Physik- oder Chemieolympiade), die als externe Anreize die Schülerinnen und Schüler projektbezogen für bestimmte Themenbereiche begeistern können und darüber hinaus viele Aspekte des Forschenden Lernen (selbstständige Wissensaneignung, Teamarbeit, Präsentation der Ergebnisse) aufgreifen (Seidel et al., 2016). Der Unterricht in MINT-Klassen wird daher vornehmlich als schülerzentriert beschrieben, in dem die Lehrkraft die Rolle eines/r begleitende/n Trainerin oder Trainers einnimmt (Seidel et al., 2016).

Doch wie schätzen die Schülerinnen und Schüler den Profilunterricht tatsächlich ein? Diese Frage kann mit Hilfe von Studien zur wahrgenommenen Unterrichtsqualität in Profilklassen beantwortet werden.

Unterrichtsqualität in Profilklassen

Auf Basis der vorangegangenen Beschreibung der Unterrichtskonzepte in Musik- und MINT-Klassen wurde sichtbar, dass in diesen Klassen von besonderen Lehr-Lern-Arrangements gesprochen werden kann. Doch spiegelt sich dies auch in der von den Schülerinnen und Schülern wahrgenommenen Unterrichtsqualität wider? Als die drei Basisdimensionen guten Unterrichts werden eine „strukturierte, klare und störungspräventive Unterrichtsführung“, ein „unterstützendes, schülerorientiertes Sozialklima“ sowie die „kognitive Aktivierung“ bezeichnet (Klieme & Rakoczy, 2008, S. 228). Das konstruktive, unterstützende Verhalten durch die Lehrkraft wird hierbei als besonders förderlich für die Motivation beschrieben, wohingegen die kognitive Aktivierung „für systematischen Wissensaufbau und Verstehen [...] relevant [ist]“ und gute Klassenführung „die Voraussetzung für beides darstellt“ (Klieme & Rakoczy, 2008, S. 228). Nonte, Haas und Stubbe (2019) untersuchten anhand der drei Basisdimensionen in fünf niedersächsischen MINT-Klassen des fünften und sechsten Jahrgangs sowie fünf direkten Parallelklassen ohne spezielle Profilierung derselben Jahrgänge, inwiefern sich die Wahrnehmung der Unterrichtsqualität zwischen profilierten und nicht-profilieren Klassen unterscheidet. Wie die Ergebnisse zeigen, schätzten die Schülerinnen und Schüler in den

MINT-Klassen die Klassenführung sowie die kognitive Aktivierung signifikant besser ein, als die Schülerinnen und Schüler in den Parallelklassen ohne Profil. Unter der Kontrolle individueller Merkmale, wie dem Geschlecht, dem sozioökonomischer Hintergrund, der fachbezogenen Kompetenz oder dem naturwissenschaftlichen Fähigkeitsselbstkonzept und Sachinteresse, ließen sich jedoch keine Unterschiede hinsichtlich der wahrgenommenen Unterrichtsqualität zwischen MINT-Klassen und Klassen ohne Profil feststellen. Allerdings zeigte sich im Analysemodell mit Kovariaten, dass die Schülerinnen und Schüler in MINT-Klassen, die ein hohes naturwissenschaftliches Fähigkeitsselbstkonzept aufweisen, signifikant mehr unterstützendes Verhalten wahrnehmen. Nonte et al. (2019) kommen daher zu dem Schluss, dass die Wahrnehmung von Unterrichtsqualität essentiell mit individuellen Lernvoraussetzungen (wie dem Fähigkeitsselbstkonzept) der Schülerinnen und Schüler in Zusammenhang steht.

Die Unterrichtsqualität in Musik-Klassen des fünften und sechsten Jahrgangs untersuchten Haas et al. (2019) und verglichen diese mit der wahrgenommenen Unterrichtsqualität in nicht-profilierten Parallelklassen derselben Schule. Wie bereits für die MINT-Klassen dargestellt, berichteten auch die Schülerinnen und Schüler in Musik-Klassen über eine signifikant bessere Klassenführung als ihre Mitschülerinnen und Mitschüler der Parallelklassen ohne Profil. Der Effekt blieb auch nach Kontrolle von Kovariaten (u.a. Geschlecht, sozioökonomischer Hintergrund, Musikkompetenz, musikalisches Fähigkeitsselbstkonzept und Sachinteresse) bestehen. Darüber hinaus konnten Haas et al. (2019) zeigen, dass Schülerinnen und Schüler, die über hohes fachbezogenes Fähigkeitsselbstkonzept verfügen, signifikant häufigerangaben, unterstützendes Verhalten durch die Lehrkraft wahrzunehmen. Sowohl für die MINT-Klassen als auch für die Musik-Klassen der Sekundarstufe I konnten folglich signifikante Effekte zwischen individuellen Faktoren, wie dem fachbezogenen Fähigkeitsselbstkonzept, und einzelnen Facetten der Unterrichtsqualität gezeigt werden (Haas et al., 2019; Nonte et al., 2019).

Eder (2011) beschrieb für österreichische Musik-Klassen des siebten und achten Jahrgangs ähnlich positive Ergebnisse wie Haas et al. (2019) zur wahrgenommenen Unterrichtsqualität in musischen Profilklassen. Auch Eder (2011) beobachtete in den untersuchten Musik-Klassen ein deutlich positiveres Klassenklima als in den nicht-profilierten Vergleichsklassen, wobei das Konstrukt Klassenklima verschiedene Aspekte der Unterrichtsqualität, wie schülerzentrierter Unterricht, Vermittlungsqualität, das pädagogische Engagement der Lehrkraft oder die Vermeidung von Rivalität und Störung, zu einem Gesamtwert zusammenfasst. Eder (2011) betont, der Unterricht in den Musik-Klassen sei besser auf die Schülerinnen und Schüler und ihre Bedarfe abgestimmt, weise weniger Störung und Rivalität auf und biete eine günstige Lerngemeinschaft in der Klasse. Ähnliche Befunde für den Besuch von Schulen mit musikischem Schwerpunkt finden sich auch bei Bastian (2000) sowie Weber, Spychiger und Patry (1993), die zu einer der frühesten deutschsprachigen Evaluationsstudie zum musischen Profilverricht gehört. Im zugrundeliegenden schweizer Projekt *Musik in der Schule* wurde in annähernd 50 Schulklassen des ersten

bis sechsten Jahrgangs in allen Kantonen der Schweiz ein Schulversuch zum erweiterten Musikunterricht vorgenommen und wissenschaftlich begleitet. Die sogenannten ‚Versuchsklassen‘ erhielten eine zusätzliche musikalische Förderung in Form von wöchentlichen fünf (statt zwei) Unterrichtseinheiten in Musik. Neben dem bereits berichteten positiven Effekten auf das Unterrichtsklima in den Versuchsklassen (Musik-Klassen) zeigte sich zudem ein stärkerer Gruppenzusammenhalt, weniger Konkurrenz aber auch eine höhere Disziplin im Musikunterricht. Zudem betonen Weber et al. (1993) besonders die Entwicklung der Motivation der Schülerinnen und Schüler in den Versuchsklassen (gemessen an der Einstellung gegenüber der Schule im Allgemeinen und dem Musikunterricht im Speziellen sowie der Einstellung gegenüber musikalischen Themen unabhängig vom Schulkontext).

Es stellt sich folglich die Frage, welche pädagogischen Ziele oder Potentiale mit Musik- und MINT-Profilklassen verbunden werden und in welchem Zusammenhang der Besuch einer Profilklass mit individuellen lernrelevanten Faktoren, wie der Entwicklung der fachspezifischen Interessen und domänenspezifischen Fähigkeitsselbstkonzepte, stehen? Damit beschäftigt sich das folgende Kapitel.

5.6 Potentiale von Profilklassen – empirische Befunde

In der Beschreibung der Unterrichtskonzepte der Musik- und MINT-Klassen im vorherigen Kapitel wurde deutlich, dass zentrale Gestaltungsmerkmale sowohl für Profilangebote im musikalischen als auch im naturwissenschaftlichen Bereich gelten. Als charakteristische Merkmale von Musik- und MINT-Klassen können demnach zusätzliche Unterrichtsstunden, praxisbezogene Lehr-Lern-Formate, der Einbezug externer Lehrkräfte beziehungsweise Lernorte sowie die Einbettung der Unterrichtsinhalte in öffentlichkeitswirksame Aktivitäten (z.B. Teilnahme an Auftritten oder Wettbewerben) herausgestellt werden. Jedes Gestaltungsmerkmal für sich ist mit bestimmten pädagogischen Potentialen assoziiert, die in den Unterrichtskonzepten der Profilklassen gebündelt genutzt werden können. Ganz allgemein wird in der Forschungsliteratur der Besuch einer Profilklass mit der Förderung der allgemeinen schulischen Leistungen und des Sozialverhaltens verbunden. Darüber hinaus werden positive Effekte auf das domänenspezifische Interesse sowie das Fähigkeitsselbstkonzept der Schülerinnen und Schüler erwartet (vgl. Eder, 2011; Göllner, 2017; Seidel et al., 2016). Differenzierte Untersuchungen zu Effekten von Profilklassen in der Sekundarstufe I auf die Entwicklung des domänenspezifischen Interesses und des Fähigkeitsselbstkonzepts der Schülerinnen und Schüler sind aktuell jedoch nur sehr begrenzt vorhanden (Eder, 2011; Koivuhovi et al., 2021; Krieg & Stubbe, 2018). Hingegen sind die einzelnen Gestaltungsmerkmale der Profilangebote hinsichtlich ihrer pädagogischen Potentiale vergleichsweise gut beforscht und können näherungsweise einen Beitrag zum besseren Verständnis möglicher Effekte der Profilklassen leisten, sofern angenommen wird, dass die Einzelmaßnahmen in

den Profilangeboten ähnliche Wirkungen entfalten. Im Folgenden werden die verschiedenen Gestaltungsmerkmale der Profilklassen hinsichtlich ihrer postulierten Potentiale dargestellt, wobei der Fokus auf mögliche Effekte auf das Interesse sowie das Fähigkeitsselbstkonzept der Schülerinnen und Schüler in den Domänen Musik und Naturwissenschaften gelegt wird.

Praxisbezogene Lehr-Lern-Formate im Musikunterricht

Wie aus der Beschreibung der Unterrichtsgestaltung in den Musik-Klassen deutlich wurde (s. Kapitel 5.5), stellt das gemeinsame praktische Musizieren im Klassenverband ein zentrales Charakteristikum dar (Göllner, 2017; Heß, 2011a; Heß, 2017a; Pabst-Krueger, 2020). In der Musikdidaktik sind praxisorientierte Unterrichtskonzepte wie das *Klassenmusizieren* auch unter der Bezeichnung des *tätigkeitsorientierten, aufbauenden* beziehungsweise *erfahrungsorientierten Musikunterrichts* bekannt (Heß, 2011a; vgl. Göllner, 2017; Jank & Stroh, 2006; Jünger, 2013).

In einem Vergleich zwischen Schülern⁸ des fünften bis zwölften Jahrgangs an Gymnasien mit musischem Schulprofil gegenüber Schülerinnen und Schülern an unprofilieren Gymnasien konnten Fritzsche et al. (2011) zeigen, dass die Schüler an den musisch-profilieren Gymnasien tatsächlich deutlich häufiger musikalische Aktivitäten ausüben als Schülerinnen sowie Schüler an nicht-profilieren Schulen. Die musikalische Aktivität bezog sich dabei auf die wöchentliche Dauer des besuchten Instrumental- oder Gesangsunterrichts, die durchschnittliche außerschulische instrumentale Übezeit, den zeitlichen Umfang der wöchentlichen Ensembleproben sowie die Häufigkeit der Teilnahme an Auftritten mit dem Klassenensemble. Schüler an musisch-profilieren Gymnasien wiesen in allen vier Kriterien die höchsten Werte auf (Fritzsche et al., 2011). Vor allem das gemeinsame Musizieren in den Musik-Klassen förderte demnach nachhaltig die Bereitschaft und die Begeisterung für musikalische Aktivitäten, das heißt des persönlichen Wertes gegenüber Musik (Göllner, 2017). Auf Basis eines schwedischen Samples von $n = 3820$ Erwachsenen konnten zudem Theorell, Lennartsson, Madison, Mosing und Ullén (2015) herausstellen, dass Schülerinnen und Schüler, die während der Schulzeit zusätzliche Musikstunden erhalten haben, auch im Erwachsenenalter häufiger musikalisch aktiv sind und beispielsweise ein Instrument spielen.

Das Engagement der Schülerinnen und Schüler über einen längeren Zeitraum ein Instrument zu erlernen, wird laut Forschungsbefunden mit einer Stärkung des Durchhaltevermögens und der Selbstdisziplin belohnt (Heß 2017a; vgl. Feuerborn, Hollenstein, Janssen, Kürner & Staub, 2002). Aber auch negative Effekte im Hinblick auf das musikbezogene Selbstkonzept konnten im Zusammenhang mit einer „mühevollen, intensiven Auseinandersetzung mit dem Instrument“ beobachtet werden (Kranefeld, Busch & Dücker, 2015, S. 64; vgl. Busch, 2013). Die Bildung

⁸ In der Studie von Fritzsche et al. (2011) wurde explizit die musische Aktivität von Schülern an Gymnasien mit Musikprofil untersucht. Als Vergleichsgruppen wurden zwei Teilstichproben herangezogen: (a) Schülerinnen an unprofilieren Gymnasien und (b) Jungen an unprofilieren Gymnasien.

eines Klassenensembles und das gemeinsame Musizieren im Klassenverband kann zudem zur Förderung der Teamfähigkeit, der sozialen Kompetenzen sowie des Klassenzusammenhalts beitragen (Heß, 2017a; vgl. Eder, 2011). Bähr und Schwab (2002) berichten diesbezüglich in ihrer Evaluationsstudie eines Modellversuchs zu praxisorientiertem Musikunterricht, der in Kooperation mit einer externen Musikschule gestaltet wurde, dass die Schülerinnen und Schüler mit spezieller musikalischer Förderung in Musik erheblich besser sozial integriert waren als Kinder ohne eine solche spezielle Förderung.

Heß (2017a) legt auf Basis einer umfangreichen Recherche von Informationsmaterialien zu Bläserklassen eine Übersicht zu den verschiedenen Zielsetzungen von Musik-Profilklassen vor. Auch Heß (2017a) betont, dass viele der mit den Bläserklassen verbundenen Potentiale nicht (ausschließlich) mit dem Musiklernen an sich verknüpft sind, sondern vielmehr einen „alltags- und zukunftsrelevanten Nutzen“ aufweisen (Heß, 2017a, S. 79). Auf der Ebene der Schülerinnen und Schüler benennt Heß (2017a) soziale Effekte (Förderung der Teamfähigkeit, Rücksichtnahme), kognitive Effekte (Steigerung der Konzentrations- und Leistungsfähigkeit) sowie positive Effekte auf die Persönlichkeitsentwicklung (Stärkung des Selbstbewusstseins und des Selbstwertgefühls sowie der Ausdauer, Verlässlichkeit, Zielorientierung und der individuellen Identität). Darüber hinaus wird „Spaß/Freude am (gemeinsamen) Musizieren“ sowie eine gesteigerte Motivation angeführt, die durch die Teilnahme an Auftritten zusätzlich positiv beeinflusst werden kann (Heß, 2017a, S. 81; s.a. dieses Kapitel weiter unten).

Quantitative Daten stützen ebenfalls die These des positiven Zusammenhangs zwischen einem praxisorientierten Musikunterricht und der Stärkung des Fachinteresses der Schülerinnen und Schüler. So konnte Heß (2011a) für regulären, nicht-profilierten Musikunterricht der Sekundarstufe I und II (5. bis 13. Jahrgangsstufe) zeigen, dass Schülerinnen und Schüler, die aktiv musizieren, über ein höheres musikalisches Fachinteresse verfügen. Zudem stellte die Autorin einen signifikant positiven Zusammenhang zwischen der inhaltlichen Gestaltung und der Bewertung des Musikunterrichts durch die Schülerinnen und Schüler fest. Die stärksten korrelativen Effekte unter den aktiv musizierenden Schülerinnen und Schülern waren dabei für die Gestaltungsmerkmale ‚Musikmachen‘ und ‚Hören‘ zu beobachten (Heß, 2011a). Fritzsche et al. (2011) untersuchten darüber hinaus, inwiefern die musikalische Aktivität von Schülern an musisch-profilierten Gymnasien mit dem Sachinteresse in Musik (intrinsicischer Wert an musikalischer Partizipation) sowie dem musikalischen Selbstkonzept zusammenhängt. Als signifikanten Indikator für die musikalische Aktivität stellte sich nur das musikalische Selbstkonzept der Schüler an Schulen mit Musik-Profil heraus, nicht aber deren musikalisches Sachinteresse. Diesen Befund stützen auch Bernecker, Haag und Pfeiffer (2006), die bei musikalisch aktiven Gymnasiastinnen und Gymnasiasten ein höheres musikalisches Selbstkonzept beobachteten als bei nicht-aktiven Schülerinnen und Schülern. McPherson und McCormick (2006) benennen hingegen im Rahmen einer Untersuchung von australischen Schülerinnen und Schülern des ersten bis achten Jahrgangs

das musikalische Selbstkonzept als einen entscheidenden Faktor, um eine Teilnahme am Instrumentalunterricht vorauszusagen. Nonte und Schurig (2019) wiederum konnten keinen Zusammenhang zwischen dem in der dritten Jahrgangsstufe erhobenen musikalischen Selbstkonzept von Schülerinnen und Schülern, die in der Primarstufe am instrumentalen Förderprogramm *Jedem Kind ein Instrument* (JeKi) teilgenommen haben, und einer späteren Teilnahme am Instrumentalunterricht in der Sekundarstufe I (6. und 7. Jahrgangsstufe) feststellen.

Die Forschungsergebnisse zum Zusammenhang zwischen praxisorientiertem Musikunterricht und dem musikalischen Fähigkeitsselbstkonzept der Schülerinnen und Schüler sind demnach eher als kontrovers zu beschreiben: So stellte Bastian (2000) in seiner viel beachteten und kritisch diskutierten Studie⁹ *Musik(erziehung) und ihre Wirkung* fest: „Das Selbstkonzept ist eindeutig unabhängig von der Tatsache, ob Kinder eine »erweiterte Musikerziehung« [Hervorhebung im Original] erleben oder nicht. [...] Musikerziehung hat keinen Einfluss auf das Selbstkonzept von Kindern im Alter zwischen sieben/acht und elf/zwölf Jahren“ (Bastian, 2000, S. 453). Bastian (2000) schloss daraus, dass sich entsprechende Effekte des aktiven Musizierens auf das Fähigkeitsselbstkonzept der Schülerinnen und Schüler erst nach Beendigung der Grundschule einstellen. Diesen Befund untermauern auch Nonte und Schwippert (2012), die (im Rahmen des Projektes SIGrun¹⁰) das musikalische Fähigkeitsselbstkonzept von Schülerinnen und Schülern der Primarstufe untersuchten, die an musisch-profilierten Schulen am Programm *Jedem Kind ein Instrument* (JeKi)¹¹ teilnahmen. Weder für das Selbstkonzept Singen¹² noch für das Selbstkonzept Musizieren konnten die Autorin und der Autor einen Zusammenhang mit dem Besuch des musikalischen Förderprofils JeKi feststellen. Allerdings zeigte sich ein signifikanter Effekt zwischen der wöchentlichen Übedauer auf dem Instrument und dem Selbstkonzept im Bereich Musizieren (Nonte & Schwippert, 2012). Somit lässt sich ableiten, dass nicht der Besuch des Förderprogramms JeKi an sich, sondern stattdessen der Anteil des aktiven Musizierens das musikalische Selbstkonzept begünstigen könnte.

⁹ vgl. u.a. Nonte und Schurig (2019)

¹⁰ *Studie zum Instrumentalunterricht in Grundschulen* (SIGrun)

¹¹ Für eine Übersicht der im JeKi-Kontext entstandenen Forschungsergebnisse siehe Kranefeld (2015) sowie Schwippert, Lehmann-Wermser und Busch (2019).

¹² Im *Shavelson-Modell* (Shavelson et al., 1976), das die verschiedenen Selbstkonzeptdimensionen beschreibt (s. Kapitel 3.2), ist das musikalische Fähigkeitsselbstkonzept ursprünglich nicht als eigene Selbstkonzeptfacette enthalten gewesen. In der musikpädagogischen Forschung entwickelte sich daher eine eigene Konzeption, in der das musikalische Selbstkonzept als mehrdimensionales Konstrukt beschrieben wird (Spychiger, 2013). Neben einem akademischen Bereich des musikalischen Selbstkonzepts (mit eigenen Unterdimensionen) wird ein nicht-akademischer Bereich (mit eigenen Subfacetten) unterschieden. Der akademische Bereich umfasst die Fähigkeitsselbsteinschätzung verschiedener musikalischer Aktivitäten (z.B. Singen, Musizieren, Komponieren etc.; für eine Übersicht s. Fiedler & Hasselhorn, 2020).

Dies bestätigen auch Kranefeld et al. (2015), die ebenfalls im Kontext der JeKi-Begleitforschung instrumentalen Gruppenunterricht in der Grundschule (Projekt BEGIn¹³) und dessen Effekte für das Selbstkonzept von Schülerinnen und Schülern untersuchten. Den stärksten Zusammenhang fanden Kranefeld et al. (2015) zwischen ‚Gefallen am Üben‘ und dem musikalischen Selbstkonzept im Bereich ‚Musik machen‘. Kinder, die gern auf ihren Instrumenten üben, wiesen demnach auch eine höhere individuelle Selbsteinschätzung bezüglich ihrer Kompetenz im ‚Musik machen‘ auf (Kranefeld et al., 2015). Zudem zeigte sich ein indirekter Effekt zwischen Übehäufigkeit, Gefallen am Üben und dem musikalischen Selbstkonzept. Wer also gern übt, tut dies öfter, wovon schließlich auch das musikalische Selbstkonzept profitiert. Im Sinne der EVT-Determination (vgl. Kapitel 4.1) könnte der Indikator ‚Gefallen am Üben‘ auch als Freude oder Spaß am Üben und damit als intrinsischer Wert (Interesse) übersetzt und verstanden werden. Demnach wären positive Effekte zwischen dem persönlichen Wert/Interesse am aktiven Musizieren und dem musikalischen Selbstkonzept zu erwarten, sofern diese angemessen operationalisiert und gemessen werden. Kranefeld et al. (2015) stellten weiterhin fest, dass die Teilnahme an außerschulischem Instrumentalunterricht sowohl mit dem Selbstkonzept ‚Musik machen‘ als auch dem Selbstkonzept ‚Singen‘ positiv zusammenhängt. Das Ausscheiden aus dem JeKi-Programm (und damit die Beendigung des erweiterten Musikunterrichts) im Laufe der Grundschule zeigte hingegen einen signifikant negativen Effekt auf die Entwicklung des musikalischen Selbstkonzepts (Busch & Kranefeld, 2013; Kranefeld et al., 2015).

Nationale ebenso wie internationale Forschungsbefunde verweisen zudem hinsichtlich der bevorzugten musikalischen Aktivitäten auf geschlechtsspezifische Unterschiede (vgl. Green, 2002; Hall, 2005; Harrison, 2007; Heß, 2011a; Nonte et al., 2021; Warzecha, 2013). Demnach sind ‚sanftere‘ musikalische Aktivitäten, wie Singen oder das Spielen eines Orchesterinstrumentes eher weiblich konnotiert, während sich Jungen eher für laute, elektronische Instrumente und Popmusik (oft auch im außerschulischen Kontext) interessieren (Abramo, 2011; Green, 2002; Harrison, 2005). Auch hinsichtlich der Instrumentenwahl konnten Geschlechterunterschiede beobachtet werden: Während Mädchen häufig die Instrumente Geige, Flöte und Klarinette spielten, bevorzugten Jungen die Instrumente Schlagzeug, Posaune und Trompete (Abeles, 2009; Green, 2002; Harrison, 2005). Spielten Mädchen und Jungen hingegen gemeinsam in einer Band, so wählten Mädchen häufiger geschlechtsatypische Instrumente als Jungen (Abeles, 2009) und jüngere, bisher unerfahrene Kinder waren in der Wahl ihrer Instrumente offener als ältere Musikerinnen und Musiker (Wrape, Dittloff & Callahan, 2014). Dies könnte bei der inhaltlichen Konzeption und zeitlichen Verortung in der Schullaufbahn des Musik-Klassen-Angebots berücksichtigt werden (vgl. McGregor & Mills, 2006).

¹³ Bielefelder Studie zum Gruppen-Instrumentalunterricht (BEGIn)

Teilnahme an Auftritten im Klassenensemble

Wie im vorherigen Kapitel 5.5 dargestellt, wird das gemeinsame Auftreten (z.B. im Rahmen von Schulveranstaltungen, Tag der offenen Tür, Weihnachtsfeiern u.a.) als zentrales Element der Musik-Klassen benannt. Die Auftritte finden regelmäßig und mehrmals pro Schuljahr statt und dienen somit einerseits der Strukturierung des Schuljahrs, sind andererseits aber auch Mittel zur Außendarstellung und Imagebildung der Schule (Göllner, 2017; Nonte, 2022). Für die Schülerinnen und Schüler werden durch die Auftritte die Unterrichtsinhalte in einen realen praktischen Rahmen eingebettet beziehungsweise mit einem konkreten Ziel verknüpft, auf das die Musik-Klasse gemeinsam hinarbeitet (Göllner, 2017). Die Veranstaltung von Auftritten oder Konzerten und die dabei erlebte positive Rückmeldung von außen (Applaus; Anerkennung) wird mit positiven Effekten für die Motivation der Schülerinnen und Schüler sowie der Gruppendynamik in der Klasse verknüpft (Jank & Stroh, 2006; Pabst-Krueger, 2020; vgl. Terhag, 2012). So stellte Göllner (2017) in seiner Interviewstudie zu Bläserklassen wiederholt das Motiv des „Erlebens einer Klassengemeinschaft“ als sozial verbindendes Element in den Vordergrund der Zielsetzung von Bläserklassen (Göllner, 2017, S. 201). Die Schülerinnen und Schüler schätzten besonders, „Teil einer Gruppe zu sein, gemeinsam Auftritte zu absolvieren und positive Erfahrungen im gemeinsamen Musizieren zu sammeln“ (Göllner, 2017, S. 184). Das daraus entstehende Gemeinschaftsgefühl in der Musik-Klasse war demnach stärker wahrnehmbar als den Parallelklassen ohne Musik-Profil. Zudem gebe es „keine Außenseiter“, die von der Klassengemeinschaft ausgeschlossen werden (Göllner, 2017, S. 211). Davon profitiert neben der Arbeitsatmosphäre auch die Lern- und Hilfsbereitschaft untereinander, vor allem aber die Motivation der Schülerinnen und Schüler (Göllner, 2017; vgl. Jank & Stroh, 2006, Weber et al., 1993).

Praxisbezogene Lehr-Lern-Formate im MINT-Unterricht

In ihrer theoretischen Konzeption sind praxisorientierte Lehr-Lern-Formen im deutschsprachigen Raum auch unter den Begriffen *Forschendes Lernen*, *Forschungsorientiertes Lernen*, *Forschungsbasiertes Lernen* sowie in der internationalen Forschungsliteratur als *inquiry-based learning* oder *context-based approaches* bekannt (vgl. Bennett, Lubben & Hogarth, 2007; Habig et al., 2018; Huber, 2014; Wulf, Haberstroh & Petersen, 2020). Entsprechende Ansätze finden sich von der Primarstufe bis zum tertiären Bildungsbereich sowie in allen Fachdisziplinen, unterscheiden sich aber hinsichtlich des Aktivitätsniveaus der Lernenden (Messner, 2012; Wulf et al., 2020).

Zahlreiche nationale und internationale Studien haben sich bereits den Potentialen von praxis-, projekt- oder kontextorientierten Unterrichtskonzepten sowie Forschendem Lernen in unterschiedlichen Altersstufen und Fachdisziplinen gewidmet (für eine Übersicht s. Potvin & Hasni, 2014): Neben einem gesteigerten Interesse am Forschen und Experimentieren (Gess, Rueß & Deicke, 2014; Höffler,

Lüthjohann & Parchmann, 2014), konnten ebenfalls positive Effekte auf die Einstellungen der Lernenden gegenüber Naturwissenschaften allgemein sowie naturwissenschaftlichen Themen im schulischen Kontext herausgestellt werden (Bennett et al., 2007; Dolenc & Kazanis, 2020; Gibson & Chase, 2002). Zudem wird die Förderung von sogenannten Soft-Skills, das heißt beispielsweise Präsentationstechniken, Problemlösekompetenz, eigenverantwortliches Lernen sowie Kommunikations- und Teamfähigkeit, prognostiziert (Wulf, Thiem & Gess, 2020; Huber, 2009).

Wie Holmeier (2017) in ihrer Evaluationsstudie eines MINT-Klassen-Konzepts eines Schweizer Gymnasiums zeigen konnte, bewerteten die Schülerinnen und Schüler die Praxisphasen als jenes Gestaltungsmerkmal, das ihnen insgesamt an der MINT-Klasse am besten gefällt. Neben dem hohen Praxisanteil beurteilten die Schülerinnen und Schüler außerdem positiv, dass ihre Interessen im Unterricht der MINT-Klasse aufgegriffen werden (Holmeier, 2017). Doch wie wird praxisnahes Unterrichten in den MINT-Fächern konkret umgesetzt?

Zum einen steht die Gestaltung von authentischem Unterricht sowie die Herstellung von mehr Lebensweltbezug und Aktualität im Vordergrund (vgl. kontextorientierte Formate; Habig et al., 2018; van Vorst, Fechner & Sumfleth, 2012). Das bedeutet, Themen, thematische Kontexte oder naturwissenschaftliche Fragestellungen zu wählen, die die Schülerinnen und Schüler in ihrer aktuellen Alltagsrealität mit ihren Perspektiven und Interessen abholen und aus denen eine persönliche Relevanz (ggf. auch für ihre berufliche Zukunft) deutlich wird (Bennett et al., 2007; Höttecke, 2012; Mitchell, 1993; Seidel et al., 2016). Bereits Häußler und Hoffmann (1995) konnten diesbezüglich im Rahmen der IPN-Interessenstudie (vgl. Hoffmann et al., 1998) feststellen, dass das Interesse der Schülerinnen und Schüler an Physik (Sachinteresse) höher ausgeprägt ist, wenn der Unterrichtsinhalt in „anwendungsbezogene Kontexte eingebettet ist“ und einen „lebenspraktischen Nutzen“ aufweist (Häußler & Hoffmann, 1995, S. 112f.; vgl. Kapitel 2.5).

Zum anderen wird den Schülerinnen und Schülern durch eigenes Ausprobieren und selbst durchgeführte Experimente beziehungsweise Projekte selbstständiges, eigenverantwortliches Lernen ermöglicht (vgl. Kapitel 5.5; Messner, 2012; Seidel et al., 2016). Schon Palmer (2009) konnte in seiner Untersuchung zu forschungsorientiertem Naturwissenschaftsunterricht für australische Schülerinnen und Schüler der neunten Jahrgangsstufe zeigen, dass das situationale Interesse¹⁴ der Schülerinnen und Schüler während der Experimentierphasen höher war als in passiveren Lernphasen. Ergänzend stellten Swarat, Ortony und Revelle (2012) in ihrer Untersuchung von Mittelstufenschülerinnen und -schülern in den USA heraus, dass die Einschätzung der Interessantheit vor allem von der Art der Aktivität und weniger vom Unterrichtsinhalt oder dem Lernziel abhing. Praxisnahe Unterrichtsformen,

¹⁴ Palmer (2009) untersuchte Effekte auf das sogenannte *situationale Interesse*, das im Gegensatz zum zeitlich stabileren *individuellen Interesse* durch externe Faktoren hervorgerufen und als temporärer, situationsspezifischer motivationaler Zustand beschrieben wird (s. Kapitel 2.3 sowie Krieg, Stubbe, Tarelli & Willems, 2021).

die aktives Experimentieren oder den Umgang mit Technik ermöglichen, könnten demnach verstärkt Interesse bei den Schülerinnen und Schülern wecken. Die Autorin und Autoren schlossen daraus, dass sich positive Effekte auf das Interesse an Naturwissenschaften relativ leicht erreichen ließen, sofern die Lernumgebungen aktiver und interessanter gestaltet werden (Swarat et al., 2012). Sowohl Palmer (2009) als auch Swarat et al. (2012) maßen den Grad des Interesses jedoch als Momentaufnahme nach einer kurzen praxisorientierten naturwissenschaftlichen Unterrichtseinheit (situationales Interesse). Die Ergebnisse lassen demnach keine Schlüsse über Langzeiteffekte zu.

Höffler et al. (2014) untersuchten hingegen kontextorientierten, fächerübergreifenden MINT-Unterricht über die ersten Jahre der Sekundarstufe I hinweg (5. bis 7. Jahrgangsstufe) und differenzierten zudem zwischen dem naturwissenschaftlichen Fach- und Sachinteresse. Aufgrund des bereits vielfach berichteten Interessenrückgangs während der Sekundarstufe I, insbesondere auch in den MINT-Fächern (vgl. u.a. Daniels, 2008; Baumert & Köller, 1998; Häußler & Hoffmann, 1995; Swarat et al., 2012; s. Kapitel 2.5), kann nach Höffler et al. (2014) ein praxisnahes Lehr-Lern-Konzept bereits dann als erfolgreich bezeichnet werden, wenn der Abwärtstrend der Interessen gestoppt wird und sich das Interessenlevel stabilisieren ließe. Und tatsächlich konnten Höffler et al. (2014) für die befragten Schülerinnen und Schüler in nicht-gymnasialen Schulen in Schleswig-Holstein feststellen, dass sowohl das naturwissenschaftliche Fach- als auch Sachinteresse zu Beginn der fünften Jahrgangsstufe hoch ausgeprägt war und beide Interessenfacetten im Zeitverlauf auf einem vergleichsweise hohen Niveau verblieben, auch wenn sich im Zeitverlauf signifikante Veränderungen im Fach- und Sachinteresse zeigten. Das Fachinteresse war dabei nominell höher ausgeprägt als das beobachtete Sachinteresse in Physik, Biologie und Chemie. Darüber hinaus beobachteten die Autoren, dass praxisnah und fächerübergreifend gestalteter MINT-Unterricht das Fähigkeitsselbstkonzept der Schülerinnen und Schüler während der Sekundarstufe I langfristig signifikant stärkte (Höffler et al., 2014). Dieser Befund lässt sich durch ein weiteres Merkmal von praxisbezogenen Lehr-Lern-Formaten erklären: Neben Alltagsbezug und gesteigerter Aktivität im Unterricht wird als drittes Charakteristikum praxisbezogener Lehr-Lern-Formate hervorgehoben, dass die Schülerinnen und Schüler dabei „Fehler als Lerngelegenheit [...] nutzen“ (Seidel et al., 2016, S. 26), wodurch „das Vertrauen in die eigenen Fähigkeiten gefördert“ werden kann (Seidel et al., 2016, S. 11).

Anhand von Regressionsanalysen konnten Höffler et al. (2014) darüber hinaus feststellen, dass das naturwissenschaftliche Fachinteresse eng mit dem Selbstkonzept verknüpft war, wobei das naturwissenschaftliche Fachinteresse zu Beginn des fünften Jahrgangs einen signifikanten Effekt auf das spätere Selbstkonzept der Schülerinnen und Schüler aufwies. Einschränkend muss an dieser Stelle angemerkt werden, dass in der Studie von Höffler et al. (2014) ursprünglich vier Messzeitpunkte zu nur zwei Messzeitpunkten zusammengefasst wurden, wobei der erste Messzeitpunkt einheitlich im fünften Jahrgang stattfand, die folgenden Messzeitpunkte jedoch in ihrer Durchführung variierten (Ende 5. Jahrgangsstufe bis Beginn

8. Jahrgangsstufe). Die berichteten Ergebnisse sollten daher lediglich als Tendenzen interpretiert werden und geben keinen differenzierten Einblick in die Entwicklung der Interessenfacetten sowie des Selbstkonzepts der Schülerinnen und Schüler zwischen dem fünften und sechsten Jahrgang.

Außerschulische Lernorte im MINT-Unterricht

In MINT-Klassen sind es hauptsächlich Schülerlabore, Schülerforschungszentren, Science Center oder Museen, die als außerschulische Lernorte genutzt werden können (vgl. Potvin & Hasni, 2014). Durch den Besuch von außerschulischen Lernorten erhalten Schülerinnen und Schüler authentische Einblicke in naturwissenschaftliche Forschung, können Forschungsmethoden kennenlernen und darüber hinaus mit naturwissenschaftlichen Berufen in Kontakt treten (vgl. Braund & Reiss, 2006; Euler, Schüttler & Hausmann, 2015; Seidel et al., 2016). So bieten beispielsweise Schülerlabore die Möglichkeit MINT-Themen durch selbstständiges Ausprobieren und Experimentieren aktiv und praxisnah im Klassenverband zu vermitteln. Dazu wird in Schülerlaboren eine spezielle Ausstattung bereitgestellt und eine individuelle – meist in Kleingruppen gestaltete – Betreuung der Schülerinnen und Schüler durch Expertinnen und Experten von Universitäten oder externen Forschungsinstituten angeboten (Schwarzer & Parchmann, 2018). Um den Besuch im Schülerlabor langfristig nachhaltig zu gestalten, fügt sich der Schülerlaborbesuch idealerweise thematisch-passend in die Unterrichtskonzeption des MINT-Unterrichts ein und wird durch eine angemessene Vor- und Nachbereitung im schulischen Kontext begleitet (ebd.; Reimann, Herzog, Parchmann & Schwarzer, 2020; vgl. Streller, 2015).

Zentrales Ziel von außerschulischen naturwissenschaftlichen Lernorten allgemein und Schülerlaboren im Speziellen ist die Initiierung und die Förderung von naturwissenschaftlichem Interesse (Euler et al., 2015; Haupt & Hempelmann, 2013; Seidel et al., 2016). Verschiedene Evaluations- und Begleitforschungen haben bereits die Wirksamkeit von Schülerlaborbesuchen auf die Interessenentwicklung und -steigerung untersucht (vgl. u.a. Engeln & Euler, 2004; Mokhonko, 2016; Mokhonko, Nickolaus & Windaus, 2014; Pawek, 2009; Potvin & Hasni, 2014; Röllke, 2019; Simon, 2019; Streller, 2015). Die Studien unterscheiden sich zwar hinsichtlich der untersuchten Altersstufen sowie der angesprochenen Interessenfacetten (Fach- bzw. Sachinteresse, emotionale, wertbezogene, epistemische/kognitive Interessenskomponente¹⁵) und deren Operationalisierung, dennoch können zentrale Tendenzen abgeleitet werden (Euler et al., 2015). Insgesamt zeigt sich in den Studien ein durchaus positives Bild: So stellte Simon (2019) beispielsweise für Schülerinnen und Schüler der Sekundarstufe II fest, dass unabhängig von der Schulart (Gymnasium oder Oberschule) bei der Mehrheit der Schülerinnen und Schüler situationales (d.h.

¹⁵ Nach dem Münchener Konzept der Person-Gegenstand-Relation beziehen sich Interessen auf bestimmte Interessengegenstände (z.B. Unterrichtsthemen oder -fächer) und sind durch eine emotionale, wertbezogene und kognitive Komponente gekennzeichnet (vgl. Kapitel 2.2; zur Übersicht s. Krapp, 2002 sowie Krieg et al., 2021).

ein eher kurzfristig aktuelles) Interesse in Physik durch den Schülerlaborbesuch ausgelöst werden konnte. Darüber hinaus waren die Effekte auf das Fachinteresse etwas höher als die Effekte auf das Sachinteresse ausgeprägt und Mädchen sowie bisher gering Interessierte profitierten stärker vom Schülerlaborbesuch (vgl. dazu auch Pawek, 2009 und Streller, 2015 sowie Brandt, Möller & Kohse-Höinghaus, 2008 für die Domäne Chemie). Die Frage nach der Nachhaltigkeit der interessenfördernden Effekte durch Schülerlabore wird in den Studien mittels Prä-/Post-Designs nachgegangen, die eine erneute Befragung der Schülerinnen und Schüler in zeitlichem Abstand von mehreren Wochen oder Monaten vorsehen. Relativ einheitlich konnte in verschiedenen Studien beobachtet werden, dass das durch den Schülerlaborbesuch hervorgerufene Interesse eher von kurzfristiger Wirkung war und durch einen erneuten Besuch an einem praxisorientierten außerschulischen Lernort erneuert werden musste (Brandt et al., 2008; Simon, 2019; Streller, 2015). Dies betraf insbesondere die emotionale und epistemische Interessenfacette und weniger die wertbezogene Komponente. Das bedeutet, dass die Schülerinnen und Schüler mit zeitlichem Abstand weniger die positive Wahrnehmung (Spaß) und das Bestreben, mehr über das Thema zu erfahren, in den Vordergrund stellten, den Interessengegenstand aber weiterhin als für sie subjektiv bedeutsam werteten (Pawek, 2009). Längerfristig wirksame Effekte lassen sich laut Euler et al. (2015) durch „ein besonders intensives Lernerlebnis ermöglichen“, während „bei eher schulnah ausgerichteten Laboren [...] diese Effekte weniger deutlich [sind]“ (Euler et al., 2015, S. 767).

Das zweite zentrale Ziel von Schülerlaboren ist die Förderung des Fähigkeitsselbstkonzepts der Schülerinnen und Schüler, in dem „Lernmöglichkeiten und Möglichkeiten zur Erfahrung und Entfaltung individueller Stärken“ geschaffen werden (Euler et al., 2015, S. 763). Verschiedene Autorinnen und Autoren untersuchten bereits die Zusammenhänge zwischen dem Besuch eines Schülerlabors und des Fähigkeitsselbstkonzepts von Schülerinnen und Schülern (Brandt et al., 2008; Mokhonko, 2016; Mokhonko et al., 2014; Pawek, 2009; Simon, 2019; Streller, 2015). Insbesondere das aktive Ausprobieren und Experimentieren zeigte dabei positive Effekte auf das Fähigkeitsselbstkonzept, wobei erneut Mädchen besonders davon profitierten (Simon, 2019). Hinsichtlich möglicher Unterschiede in Abhängigkeit des Leistungsniveaus waren hingegen keine Effekte zu beobachten, woraus Simon (2019) schließt, dass Schülerlabore das Fähigkeitsselbstkonzept sowohl von leistungsschwachen als auch leistungsstarken Schülerinnen und Schüler positiv beeinflussen kann. Auch Brandt et al. (2008), Mokhonko et al. (2014) und Pawek (2009) fanden positive Effekte des Schülerlaborbesuchs auf das Fähigkeitsselbstkonzept in Chemie beziehungsweise Physik. Jedoch merken die Autorinnen und Autoren an, dass die Effekte mit einem zeitlichen Abstand von drei (Mokhonko et al., 2014) beziehungsweise vier Monaten (Brandt et al., 2008) verschwunden waren und durch sogenannte „Booster oder Auffrischungsmaßnahmen“, also erneuten außerschulischen Praxisphasen, reaktiviert werden müssten (Brandt et al., 2008, S. 10).

In MINT-Klassen könnte genau dies umgesetzt werden, da die Unterrichtskonzepte allgemein durch einen hohen Praxisbezug und eigenständiges Ausprobieren/Experimentieren gekennzeichnet sind und längerfristige Praxisphasen ermöglichen (s. Kapitel 5.5). Jedoch merken Schwarzer und Parchmann (2018) an, es fehle diesbezüglich aktuell an längsschnittlich angelegten Studien. Insbesondere die Veränderbarkeit des Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts durch außerschulische Lernangebote wurde dabei bisher vernachlässigt beziehungsweise konnte durch die angewendeten Prä-Post-Designs, mit einer erneuten Befragung der Schülerinnen und Schüler im Abstand von nur wenigen Wochen, nicht geklärt werden.

Teilnahme an naturwissenschaftlichen Wettbewerben

Die Teilnahme an MINT-Schülerwettbewerben (z.B. Jugend forscht, Robotik-Wettbewerbe, Biologie-, Physik- oder Chemieolympiade) stellt ein zentrales Merkmal von MINT-Profilklassen dar (s. Kapitel 5.5). Naturwissenschaftliche Schülerwettbewerbe können zur Förderung besonders leistungsstarker Schülerinnen und Schüler beitragen, werden aber vor allem mit dem Potential verbunden, „durch ihren inhärent hohen Motivationscharakter das Interesse einer breiteren Schülerschaft [zu] wecken“ (Blankenburg, Höffler & Parchmann, 2015, S. 76) und so möglicherweise auch für einen MINT-Beruf zu begeistern (Steegeh, Höffler, Höft & Parchmann, 2021a). Diesbezüglich untersuchten Köhler, Bernholt, Höffler und Parchmann (2015), inwiefern sich Schülerinnen und Schüler der Sekundarstufe, die an der JuniorScienceOlympiade teilnahmen, von nicht-teilnehmenden Schülerinnen und Schülern hinsichtlich ihres Interesses für schulische naturwissenschaftliche Aktivitäten (Sachinteresse) sowie bezüglich des domänenspezifischen Selbstkonzepts unterscheiden und konnten herausstellen, dass die Wettbewerbsteilnehmenden sowohl über ein signifikant höheres naturwissenschaftliches Interesse als auch fachspezifisches Selbstkonzept verfügten, als ihre nicht-teilnehmenden Mitschülerinnen und Mitschüler (Köhler et al., 2015). Entsprechende Ergebnisse finden sich auch international. So berichten beispielsweise Welch und Huffman (2011) für amerikanische Schülerinnen und Schüler an der High-School, die an außerschulischen Robotik-Wettbewerben teilnahmen, gegenüber nicht-teilnehmenden Schülerinnen und Schülern eine positivere Einstellung zu Naturwissenschaften sowie zu wissenschaftlicher Forschung. Hinsichtlich der Freude am Naturwissenschaftsunterricht unterschieden sich die beiden Gruppen jedoch nicht.

Da sich die an MINT-Schülerwettbewerben teilnehmenden Schülerinnen und Schüler in zahlreichen Wettbewerbsrunden gegen andere Wettbewerbsteilnehmende durchsetzen müssen, herrscht während der Wettbewerbe jedoch auch Leistungsdruck. Zudem ist die Wettbewerbsteilnahme mit einem erheblichen Zeitaufwand (z.T. über ein Jahr) verbunden, woraus Nick, Urhahne, Stang und Parchmann (2015) schlussfolgern, dass es neben fachspezifischen Kompetenzen insbesondere

auch individueller Wertüberzeugungen (fachlichem Interesse) und Erfolgserwartungen (Fähigkeitsselbstkonzept) bedarf, um in Schülerwettbewerben erfolgreich zu sein. Für die Teilnahme an einer Chemieolympiade konnten sie zeigen, dass das Fachinteresse sowie die Erfolgserwartung entscheidende Indikatoren für die Wettbewerbsleistung darstellten.

Auch andere Studien greifen auf die beiden Komponenten der Erwartungswert Theorie (vgl. Eccles & Wigfield, 2002; s. Kapitel 4) zurück, um die Wettbewerbsteilnahme und den -erfolg erklären zu können (vgl. u.a. Stang, Urhahne, Nick & Parchmann, 2014; Steegh et al., 2021a; Steegh et al., 2021b). Steegh et al. (2021b) heben hervor, dass Mädchen allgemein seltener an mathematischen und naturwissenschaftlichen Schülerwettbewerben teilnehmen als Jungen. Sowohl für Jungen als auch für Mädchen konnten Steegh et al. (2021b) einen signifikant positiven Zusammenhang zwischen dem naturwissenschaftlichen Selbstkonzept und der Wettbewerbsleistung beobachten. Die Bereitschaft an einer weiteren Chemieolympiade teilzunehmen, hing hingegen sowohl für Jungen als auch bei Mädchen mit dem Fachinteresse zusammen. Während bei den Jungen vor allem das fachspezifische Interesse für das Erreichen der nächsten Wettbewerbsrunde eine Rolle spielte, war es bei den Mädchen das domänenspezifische Selbstkonzept. Zudem fanden Steegh et al. (2021b) einen indirekten negativen Zusammenhang zwischen dem naturwissenschaftlichen Geschlechtsrollen-Selbstkonzept der Mädchen und der Bereitschaft zur weiteren Wettbewerbsteilnahme, der durch das fachspezifische Interesse mediiert wurde. Allerdings basieren die beschriebenen Befunde auf querschnittlichen Daten, sodass Aussagen über längerfristige Zusammenhänge zwischen Schülerwettbewerben und dem Fachinteresse beziehungsweise Selbstkonzept nicht getroffen werden können.

5.7 Entwicklung des Interesses und des Fähigkeitsselbstkonzepts in Profilklassen

Koivuhovi et al. (2021)¹⁶ untersuchten Klassen mit und ohne Profil in Finnland und analysierten mögliche Unterschiede zwischen den Klassen in der Entwicklung des mathematischen Fähigkeitsselbstkonzepts zwischen der vierten und sechsten Jahrgangsstufe. Elf der insgesamt 47 untersuchten Klassen wiesen ein Profil auf, davon sechs Klassen ein sprachliches Profil, vier Klassen ein Musik-Profil und eine Klasse ein Sport-Profil. Aufgrund der geringen Anzahl an profilierten Klassen differenzieren die Autorinnen jedoch nicht zwischen den einzelnen thematischen Schwerpunkten, sondern fassten alle Profilklassen zusammen und verglichen diese mit nicht-profilieren Klassen. Sowohl in den Profilklassen als auch in den nicht-profilieren

¹⁶ Die Publikation ist nur in finnischer Sprache erschienen. Eine Zusammenfassung der Ergebnisse wurde von der Erstautorin im Rahmen ihrer Dissertation in englischer Sprache veröffentlicht (Koivuhovi, 2021) und für die vorliegende Arbeit verwendet.

Klassen beobachteten Koivuhovi et al. (2021) eine Abnahme des mathematischen Fähigkeitsselbstkonzepts zwischen der vierten und sechsten Klassenstufe. In den Klassen ohne Profil war der Rückgang gravierender als in den Profilklassen. Unter Kontrolle von Hintergrundmerkmalen (Geschlecht, vorherige Mathematikleistung, Bildungsniveau der Mutter) unterschieden sich die beiden Gruppen jedoch nicht in der Entwicklung des mathematischen Fähigkeitsselbstkonzepts. Interessanterweise stellten Koivuhovi et al. (2021) signifikante geschlechterspezifische Unterschiede im mathematischen Fähigkeitsselbstkonzept nur in Klassen ohne Profil fest – hier zugunsten der Jungen. In den Profilklassen hingegen hatte das Geschlecht keinen signifikanten Effekt auf die Entwicklung des mathematischen Fähigkeitsselbstkonzepts. Weiterhin fanden die Autorinnen Hinweise auf den Big-Fish-Little-Pond-Effekt in den Profilklassen (s. Kapitel 3.3), das heißt negative Effekte eines durchschnittlich hohen Leistungsniveaus der Klasse auf das individuelle Fähigkeitsselbstkonzept der Schülerinnen und Schüler. Der Reflected-Glory-Effekt konnte hingegen nicht beobachtet werden (Koivuhovi et al., 2021; vgl. Koivuhovi, 2021).

Auf Basis einer Längsschnittuntersuchung von Klassen mit erweitertem Musikunterricht (Musik-Klassen) des siebten und achten Jahrgangs stellte Busch (2013) signifikante Unterschiede in der Entwicklung der Selbstwirksamkeit¹⁷ von Schülerinnen und Schülern in Musik-Klassen gegenüber jenen in Regelklassen (mit regulärem Musikunterricht nach Stundentafel) fest. Demnach wiesen die Schülerinnen und Schüler in Musik-Klassen zu allen drei Messzeitpunkten signifikant höhere Werte der musikalischen Selbstwirksamkeit auf als ihre Mitschülerinnen und Mitschüler in den Regelklassen. Dies traf vor allem für Musik-Klassen zu, die zwischen vier und sechs Unterrichtsstunden in Musik pro Woche erhalten hatten. Busch (2013) bezieht sich zur Interpretation der Ergebnisse auf den Reflected-Glory-Effekt und schlussfolgert, dass die Teilnahme an der Musik-Klasse und die dort erhaltene positive Fremdeinschätzung die musikalische Selbstwirksamkeit der Schülerinnen und Schüler stärkt. Allerdings lag die musikalische Selbstwirksamkeitserwartung der Schülerinnen und Schüler in den Musik-Klassen bereits zu Beginn des Beobachtungszeitraums (7. Jahrgangsstufe) weit über der musikalischen Selbstwirksamkeitserwartung der Regelklassenschülerinnen und -schüler und der Unterschied zwischen den Vergleichsgruppen vergrößerte sich bis zur Ende der Untersuchung (Ende 8. Jahrgangsstufe) nur unwesentlich. Dieses Ergebnis begründet Busch (2013) mit möglichen Tendenzen zum Big-Fish-Little-Pond-Effekt: Durch die stärkere Konkurrenz in den Musik-Klassen sei demnach eine positive Entwicklung der Selbstwirksamkeit beschränkt. Zudem kämen weitere Faktoren, wie instrumentale Vorkenntnisse und das kulturelle Kapital der Herkunftsfamilien, in der Entwicklung der musikalischen Selbstwirksamkeit zum Tragen.

¹⁷ Das Konzept der *Selbstwirksamkeit* (Bandura, 1997) bezeichnet „die subjektive Wahrscheinlichkeit, neue und/oder schwierige Situationen aufgrund eigener Kompetenz bewältigen zu können“ (Möller & Trautwein, 2020, S. 203). Demnach ist das Konzept der Selbstwirksamkeit dem Konstrukt des Selbstkonzepts ähnlich, bezieht sich aber auf konkrete Aufgaben, Herausforderungen oder Kontexte und ist demnach deutlich situationaler angelegt (Busch, 2013; Möller & Trautwein, 2020).

In einer vertiefenden Analyse verglich Busch (2013) zudem Musik-Klassen, die maximal drei Stunden Musikunterricht erhalten hatten, mit Musik-Klassen, in denen vier bis sechs Unterrichtsstunden Musik stattfanden (Musik-Klassen+). In den Musik-Klassen+ war dabei ein stärkerer Anstieg der musikalischen Selbstwirksamkeit zu beobachten als in den Musik-Klassen, die weniger erweiterten Musikunterricht erhielten. Die Mittelwertvergleiche zu den drei Messzeitpunkten zeigten jedoch keine signifikanten Unterschiede zwischen den Musik-Klassen (mit drei Musikstunden/Woche) und den Musik-Klassen+. Somit können keine Rückschlüsse vom Umfang des erweiterten Musikunterrichts auf die Entwicklung der musikalischen Selbstwirksamkeit gezogen werden. Außerdem sind Schulformeffekte möglich, da die Musik-Klassen mit bis zu drei Unterrichtsstunden vornehmlich an Gymnasien verortet waren, während sich die Musik-Klassen+ vor allem an Gesamtschulen fanden (Busch, 2013). Weiterhin konnte Busch (2013) zeigen, dass die musikalische Selbstwirksamkeit der stärkste Prädiktor für das musikalische Selbstkonzept der Schülerinnen und Schüler war. Somit scheint die kurzfristige, situationsspezifische Wahrnehmung von musikalischer Selbstwirksamkeit sich auch in einem längerfristigen musikalischen Selbstkonzept zu manifestieren. Darüber hinaus erwies sich die Zugehörigkeit zu einer Musik-Klasse ebenfalls als signifikanter Prädiktor für das musikalische Selbstkonzept (Busch, 2013). Unterschiede zwischen musisch-profilierten und nicht-profilierten Klassen hinsichtlich der Entwicklung des Selbstkonzepts waren jedoch nicht Bestandteil der Untersuchung von Busch (2013).

Entsprechende Analysen finden sich bei Eder (2011), der in seiner im Schuljahr 2005/2006 durchgeführten Untersuchung von österreichischen Musik-Profilklassen der siebten und achten Jahrgangsstufe jedoch keine Unterschiede zwischen dem Selbstkonzept der Schülerinnen und Schüler in Musik-Klassen und in Klassen ohne Profil feststellen konnte. Dennoch waren die Werte des Leistungsselbstkonzepts, des sozialen Selbstkonzepts und des allgemeinen Selbstwertgefühls in den Musik-Klassen nominell höher ausgeprägt als in den nicht-profilierten Klassen.

Interessenentwicklung in Profilklassen

Während Entwicklungsprozesse des naturwissenschaftlichen und musikalischen Fähigkeitsselbstkonzepts der Schülerinnen und Schüler in Profilklassen sowie dessen Unterschiede zu nicht-profilierten Klassen bereits in einigen wenigen Studien untersucht wurde, fehlt es hinsichtlich der Frage nach der Entwicklung des domänenspezifischen Fach- und Sachinteresse in Profilklassen fast gänzlich an Forschungsbefunden.

Ganz allgemein werden MINT-Klassen mit dem Ziel verbunden, „die Motivation und das Interesse an den MINT-Fächern zu wecken“ und bieten darüber hinaus die „Gelegenheit zur Vertiefung schon vorhandener Interessen“ (Seidel et al., 2016, S. 27f.). Ein ähnliches Bild wird auch von Musik-Klassen entworfen: „Schülerinnen und Schüler [...] erleben in Musikklassenangeboten besonders viel Freude und seien darum besonders ‚motiviert‘ [Hervorhebung im Original]“ (Göllner,

2017, S. 54). Allerdings ist laut Göllner (2017) umstritten und bisher nicht erforscht, inwiefern die beschriebene Freude in Musik-Klasse tatsächlich auch Interesse an Musik und am Musikunterricht meint. Heß (2005) konstatiert hingegen, „dass das Konzept Klassenmusizieren die Schüler polarisiert und dass sich negative Erfahrungen innerhalb des Projektes in der Folge extrem ungünstig auf die Motivation auswirken“ (Heß, 2005, S. 88).

Die bisher einzige Untersuchung, die die Entwicklung des musikalischen Fach- und Sachinteresses zwischen fünfter und sechster Jahrgangsstufe in Musik-Klassen differenziert analysiert, stammt aus dem Projekt *Profilbildung an Niedersächsischen Gymnasien und Integrierten Gesamtschulen* (ProBiNi; Stubbe, Nonte et al., 2022). Krieg und Stubbe (2018) untersuchten, wie sich die Entwicklung des musikalischen Fach- und Sachinteresses zwischen Schülerinnen und Schülern in Klassen mit und ohne musikalischen Profil (auch unter Kontrolle der Eingangsselektivität) unterscheidet. Sie konnten zeigen, dass die Musik-Klassen sowohl im fünften auch als im sechsten Jahrgang über ein signifikant höheres musikalischen Fach- und Sachinteresse verfügten als die untersuchten Parallelklassen ohne Profil. Um eine mögliche Eingangsselektivität in der Interessenentwicklung zu berücksichtigen, wurde ein *Propensity Score Matching* (Rosenbaum & Rubin, 1983, 1985) genutzt. Die zugrundeliegende Idee der Methode besteht darin, theoriegeleitet diejenigen Variablen auszuwählen, die die Wahrscheinlichkeit zur Treatmentgruppe zu gehören (also in diesem Fall in eine Musik-Klasse zu gehen) beeinflussen. Beim Propensity Score Matching werden aus der Kontrollgruppe (Vergleichsklassen ohne Profil) diejenigen Schülerinnen und Schüler ausgewählt, die hinsichtlich dieser ausgewählten Variablen ähnliche Werte aufweisen wie die Schülerinnen und Schüler in den Musik-Klassen (=Bildung statistischer Zwillinge). Anschließend werden nur noch die gematchten Fälle analysiert (*average treatment effect for the treated*, ATT). Für das Matching wurden unter anderem der sozioökonomische Status, das Bildungsniveau der Eltern (ISCED), das Geschlecht der Schülerinnen und Schüler sowie das musikalische Fach- und Sachinteresse und die Musikkompetenz einbezogen. Für die Analysen wurden fehlende Werte zunächst in R mit dem Package *MICE* (van Buuren & Groothuis-Oudshoorn, 2009) 10-fach imputiert und das Matching anschließend dem R-Package *MatchIt* (Ho, Imai, King & Stuart, 2011) durchgeführt. Nach dem Matching unterschieden sich die beiden Gruppen zum ersten Messzeitpunkt (5. Jahrgangsstufe) weder im musikalischen Fach- noch im Sachinteresse signifikant voneinander. Nach einem Jahr Musik-Klassenunterricht (6. Jahrgangsstufe) ließ sich hingegen auch in der gematchten Stichprobe ein signifikanter Unterschied im musikalischen Fachinteresse zwischen den Musik-Klassen und den Klassen ohne Profil feststellen – zugunsten der Musik-Klassen. Dieser Befund deutet darauf hin, dass das musikalische Fachinteresse durch den Besuch einer Musik-Klasse tatsächlich gestärkt werden konnte – und zwar unabhängig von möglichen Effekten der Eingangsselektivität (Krieg & Stubbe, 2018). Entsprechende Ergebnisse zur Interessenentwicklung in MINT-Klassen liegen nicht vor.

6 Zwischenfazit und Forschungsdesiderate

Wie in Kapitel 2.6 aus der Darstellung der empirischen Forschungsergebnisse zur Interessenentwicklung deutlich wurde, verringert sich nach dem Übergang in die Sekundarstufe I das durchschnittliche Interesse der Schülerinnen und Schüler erheblich (vgl. u.a. Anderhag et al., 2016; Daniels, 2008; Ferdinand, 2014; Fredricks & Eccles, 2002; Frenzel et al., 2010; Gottfried et al., 2001; Häußler & Hoffmann, 1995; Hoffmann & Lehrke, 1986; Jacobs et al., 2002; Köller et al., 2001; Potvin & Hasni, 2014; Watt, 2004; Wigfield et al., 1997). Dabei zeigten sich Unterschiede zwischen der Entwicklung des Fach- und Sachinteresses (vgl. Häußler & Hoffmann, 1995) sowie hinsichtlich der betrachteten Altersstufen, wobei bis zur siebten Jahrgangsstufe fächerübergreifend der Interessenverlust besonders ausgeprägt war, sich jedoch zum Ende der Sekundarstufe I wieder abschwächte (vgl. Daniels, 2008; Frenzel et al., 2010; Gottfried et al., 2001; Häußler & Hoffmann, 1995). Stark vom Interessenrückgang betroffen scheinen die mathematisch-naturwissenschaftlichen Fächer zu sein (vgl. Anderhag et al., 2016; Daniels, 2008; Frenzel et al., 2010; Gottfried et al., 2001; Häußler & Hoffmann, 1995) sowie das Fach Musik (Wigfield et al., 1997). Der Analyse der Interessententwicklung zu Beginn der Sekundarstufe I sowie der Suche nach Möglichkeiten, um den Interessenrückgang in dieser Phase abzumildern oder gar zu stoppen, kommt daher besondere Bedeutung zu und wird in der vorliegenden Arbeit aufgegriffen. Die negative Entwicklung des Interesses war

zudem häufig mit einem ebenfalls negativen Trend des Fähigkeitsselbstkonzepts assoziiert (vgl. u.a. Daniels, 2008; Häußler & Hoffmann, 1995; Wigfield et al., 1997).

Entsprechend wurde das Konstrukt des Fähigkeitsselbstkonzepts in Kapitel 3 zunächst theoretisch eingeführt und anschließend die Entwicklung des Fähigkeitsselbstkonzepts im Laufe der Schulzeit anhand nationaler und internationaler Forschungsbefunde dargelegt. Weitgehend übereinstimmend verwiesen die berichteten Ergebnisse dabei auf eine Verringerung des mittleren fachspezifischen Fähigkeitsselbstkonzepts im Schulverlauf (vgl. u.a. Dickhäuser et al., 2017; Jacobs et al., 2002; Marsh & Ayotte, 2003; Marsh et al., 1984; Scherrer & Preckel, 2019; Spinath & Spinath, 2005; Watt, 2004). Da sich am Übergang zwischen der Primar- und Sekundarstufe die Lernumgebung und damit auch die sozialen Vergleichsmaßstäbe als Faktor der Selbstkonzeptgenese ändern (Arens, 2017; Eccles, Midgley et al., 1993; s. Kapitel 3.3), stellt der Übergang in die weiterführende Schule ein besonders spannendes Forschungsfeld dar. Hinweise auf negative Effekte des Schulübergangs auf die Entwicklung des fachspezifischen Fähigkeitsselbstkonzepts fanden sich sowohl für deutsche Schülerinnen und Schüler als auch in Forschungsbefunden aus den USA (vgl. Arens et al., 2013; Wigfield & Eccles, 1994).

Die beiden zunächst parallel aufgebauten Forschungsstränge zur Interessenentwicklung sowie zur Fähigkeitsselbstkonzeptgenese konnten in Kapitel 4 im Rahmen der Expectancy-Value Theory (EVT; dt. Erwartungs-Wert Theorie; Eccles & Wigfield, 2002) zusammengeführt werden. Die EVT konzeptualisiert die Erfolgserwartung (Fähigkeitsselbstkonzept) und die Wertüberzeugung (Fach- bzw. Sachinteresse) als getrennte Komponenten, deren positiver Zusammenhang bereits vielfach empirisch belegt werden konnte (vgl. u.a. Arens et al., 2019; Denissen et al., 2007; Dinkelmann & Buff, 2016; Lauer mann et al., 2017; Marsh et al., 2005; Spinath & Steinmayr, 2008; Trautwein et al., 2012; Wigfield et al., 1997). In der Darstellung der empirischen Forschungsergebnisse wurde auch deutlich, dass sich die korrelativen Zusammenhänge der Erwartungs- und Wertkomponente im Schulverlauf intensivieren (vgl. u.a. Denissen et al., 2007; Fredricks & Eccles, 2002; Wigfield et al., 1997) und folglich die Entwicklung des Interesses zunehmend enger mit der Entwicklung des Fähigkeitsselbstkonzepts verknüpft ist.

Während sich für die Erwartungskomponente eine Operationalisierung über das Fähigkeitsselbstkonzept etabliert hat (vgl. u.a. Arens et al., 2019; Guo et al., 2017; Trautwein et al., 2012), wird der intrinsische Wert der Wertkomponente sowohl über Skalen zum Fachinteresse (vgl. u.a. Arens et al., 2019; Gaspard, Lauer mann et al., 2019), als auch über Skalen zum Sachinteresse gemessen (vgl. u.a. Fritzsche et al., 2011, 2015; Marsh et al., 2005; Wigfield et al., 1997). Je nach angesprochenem Bereich (Interesse am Unterrichtsfach vs. Interesse am Themenbereich) sind jedoch unterschiedliche Ergebnisse zu erwarten, da das Fach- und Sachinteresse getrennte Konstrukte darstellen und nicht notwendigerweise gleichzeitig auftreten (Häußler & Hoffmann, 1995; Heß, 2017b; vgl. Kapitel 2.4). Da bisherige Forschungsarbeiten im Kontext der EVT jedoch entweder die eine oder die andere Interessenfacette

verwendeten, fehlt es aktuell an einer Gegenüberstellung möglicher Ergebnisunterschiede. Diesem Forschungsdesiderat wird in der vorliegenden Arbeit Rechnung getragen.

In der Vergangenheit ebenfalls unterschiedlich beantwortet wurde die Frage nach der Direktionalität des Zusammenhangs der Erwartungs- und Wertkomponente. In den berichteten empirischen Befunden traten häufiger Effekte zwischen dem früheren Fähigkeitsselbstkonzept und dem späteren Interesse auf als umgekehrt, wobei sich Unterschiede hinsichtlich der betrachteten Unterrichtsfächer ergaben (vgl. Arens et al., 2019; Lauer mann et al., 2017; Marsh et al., 2005; vgl. Kapitel 4.2). In bisherigen längsschnittlichen Zusammenhangsanalysen ausgespart blieben die Domänen Naturwissenschaften sowie Musik, da sich vergangene Studien vornehmlich auf die Fächer Mathematik, Englisch und Deutsch fokussierten (vgl. Arens et al., 2019; Lauer mann et al., 2017; Marsh et al., 2005). Somit wird mit den in dieser Arbeit präsentierten Analysen auch dieses Forschungsdesiderat aufgegriffen.

Der zweite für diese Arbeit zentrale Themenblock beschäftigte sich mit Profilierungsprozessen auf Schul- und Klassenebene (Kapitel 5). Als zentrale organisatorische und didaktische Gestaltungsmerkmale des musikalischen sowie naturwissenschaftlichen Profilverunterrichts konnten zusätzliche Unterrichtsstunden, praxisbezogene Lehr-Lern-Formate, der Einbezug externer Lernorte sowie die Einbettung der Unterrichtsinhalte in öffentlichkeitswirksame Aktivitäten (z.B. Teilnahme an Auftritten oder Wettbewerben) herausgestellt werden. Da hinsichtlich der Interessen- und Selbstkonzeptentwicklung in Profilklassen ein Forschungsdesiderat besteht, wurden mögliche Effekte einzelner profilspezifischer Gestaltungsmerkmale hinsichtlich ihrer postulierten pädagogischen Potentiale untersucht und entsprechende Forschungsbefunde eingehend beschrieben (vgl. Kapitel 5.6).

Bezüglich praxisorientierter Lehr-Lern-Formen im Musikunterricht konnte gezeigt werden, dass Schüler¹⁸ an musisch-profilierten Gymnasien deutlich häufiger musikalische Aktivitäten ausüben als Schülerinnen und Schüler an nicht-profilierten Schulen (Fritzsche et al., 2011). Zudem wiesen Schülerinnen und Schüler, die aktiv musizieren, ein höheres musikalisches Fachinteresse auf als weniger musikalisch aktive Schülerinnen und Schüler (Heß, 2011a). Zwischen der Häufigkeit des aktiven Musizierens und dem musikalischen Fähigkeitsselbstkonzept zeichneten sich ebenfalls positive Effekte ab, die zum Teil über das Interesse am Instrumentalspiel vermittelt waren (Bernecker et al., 2006; Fritzsche et al., 2011; Kranefeld et al., 2015; Nonte & Schwippert, 2012).

Auch im MINT-Bereich haben sich bereits zahlreiche Studien den Potentialen von praxis-, projekt- oder kontextorientierten Unterrichtskonzepten gewidmet: Neben einem gesteigerten Interesse am Forschen und Experimentieren (Gess et al., 2014; Höffler et al., 2014), konnten ebenfalls positive Effekte auf die Einstellungen

¹⁸ In der Studie von Fritzsche et al. (2011) wurde explizit die musische Aktivität von Schülern an Gymnasien mit Musikprofil untersucht.

der Lernenden gegenüber Naturwissenschaften allgemein sowie naturwissenschaftlichen Themen im schulischen Kontext herausgestellt werden (Bennett et al., 2007; Dolenc & Kazanis, 2020; Gibson & Chase, 2002). Zudem zeigte sich, dass das situationale Interesse der Schülerinnen und Schüler während Experimentierphasen höher war als in passiveren Lernphasen (Palmer, 2009) und die Einschätzung der situativen Interessantheit eines Unterrichtsinhaltes vor allem von der Art der Unterrichtsaktivitäten abhing (Swarat et al., 2012). Darüber hinaus beobachteten Höffler et al. (2014), dass praxisnah und fächerübergreifend gestalteter MINT-Unterricht das Fähigkeitsselbstkonzept der Schülerinnen und Schüler während der Sekundarstufe I langfristig stärken kann.

Im MINT-Bereich sind es hauptsächlich Schülerlabore, Schülerforschungszentren, Science Center oder Museen, die als außerschulische Lernorte für praxisnahen Unterricht genutzt werden. Verschiedene Evaluations- und Begleitforschungen haben bereits die Wirksamkeit des Besuchs von außerschulischen Lernorten auf die Interessenentwicklung untersucht (vgl. u.a. Engeln & Euler, 2004; Mokhonko, 2016; Mokhonko et al., 2014; Pawek, 2009; Potvin & Hasni, 2014; Röllke, 2019; Simon, 2019; Streller, 2015) und zeichneten ein durchaus positives Bild. Allerdings waren die (beispielweise durch einen Schülerlaborbesuch hervorgerufenen) positiven Effekte auf das Interesse eher von kurzfristiger Wirkung und mussten und durch einen erneuten Besuch an einem praxisorientierten außerschulischen Lernort erneuert werden (Brandt et al., 2008; Simon, 2019; Streller, 2015). Das aktive Ausprobieren und Experimentieren in außerschulischen Lernorten zeigte zudem positive Effekte auf das naturwissenschaftliche Fähigkeitsselbstkonzept (vgl. Brandt et al., 2008; Mokhonko et al., 2014, Pawek, 2009; Simon, 2019). Jedoch waren auch hier die positiven Effekte mit einem zeitlichen Abstand von wenigen Monaten verschwunden und müssten durch wiederholte Praxisphasen reaktiviert werden (Brandt et al., 2008; Mokhonko et al., 2014).

Bezüglich der Teilnahme an MINT-Schülerwettbewerben (z.B. Jugend forscht, Robotik-Wettbewerbe, Biologie-, Physik- oder Chemieolympiade) konnten Köhler et al. (2015) beobachten, dass die Wettbewerbsteilnehmenden sowohl über ein signifikant höheres naturwissenschaftliches Interesse als auch fachspezifisches Selbstkonzept verfügten, als ihre nicht-teilnehmenden Mitschülerinnen und Mitschüler. Steegh et al. (2021b) stellten heraus, dass die Bereitschaft an einem naturwissenschaftlichen Wettbewerb teilzunehmen, vom Fachinteresse der Schülerinnen und Schüler abhing, während die Wettbewerbsleistung mit dem naturwissenschaftlichen Fähigkeitsselbstkonzept assoziiert war.

Die Teilnahme der Musik-Klassen an Auftritten oder Konzerten und die dabei erlebte positive Rückmeldung von außen (Applaus; Anerkennung) wurde zudem mit positiven Effekten für die Motivation der Schülerinnen und Schüler sowie der Gruppendynamik in der Klasse verknüpft (Jank & Stroh, 2006; Pabst-Krueger, 2020; vgl. Terhag, 2012).

Wie bereits angemerkt, besteht hinsichtlich der Selbstkonzeptentwicklung in Profilklassen ein Forschungsdesiderat, da sich bisher nur wenige Studien auf mögliche Effekte des Profilklassenbesuchs auf die Fähigkeitsselbstkonzeptgenese fokussierten (Busch, 2013; Eder, 2011; Koivuhovi et al., 2021). Eder (2011) konnte für Schülerinnen und Schüler in österreichischen Musik-Klassen ein nominell höheres Selbstkonzept feststellen als für Schülerinnen und Schüler in Klassen ohne Profil. Busch (2013) fand signifikante Unterschiede in der Entwicklung der musikalischen Selbstwirksamkeit von Schülerinnen und Schülern in Musik-Klassen gegenüber jenen in Regelklassen – zugunsten der Musik-Klassen. Darüber hinaus erwies sich die Zugehörigkeit zu einer Musik-Klasse als signifikanter Prädiktor für das musikalische Selbstkonzept (Busch, 2013).

Für die Entwicklung des mathematischen Fähigkeitsselbstkonzepts zwischen der vierten und sechsten Jahrgangsstufe berichteten Koivuhovi et al. (2021) für ein finnisches Sample, dass der Rückgang des Fähigkeitsselbstkonzepts in Klassen ohne Profil gravierender war als in den Profilklassen. Aufgrund der geringen Anzahl an profilierten Klassen konnten Koivuhovi et al. (2021) jedoch nicht zwischen den einzelnen thematischen Schwerpunkten (Sprach-, Musik- oder Sport-Profil) differenzieren, sodass an dieser Stelle offenbleibt, ob die Entwicklung des fachspezifischen Fähigkeitsselbstkonzepts durch den Besuch einer Profilklassse in derselben Domäne positiv beeinflusst werden kann (z.B. positiver Effekt des Besuchs einer MINT-Profilklassse auf das naturwissenschaftliche Fähigkeitsselbstkonzept).

Auch hinsichtlich der Entwicklung des Fach- und Sachinteresses in Profilklassen besteht aktuell ein Forschungsdesiderat. In der bisher einzigen Untersuchung, die die Entwicklung des musikalischen Fach- und Sachinteresses zwischen fünfter und sechster Jahrgangsstufe in Musik-Klassen differenziert analysiert, konnte herausgestellt werden, dass (nach Kontrolle der Eingangsselektivität) Schülerinnen und Schüler in Musik-Klassen nach einem Jahr Profilverunterricht ein signifikant höheres musikalisches Fachinteresse aufweisen als Schülerinnen und Schüler in Klassen ohne Profil. Entsprechende Effekte für das musikalische Sachinteresse waren hingegen nicht zu beobachten (Krieg & Stubbe, 2018).

7 Forschungsfragen und Hypothesen

Auf Grundlage der in den vorherigen Kapiteln beschriebenen theoretischen Annahmen sowie der dargestellten Forschungsbefunde konnten verschiedene Forschungsdesiderate herausgearbeitet werden, die in der vorliegenden Arbeit bearbeitet werden sollen: Ein Mehrwert der vorliegenden Arbeit besteht in der Analyse der Entwicklungsprozesse des domänenspezifischen Interesses sowie des Fähigkeitsselbstkonzepts zu Beginn der Sekundarstufe I. Die Eingangserhebung direkt nach dem Übergang in die Sekundarstufe I ermöglicht es dabei, Effekte des Wechsels auf die weiterführende Schule in den Blick zu nehmen. Darüber hinaus können die angenommenen Potentiale der Profilklassen, die unter anderem in der besonderen Förderung des Interesses und des Fähigkeitsselbstkonzepts der Schülerinnen und Schüler liegen, empirisch geprüft werden. Bezugnehmend auf die Erwartungs-Wert Theorie (vgl. Eccles & Wigfield, 2002) können außerdem Unterschiede in der Operationalisierung der Wertkomponente als Interesse am Unterrichtsfach (Fachinteresse) oder Interesse an musikalischen beziehungsweise naturwissenschaftlichen Themen (Sachinteresse) herausgestellt und entsprechende Ergebnisse gegenübergestellt werden. Schließlich geben die Analysen zur längsschnittlichen Beziehungsstruktur zwischen der Wertkomponente (domänenspezifisches Fach- bzw. Sachinteresse) und der Erwartungskomponente (domänenspezifisches Fähigkeitsselbstkonzept) Aufschluss über die Veränderbarkeit und Zusammenhänge der Komponenten zwischen der fünften und siebten Jahrgangsstufe. Dabei wird zusätzlich ein

besonderer Fokus auf den Effekt des Klassenprofils gelegt und somit neue Erkenntnisse zu möglichen Potentialen von Profilierungsmaßnahmen generiert. Die übergeordnete Kernfrage der vorliegenden Arbeit lässt sich wie folgt formulieren:

Wie entwickelt sich das musikalische und naturwissenschaftliche Fach- beziehungsweise Sachinteresse sowie das domänenspezifische Fähigkeitsselbstkonzept der Schülerinnen und Schüler während der Sekundarstufe I und welche Rolle spielt dabei der Besuch einer Musik- beziehungsweise MINT-Profilklasse?

Um Entwicklungsprozesse zu untersuchen und dabei Gruppenunterschiede zwischen Profilklassen und Klassen ohne Profil herausstellen zu können, werden in einem ersten Schritt methodische Aspekte in den Fokus gerückt und somit die Bedingungen für die Vergleichbarkeit der Konstrukte im Längsschnitt sowie im Multigruppenvergleich geschaffen. Die ersten drei Forschungsfragen adressieren demnach die Konstruktvalidität sowie die Messinvarianz der drei zentralen Konstrukte (Fach- bzw. Sachinteresse und Fähigkeitsselbstkonzept) und lauten:

Forschungsfrage 1: Weisen die Konstrukte des domänenspezifischen (d.h. musikalischen und naturwissenschaftlichen) Fach- beziehungsweise Sachinteresses sowie des domänenspezifischen Fähigkeitsselbstkonzepts eine angemessene Konstruktvalidität auf?

Forschungsfrage 2: Sind die Konstrukte des domänenspezifischen Fach- beziehungsweise Sachinteresses sowie des domänenspezifischen Fähigkeitsselbstkonzepts im Längsschnitt messinvariant, um die Bedingungen der Vergleichbarkeit der Konstrukte zu gewährleisten?

Forschungsfrage 3: Sind die Konstrukte des domänenspezifischen Fach- beziehungsweise Sachinteresses sowie des domänenspezifischen Fähigkeitsselbstkonzepts im Multigruppenvergleich messinvariant, um die Bedingungen der Vergleichbarkeit der Konstrukte zu gewährleisten?

In einem zweiten Schritt werden die durchschnittlichen Ausprägungen des domänenspezifischen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts von Schülerinnen und Schülern in Musik- und MINT-Klassen sowie deren Parallelklassen ohne Profil zwischen der fünften und siebten Jahrgangsstufe in den Blick genommen. Dabei ist die folgende Forschungsfrage leitend:

Forschungsfrage 4: Lassen sich zwischen der fünften und siebten Jahrgangsstufe Unterschiede in den durchschnittlichen Ausprägungen des domänenspezifischen Fach- und Sachinteresses sowie des domänenspezifischen Fähigkeitsselbstkonzepts zwischen Schülerinnen und Schülern in Musik- beziehungsweise MINT-Klassen und deren Parallelklassen ohne Profil beobachten?

Ausgehend von den in Kapitel 5.6 beschriebenen Forschungsbefunden zu möglichen positiven Effekten der profilspezifischen Gestaltungsmerkmale auf das domänenspezifische Interesse sowie Fähigkeitsselbstkonzept der Schülerinnen und Schüler in Profilklassen werden diesbezüglich folgende Hypothesen geprüft:

- *Hypothese 4.1a:* Das musikalische Fachinteresse der Schülerinnen und Schüler in Musik-Klassen verbleibt zu den drei Messzeitpunkten auf einem überdurchschnittlichen Niveau.
- *Hypothese 4.1b:* Das musikalische Fachinteresse der Schülerinnen und Schüler in Musik-Klassen ist zu den drei Messzeitpunkten signifikant höher ausgeprägt als in den jeweiligen Vergleichsgruppen.
- *Hypothese 4.2a:* Das naturwissenschaftliche Fachinteresse der Schülerinnen und Schüler in MINT-Klassen verbleibt zu den drei Messzeitpunkten auf einem überdurchschnittlichen Niveau.
- *Hypothese 4.2b:* Das naturwissenschaftliche Fachinteresse der Schülerinnen und Schüler in MINT-Klassen ist zu den drei Messzeitpunkten signifikant höher ausgeprägt als in den jeweiligen Vergleichsgruppen.
- *Hypothese 4.3a:* Das musikalische Sachinteresse der Schülerinnen und Schüler in Musik-Klassen verbleibt zu den drei Messzeitpunkten auf einem überdurchschnittlichen Niveau.
- *Hypothese 4.3b:* Das musikalische Sachinteresse der Schülerinnen und Schüler in Musik-Klassen ist zu den drei Messzeitpunkten signifikant höher ausgeprägt als in den jeweiligen Vergleichsgruppen.
- *Hypothese 4.4a:* Das naturwissenschaftliche Sachinteresse der Schülerinnen und Schüler in MINT-Klassen verbleibt zu den drei Messzeitpunkten auf einem überdurchschnittlichen Niveau.
- *Hypothese 4.4b:* Das naturwissenschaftliche Sachinteresse der Schülerinnen und Schüler in MINT-Klassen ist zu den drei Messzeitpunkten signifikant höher ausgeprägt als in den jeweiligen Vergleichsgruppen.
- *Hypothese 4.5a:* Das musikalische Fähigkeitsselbstkonzept der Schülerinnen und Schüler in Musik-Klassen verbleibt zu den drei Messzeitpunkten auf einem überdurchschnittlichen Niveau.
- *Hypothese 4.5b:* Das musikalische Fähigkeitsselbstkonzept der Schülerinnen und Schüler in Musik-Klassen ist zu den drei Messzeitpunkten signifikant höher ausgeprägt als in den jeweiligen Vergleichsgruppen.

- *Hypothese 4.6a:* Das naturwissenschaftliche Fähigkeitsselbstkonzept der Schülerinnen und Schüler in MINT-Klassen verbleibt zu den drei Messzeitpunkten auf einem überdurchschnittlichen Niveau.
- *Hypothese 4.6b:* Das naturwissenschaftliche Fähigkeitsselbstkonzept der Schülerinnen und Schüler in MINT-Klassen ist zu den drei Messzeitpunkten signifikant höher ausgeprägt als in den jeweiligen Vergleichsgruppen.

Im dritten Schritt werden die längsschnittlichen Zusammenhänge des domänenspezifischen Fach- beziehungsweise Sachinteresses sowie des Fähigkeitsselbstkonzepts analysiert. Dafür wird die theoretisch angenommene positive Beziehung zwischen dem Fach- beziehungsweise Sachinteresse sowie dem Fähigkeitsselbstkonzept untersucht und hinsichtlich ihrer korrelativen, autoregressiven und kreuzversetzten Effekte geprüft. Die folgende Forschungsfrage ist dabei handlungsleitend:

Forschungsfrage 5: In welcher längsschnittlichen Beziehung stehen das domänenspezifische Fach- beziehungsweise Sachinteresse und das domänenspezifische Fähigkeitsselbstkonzept?

Es werden folgende Hypothesen geprüft:

- *Hypothese 5.1a:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Fähigkeitsselbstkonzept einen positiven Effekt auf das spätere domänenspezifische Fachinteresse.
- *Hypothese 5.1b:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Fähigkeitsselbstkonzept einen positiven Effekt auf das spätere domänenspezifische Sachinteresse.
- *Hypothese 5.2a:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Fachinteresse einen positiven Effekt auf das spätere domänenspezifische Fähigkeitsselbstkonzept.
- *Hypothese 5.2b:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Sachinteresse einen positiven Effekt auf das spätere domänenspezifische Fähigkeitsselbstkonzept.

Abschließend wird explorativ der Effekt des Klassenprofils auf die längsschnittliche Beziehungsstruktur zwischen dem domänenspezifischen Fach- beziehungsweise Sachinteresse sowie dem domänenspezifischen Fähigkeitsselbstkonzept untersucht. Dafür werden die zur vorherigen Forschungsfrage aufgestellten Hypothesen 5.1a, 5.1b, 5.2a und 5.2b erneut getrennt für die Musik- beziehungsweise MINT-Klassen sowie deren direkte Parallelklasse ohne Profil geprüft. Die entsprechende Forschungsfrage lautet:

Forschungsfrage 6: Lassen sich unter Annahme skalarer Messinvarianz im Längsschnitt und im Gruppenvergleich sowie unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten und der autoregressiven Zusammenhänge innerhalb eines Konstrukts hinsichtlich der längsschnittlichen Beziehungsstruktur zwischen dem domänenspezifischen Fach- beziehungsweise Sachinteresse sowie dem Fähigkeitsselbstkonzept Unterschiede zwischen Profilklassen und Klassen ohne Profil feststellen?

Die zu prüfenden Hypothesen lauten:

- *Hypothese 6.1a:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Fähigkeitsselbstkonzept einen positiven Effekt auf das spätere domänenspezifische Fachinteresse.
- *Hypothese 6.1b:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Fähigkeitsselbstkonzept einen positiven Effekt auf das spätere domänenspezifische Sachinteresse.
- *Hypothese 6.2a:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Fachinteresse einen positiven Effekt auf das spätere domänenspezifische Fähigkeitsselbstkonzept.
- *Hypothese 6.2b:* Unter Annahme skalarer Messinvarianz und unter Kontrolle der korrelativen Zusammenhänge zwischen den Konstrukten sowie der autoregressiven Zusammenhänge innerhalb eines Konstrukts hat ein hohes domänenspezifisches Sachinteresse einen positiven Effekt auf das spätere domänenspezifische Fähigkeitsselbstkonzept.

8 Datengrundlage und methodisches Vorgehen

Die Datengrundlage für die in dieser Arbeit durchgeführten Längsschnittanalysen bildet die von der DFG geförderte Studie *Profilbildung an Niedersächsischen Gymnasien und Integrierten Gesamtschulen* (ProBiNi). Im Folgenden wird zunächst das Studiendesign von ProBiNi vorgestellt, bevor anschließend auf die Gesamtstichprobe sowie die für diese Untersuchung genutzte Substichprobe näher eingegangen wird. Im Anschluss folgt die Beschreibung der verwendeten Erhebungsinstrumente (Kapitel 8.1.2) unter besonderer Berücksichtigung des ordinalen Skalenniveaus und den sich daraus ergebenden forschungspraktischen Konsequenzen (Kapitel 8.2). In Kapitel 8.3 wird der Umgang mit fehlenden Werten thematisiert. Im letzten Kapitel 8.4 werden die verwendeten Analysemethoden detailliert vorgestellt, wobei auch hier ein besonderer Fokus auf die adäquate Berücksichtigung des ordinalen Skalenniveaus in der methodischen Umsetzung gelegt wird.

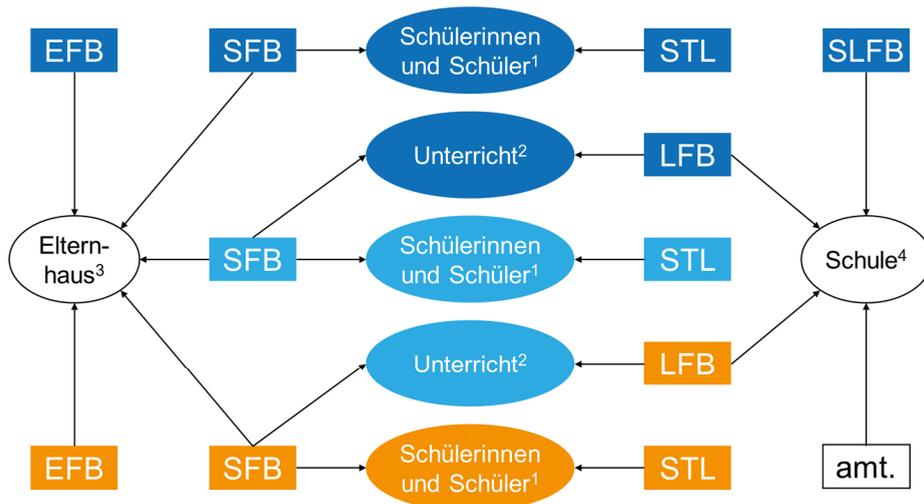
8.1 Die Studie ProBiNi

Die kombinierte Quer- und Längsschnittstudie *Profilbildung an Niedersächsischen Gymnasien und Integrierten Gesamtschulen* (ProBiNi) wurde von der Deutschen Forschungsgemeinschaft (DFG) gefördert¹⁹ und mit einer Projektlaufzeit von fünf Jahren

¹⁹ DFG-Projektnummer: 312968144

(2016 bis 2020) unter der Leitung von Prof. Dr. Tobias C. Stubbe und Prof. Dr. Sonja Nonte an der Georg-August-Universität Göttingen durchgeführt (Stubbe, Nonte et al., 2022). Im Rahmen von ProBiNi wurden Klassenprofilierungsprozesse in der Sekundarstufe I erstmalig umfassend auf Schul-, Klassen- und Schülerenebene untersucht. Auf Basis einer systematischen Analyse von Schulhomepages konnte in einer Vorstudie (Weber et al., 2021) herausgestellt werden, dass Klassenprofilierungen in Hamburg und im Regierungsbezirk Braunschweig (Niedersachsen) am häufigsten im naturwissenschaftlichen und musikalischen Bereich (sogenannte MINT-, Forscher- oder Entdeckerklassen bzw. Bläser-, Orchester- oder Streicherklassen) vorrangig an Gymnasien und integrierten Gesamtschulen vertreten sind. Daher fokussierte sich das Projekt ProBiNi auf diese zwei Profile und nahm sowohl Gymnasien als auch integrierte Gesamtschulen in den Blick. Das längsschnittliche Studiendesign ist in Abbildung 8.01 veranschaulicht (vgl. Stubbe, Krieg, Haas & Nonte, 2022). Zum ersten Messzeitpunkt wurde die Längsschnittstudie zusätzlich durch eine Querschnittsstichprobe ergänzt, um das Schulwahlverfahren zwischen Schulen, die Klassenprofile anbieten, und Schulen, an denen es keine Klassenprofile gibt, näher beleuchten zu können.

Abbildung 8.01: ProBiNi-Studiendesign der Längsschnittstudie



MZP 1: Herbst 2016 (5. Jahrgangsstufe)

MZP 2: Herbst 2017 (6. Jahrgangsstufe)

MZP 3: Herbst 2018 (7. Jahrgangsstufe)

EFB=Elternfragebogen, SFB=Schülerfragebogen,
STL=Schülerteilnahmeliste, LFB=Lehrerfragebogen,
SLFB=Schulleiterfragebogen, amt.=amtliche Statistiken

¹ Kompetenzen (Musik und Biologie) sowie
Hintergrundmerkmale der Schülerinnen und
Schüler zu Beginn des jeweiligen Schuljahres

² Merkmale des Unterrichts im jeweiligen
Schuljahr

³ Merkmale des Elternhauses

⁴ Merkmale der Schule

Der Längsschnittstudie liegt ein quasi-experimentelles Studiendesign zugrunde: Zu drei Messzeitpunkten wurden pro Schule jeweils eine Schulklasse mit musikalischem oder naturwissenschaftlichem Profil (im Folgenden als Musik- bzw. MINT-Klasse bezeichnet) sowie eine Parallelklasse desselben Jahrgangs ohne ein spezifisches Klassenprofil untersucht. Zwischen 2016 und 2018 wurden die Daten der Schülerinnen und Schüler mit Hilfe von standardisierten Paper-Pencil-Fragebögen erhoben. Die Fragebögen enthielten neben Fragen zur Schul- und Profilverwahl sowie zu Einstellungen zum Profilanangebot auch zahlreiche Fragen zu individuellen Merkmalen der Schülerinnen und Schüler, wie dem fachspezifischen Interesse sowie Fähigkeitsselbstkonzept, außerschulischen Aktivitäten oder zu verschiedenen Aspekten der Unterrichtsgestaltung und -qualität. Ergänzend wurden anhand von standardisierten Kompetenztests die Fähigkeiten in den Domänen Biologie und Musik sowie die kognitiven Fähigkeiten der Schülerinnen und Schüler erhoben. Weitere Angaben, wie das Geschlecht, das Alter oder die Halb- beziehungsweise Ganzjahresnoten wurden über die Schülerteilnahmelisten erfasst. Anhand der Schülerteilnahmelisten konnte außerdem pseudonymisiert mittels einer eindeutigen Identifikationsnummer die Zuordnung der Erhebungsmaterialien im Längsschnitt gesichert werden. Die Befragung der Schülerinnen und Schüler wurde im Klassenkontext während der regulären Unterrichtszeit von geschulten externen Testleiterinnen und Testleitern durchgeführt und nahm pro Erhebungszeitpunkt circa 90 Minuten in Anspruch. Zu allen drei Messzeitpunkten fanden die Befragungen der Schülerinnen und Schüler jeweils direkt nach den Sommerferien retrospektiv statt, sodass zum ersten Messzeitpunkt Effekte des Wechsels auf die weiterführende Schule und der Übergang in die Sekundarstufe I untersucht werden konnten.

Zum ersten und dritten Messzeitpunkt wurden außerdem die Eltern der Schülerinnen und Schüler per standardisiertem Paper-Pencil-Fragebogen befragt. Neben der differenzierten Erfassung verschiedener Indikatoren zum sozialen Status der Familien kamen auch Fragen zur Bildungsaspiration oder des Unterstützungsverhaltens, zu allgemeinen Einstellungen zu Musik und Naturwissenschaften sowie zum Wahlverhalten der weiterführenden Schule zum Einsatz (Stubbe, Krieg et al., 2022).

Um die spezifische Gestaltung des Unterrichts in den Profilklassen sowie Effekte der Klassenprofilierung auf Schulentwicklungsprozesse näher untersuchen zu können, wurden außerdem zum ersten und dritten Messzeitpunkt ausgewählte Lehrkräfte per Onlineumfrage befragt. Zum ersten Messzeitpunkt fand die Befragung der Schulleitungen via Onlineumfrage statt und wurde im Rahmen eines qualitativen Teilprojektes durch leitfadengestützte Interviews ergänzt, um die Perspektiven auf Schulprofilierung sowie handlungsleitende Orientierungen der Schulleitungen differenziert analysieren zu können (Keßler, Nonte, Krieg & Brauns, 2022).

8.1.1 Stichprobenbeschreibung

Ursprünglich sollte die Stichprobenziehung auf Basis eines *systematic sampling* umgesetzt werden (Schnell, Hill & Esser, 2018). Grundlage dafür waren Schullisten, die zu Beginn des Projektes differenziert nach Schulformen und Bezirken erstellt wurden. Mittels dieses Verfahrens konnte die Soll-Stichprobe von 24 Schulen jedoch nicht erreicht werden. Um möglichst viele Schulen für die Teilnahme an ProBiNi zu gewinnen, wurden anschließend alle Gymnasien und integrierten Gesamtschulen in Niedersachsen kontaktiert, die ein musikalisches oder naturwissenschaftliches Profil anbieten (*convenience sample*). Von den $N = 88$ kontaktierten Schulen nahmen zum ersten Erhebungszeitpunkt schließlich $n = 16$ Schulen teil (s. Tabelle 8.01). Davon verfügten zwölf Schulen über ein Musik- und/oder MINT-Profil und gehörten somit zum Längsschnittsample, das zu drei Messzeitpunkten befragt werden sollte ($n = 695$ Schülerinnen und Schüler). Vier Schulen ohne ein entsprechendes Profil bildeten das Querschnittssample und waren mit jeweils nur einer Schulklasse lediglich an der ersten Befragung beteiligt ($n = 101$ Schülerinnen und Schüler).

Tabelle 8.01: Beschreibung der Gesamtstichprobe zum ersten Messzeitpunkt

	Gymnasien	Integrierte Gesamtschulen	Gesamt
<i>Querschnitt</i>			
Schulen	1	3	4
Klassen	1	3	4
Schülerinnen und Schüler	27	74	101
<i>Längsschnitt</i>			
Schulen	9	3	12
Klassen	18	7	25
Musik-Klassen	5	2	7
MINT-Klassen	4	2	6
kein Profil	9	3	12
Schülerinnen und Schüler	508	187	695
Musik-Klassen	145	55	200
MINT-Klassen	110	60	170
kein Profil	253	72	325

Im Längsschnittsample hat eine Schule sowohl mit einer musikalischen als auch mit einer naturwissenschaftlichen Profilklassen aber nur mit einer Vergleichsklasse ohne Profil teilgenommen, sodass insgesamt von $n = 25$ Klassen Daten des ersten Messzeitpunkts vorliegen. Die Substichproben der Musik-Klassen ($n = 7$) und MINT-Klassen ($n = 6$) sind in der Gesamtstichprobe annähernd gleich groß. Eine Ungleichverteilung kann hingegen für die Schulform festgestellt werden. So waren

deutlich mehr Gymnasien als integrierte Gesamtschulen an der Längsschnittstudie beteiligt. Darüber hinaus zeigte sich, dass die in der Stichprobe enthaltenen Musik-Klassen eher in urbanen Räumen in der Nähe von Großstädten verortet sind, während sich die MINT-Klassen eher in ländlichen Räumen befinden (vgl. Nonte & Stubbe, 2017).

Von den insgesamt zwölf Schulen des Längsschnittsamples zogen zwei Schulen zum zweiten sowie eine Schule zum dritten Messzeitpunkt ihre Teilnahme an der Studie zurück, sodass neun Schulen in der längsschnittlichen Analytestichprobe verbleiben, die an allen drei Messzeitpunkten teilgenommen haben. Ausgehend von der Gesamtstichprobe des ersten Messzeitpunkts ($n = 796$), wurden für die folgenden längsschnittlichen Analysen die Querschnittsschulen ($n = 101$) sowie Schulen, die ihre Studienteilnahme nach dem ersten oder zweiten Messzeitpunkt beendet haben ($n = 139$), ausgeschlossen.

Für die Analysen der vorliegenden Arbeit können folglich die Daten von $n = 556$ Schülerinnen und Schülern einbezogen werden, die zwischen 2016 und 2018 an der ProBiNi-Studie teilgenommen haben (s. Tabelle 8.02). Von den insgesamt in der Analytestichprobe enthaltenen $n = 556$ Schülerinnen und Schülern besuchten $n = 149$ Schülerinnen und Schüler eine Musik-Profilklasse, $n = 144$ Schülerinnen und Schüler eine MINT-Profilklasse und $n = 263$ Schülerinnen und Schüler eine Klasse ohne Profil (Vergleichsklasse). Pro Schule wurde je eine Profilklasse (Musik- oder MINT-Klasse) und eine Parallelklasse desselben Jahrgangs ohne Profil (Vergleichsklasse) befragt. Von diesem Studiendesign wurde an einer Schule abgewichen, die sowohl mit einer Musik-Klasse als auch einer MINT-Klasse aber nur mit einer Parallelklasse desselben Jahrgangs ohne Profil teilgenommen hat. An dieser Schule wird die Parallelklasse ohne Profil gleichzeitig als Vergleichsklasse für die Musik-Klasse als auch als Vergleichsklasse für die MINT-Klasse herangezogen. Somit sind neben den fünf Musik-Klassen und fünf MINT-Klassen nur neun Klassen ohne Profil in der Analytestichprobe enthalten (s. Tabelle 8.02).

Tabelle 8.02: Beschreibung der Analytestichprobe

	Gymnasien	Integrierte Gesamtschulen	Gesamt
Schulen	7	2	9
Klassen	14	5	19
Musik-Klassen	4	1	5
MINT-Klassen	3	2	5
kein Profil	7	2	9
Schülerinnen und Schüler	407	149	556
Musik-Klassen	120	29	149
MINT-Klassen	84	60	144
kein Profil	203	60	263

Die Schülerinnen und Schüler in der Analysestichprobe sind zum ersten Messzeitpunkt (Beginn 5. Jahrgangsstufe) im Durchschnitt $M = 10.4$ Jahre ($SD = 0.41$) alt. Das Geschlechterverhältnis fällt in den verschiedenen Subgruppen sehr unterschiedlich aus und wird daher für die Profilklassen und die Vergleichsklassen ohne Profil getrennt betrachtet. In Tabelle 8.03 sind die Anteile der Mädchen und Jungen in den Musik- beziehungsweise MINT-Klassen sowie deren direkten Parallelklassen ohne Profil (Vergleichsklassen) dargestellt. Durch die Studienteilnahme einer Schule mit zwei Profilklassen aber nur einer Vergleichsklasse ohne Profil ist diese Vergleichsklasse sowohl in der Substichprobe der Musik-Vergleichsklassen als auch in der Substichprobe der MINT-Vergleichsklassen enthalten.

Während die Musik-Klassen von deutlich mehr Mädchen (71.1 %) als Jungen (28.9 %) besucht werden, überwiegt in den MINT-Klassen der Anteil der Jungen (60.4 %) gegenüber dem Anteil der Mädchen (39.6 %). Dieser Befund überrascht angesichts der von Kessels (2005) veröffentlichten Ergebnissen zu Geschlechterdifferenzen im präferierten Unterrichtsfach nicht. In den MINT-Vergleichsklassen ist das Geschlechterverhältnis ausgeglichen und in den Musik-Vergleichsklassen nur leicht zugunsten der Jungen (55.6 %) verschoben.

Tabelle 8.03: Anteil der Mädchen und Jungen in der Analysestichprobe – absolute und relative Häufigkeiten

	Mädchen <i>n</i>	Mädchen %	Jungen <i>n</i>	Jungen %
Musik-Klassen	106	71.1	43	28.9
Musik-Vergleichsklassen	67	44.4	84	55.6
MINT-Klassen	57	39.6	87	60.4
MINT-Vergleichsklassen	71	50.0	71	50.0

Auch hinsichtlich der Eingangsselektivität beim Übergang in die Sekundarstufe I zeigen sich Unterschiede zwischen den musikalischen und naturwissenschaftlichen Profilklassen und den Klassen ohne Profil (Vergleichsklassen). Tabelle 8.04 gibt einen Überblick über die Verteilung des sozialen Status in den verschiedenen Klassen auf Basis des gebildeten Sozialindex (zur Berechnung des Sozialindexes s. Kapitel 8.1.2). Die Ergebnisse deuten auf eine heterogene Stichprobe hin. Die Familien der Schülerinnen und Schüler in den Musik-Klassen weisen im Vergleich zu den übrigen Untersuchungsgruppen in der Stichprobe den höchsten sozialen Status auf (vgl. Lehmann-Wermser et al., 2010), gefolgt von den Schülerinnen und Schülern in den Musik-Vergleichsklassen. MINT-Klassen und deren direkte Vergleichsklassen ohne Profil unterscheiden sich hinsichtlich des sozialen Status der Schülerinnen und Schüler kaum, liegen insgesamt aber deutlich unter den Musik-Klassen und deren Vergleichsklassen. Somit scheinen sich die Unterschiede im sozioökonomischen

Hintergrund der Familien verstärkt auf Schulebene statt auf Klassenebene zu ergeben. Profilklassen und deren Vergleichsklassen an der gleichen Schule unterscheiden hinsichtlich des soziökonomischen Hintergrund der Familien entsprechend weniger als beispielsweise Musik- und MINT-Profilklassen. Da die mit MINT-Klassen teilnehmenden Schulen vornehmlich im ländlichen Raum verortet sind, könnten die Unterschiede im soziökonomischen Hintergrund zwischen den Schulen auch Ausdruck eines systematischen Stadt/Land-Effekts sein (vgl. Nonte & Stubbe, 2017).

Tabelle 8.04: Beschreibung des sozialen Status in der Analysestichprobe

	<i>M</i>	<i>(SE)</i>	<i>SD</i>
Gesamt	296	(2.7)	64
Musik-Klassen	316	(5.3)	64
Musik-Vergleichsklassen	301	(5.2)	63
MINT-Klassen	278	(5.0)	60
MINT-Vergleichsklassen	279	(5.5)	64

Der Personenschätzer (*weighted likelihood estimator*; WLE) ist normiert auf $M = 300$, $SD = 60$ für die 380 Schülerinnen und Schüler, die an allen drei Erhebungen teilgenommen haben.

8.1.2 Erhebungsinstrumente

In diesem Kapitel werden die für die Analysen verwendeten Erhebungsinstrumente vorgestellt. Dabei wird zunächst detailliert auf die Operationalisierung der drei zentralen Konstrukte des Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts eingegangen und anschließend werden die Messinstrumente weiterer Hintergrundmerkmale vorgestellt.

Den theoretischen Überlegungen entsprechend (s. Kapitel 4.1), wurde in ProBiNi die Wert- und Erwartungskomponente (vgl. Eccles & Wigfield, 2002) differenziert und fachspezifisch, das heißt für die Domänen Musik und Naturwissenschaften getrennt, erhoben (vgl. Denissen et al., 2007; Gaspard, Nagengast et al., 2019; Trautwein et al., 2012). Um die unterschiedlichen Facetten der Wertkomponente adäquat abbilden zu können, wurde zum einen der subjektive Wert des jeweiligen Schulfaches (Fachinteresse in Musik bzw. den Naturwissenschaften) und zum anderen die individuelle Wertüberzeugung für musikalische beziehungsweise naturwissenschaftliche Themen oder Aufgaben (Sachinteresse) erfasst. Die Erwartungskomponente wurde über das fachspezifische Fähigkeitsselbstkonzept operationalisiert. Die jeweiligen Fragen der Skalen des Fach- beziehungsweise Sachinteresses sowie des Fähigkeitsselbstkonzepts waren für die zwei Domänen parallel formuliert und unterschieden sich lediglich hinsichtlich des angesprochenen Schulfaches. Als

Antwortformat wurde für die sechs zentralen Konstrukte zu allen drei Messzeitpunkten eine vierstufige Likert-Skala mit den Antwortkategorien 1 = *stimmt genau* bis 4 = *stimmt gar nicht* eingesetzt. Alle Items wurden rekodiert, sodass ein niedriger Wert für eine niedrige Zustimmung und ein hoher Wert für eine starke Zustimmung steht.

Das *Fachinteresse* ist als subjektive Bedeutung, die die Schülerinnen und Schüler dem Unterrichtsfach entgegenbringen, konzeptualisiert. Das Fachinteresse wurde mittels einer Skala aus vier Items erfasst (Daniels, 2008), die im Rahmen der BIJU-Studie (vgl. Baumert et al., 2003) entwickelt wurde und ihre Ursprünge bereits in einer 1969 vom Max-Planck-Institut für Bildungsforschung Berlin durchgeführten Studie hat (vgl. Baumert, Roeder, Sang & Schmitz, 1986). Tabelle 8.05 gibt einen Überblick über die eingesetzten Items der Skala des musikalischen sowie naturwissenschaftlichen Fachinteresses. Zudem werden Mittelwerte, Standardabweichungen und Trennschärfen der Items zum ersten Messzeitpunkt in der Gesamtstichprobe berichtet.

Tabelle 8.05: Kennwerte der Items zum musikalischen und naturwissenschaftlichen Fachinteresse zum ersten Messzeitpunkt

<i>Fachinteresse – Musik</i>	<i>M</i>	<i>SD</i>	<i>r_{it}</i>
1. Mir liegt viel daran, im Fach Musik viel zu wissen.	2.94	1.01	.81
2. Ich würde gern im Fach Musik noch mehr Stunden haben als bisher.	2.61	1.21	.82
3. Ich freue mich sehr auf eine Unterrichtsstunde in Musik.	2.95	1.10	.85
4. Mir liegt viel daran, den Stoff im Fach Musik zu behalten.	2.99	1.03	.79
<i>Fachinteresse – Naturwissenschaften</i>	<i>M</i>	<i>SD</i>	<i>r_{it}</i>
1. Mir liegt viel daran, in den naturwissenschaftlichen Fächern viel zu wissen.	2.90	0.87	.72
2. Ich würde in den naturwissenschaftlichen Fächern gern noch mehr Stunden haben als bisher.	2.46	1.05	.74
3. Ich freue mich auf eine Stunde in den naturwissenschaftlichen Fächern.	2.95	0.97	.78
4. Mir liegt viel daran, den Stoff in den naturwissenschaftlichen Fächern zu behalten.	2.96	0.92	.72

Antwortkategorien: 1 = stimmt gar nicht, 2 = stimmt eher nicht, 3 = stimmt eher, 4 = stimmt genau

Die Skala zum Fachinteresse enthält neben zwei Items zum intrinsischen Wert (Item 2 und 3), der auch als Freude am Unterrichtsfach bezeichnet werden kann, ebenfalls zwei Items zur persönlichen Wichtigkeit/Bedeutsamkeit des jeweiligen Unterrichtsfachs (Item 1 und 4). Somit können sowohl die gefühlsbetonte als auch die wertbezogene Komponente des Interesses abgebildet werden (vgl. Schiefele, 1996). Die Formulierung der vier Items wurde über die drei Messzeitpunkte nicht verändert und lediglich hinsichtlich des angesprochenen Schulfachs für die Domänen Musik und Naturwissenschaften variiert.

Das *Sachinteresse* bezieht sich – unabhängig vom schulischen Kontext – auf Themen im musikalischen (Heß, 2011b) beziehungsweise naturwissenschaftlichen (Daniels, 2008) Bereich. Das Sachinteresse wurde anhand einer Skala aus vier Items gemessen, die neben der Wichtigkeit (Item 1) auch den intrinsischen Wert (Item 2), das heißt neben der wertbezogenen auch die affektive Komponente, abbilden. Darüber hinaus enthält die Skala ebenfalls den Aspekt des Erlebens von *Flow* (vgl. Csikszentmihalyi, 2010; Item 3) sowie – im Sinne der Selbstbestimmungstheorie (vgl. Deci & Ryan, 1993) – den Aspekt der Selbstintentionalität (Item 4).

Tabelle 8.06: Kennwerte der Items zum musikalischen und naturwissenschaftlichen Sachinteresse zum ersten Messzeitpunkt

<i>Sachinteresse – Musik</i>	<i>M</i>	<i>SD</i>	<i>r_{it}</i>
1. Die Beschäftigung mit musikalischen Themen ist für mich sehr wichtig, unabhängig von der Schule oder anderen Personen.	2.64	1.00	.70
2. Auf die Beschäftigung mit Musik würde ich nur ungern verzichten, einfach weil sie mir Freude macht.	3.06	0.97	.72
3. Wenn ich mich mit musikalischen Dingen befasse, kann ich darin richtig versunken sein.	2.71	1.02	.65
4. Für die Beschäftigung mit musikalischen Dingen bin ich auch bereit, meine Freizeit zu verwenden.	2.60	1.11	.69
<i>Sachinteresse – Naturwissenschaften</i>	<i>M</i>	<i>SD</i>	<i>r_{it}</i>
1. Die Beschäftigung mit naturwissenschaftlichen Themen ist für mich sehr wichtig, unabhängig von der Schule oder anderen Personen.	2.51	0.95	.72
2. Auf die Beschäftigung mit naturwissenschaftlichen Themen würde ich nur ungern verzichten, einfach weil sie mir Freude macht.	2.56	1.00	.76
3. Wenn ich mich mit naturwissenschaftlichen Dingen befasse, kann ich darin richtig versunken sein.	2.41	1.03	.71
4. Für die Beschäftigung mit naturwissenschaftlichen Dingen bin ich auch bereit, meine Freizeit zu verwenden.	2.17	1.02	.71

Antwortkategorien: 1 = stimmt gar nicht, 2 = stimmt eher nicht, 3 = stimmt eher, 4 = stimmt genau

Die Itemformulierung kann der Tabelle 8.06 entnommen werden, die darüber hinaus auch die Mittelwerte, Standardabweichungen sowie Trennschärfen der Items zum ersten Messzeitpunkt in der Gesamtstichprobe zeigt. Die Formulierung der vier Items wurde über die drei Messzeitpunkte hinweg beibehalten und lediglich hinsichtlich des angesprochenen Schulfachs für die Domänen Musik und Naturwissenschaften variiert.

Das fachspezifische *Fähigkeitsselbstkonzept* wurde anhand der etablierten Skala DISK-Gitter (Rost & Sparfeldt, 2002) erhoben. Die Skala enthält jeweils acht Items, die für die beiden Domänen Musik und Naturwissenschaften analog formuliert waren. In den Tabellen 8.07 und 8.08 sind die Einzelitems zum fachspezifischen Fähigkeitsselbstkonzept sowie die Mittelwerte, Standardabweichungen und Trennschärfen der Skala zum ersten Messzeitpunkt für die Gesamtstichprobe aufgeführt. Neben dem Aspekt der Selbsteinschätzung eigener Fähigkeiten ist in der Skala auch eine Leistungskomponente enthalten, die sich auf die Schulnoten bezieht. Da das Fähigkeitsselbstkonzept stets vom klasseninternen Bezugsrahmen/Kontext abhängt (s. Kapitel 3.3), sollte das individuelle Fähigkeitsselbstkonzept der Schülerinnen und Schüler auch im Vergleich zu ihren Mitschülerinnen und Mitschülern eingeschätzt werden.

Tabelle 8.07: Kennwerte der Items zum musikalischen Fähigkeitsselbstkonzept zum ersten Messzeitpunkt

<i>Fähigkeitsselbstkonzept – Musik</i>	<i>M</i>	<i>SD</i>	<i>r_{it}</i>
1. Ich weiß in Musik die Antwort auf eine Frage schneller als die Anderen.	2.48	0.86	.70
2. Manchmal fühle ich mich in Musik anderen überlegen und glaube, dass sie noch manches von mir lernen können.	2.22	0.94	.58
3. Ich kann in Musik Sachen selbst herausbekommen.	3.11	0.89	.74
4. Ich bin in Musik zufrieden mit meiner Fähigkeit vor der Klasse zu sprechen.	3.03	0.95	.67
5. Es fällt mir in Musik leicht, Probleme zu lösen.	2.99	0.90	.76
6. In Musik fallen mir gute Zensuren zu.	3.12	0.80	.71
7. Ich gehöre in Musik zu den Guten.	2.96	0.88	.79
8. Ich habe ein gutes Gefühl, was meine Arbeit in Musik angeht.	3.17	0.86	.74

Antwortkategorien: 1 = stimmt gar nicht, 2 = stimmt eher nicht, 3 = stimmt eher, 4 = stimmt genau

Tabelle 8.08: Kennwerte der Items zum naturwissenschaftlichen Fähigkeitsselbstkonzept zum ersten Messzeitpunkt

<i>Fähigkeitsselbstkonzept – Naturwissenschaften</i>	<i>M</i>	<i>SD</i>	<i>r_{it}</i>
1. Ich weiß im naturwissenschaftlichen Unterricht die Antwort auf eine Frage schneller als die Anderen.	2.37	0.75	.66
2. Manchmal fühle ich mich im naturwissenschaftlichen Unterricht anderen überlegen und glaube, dass sie noch manches von mir lernen können.	2.14	0.85	.53
3. Ich kann im naturwissenschaftlichen Unterricht Sachen selbst herausbekommen.	3.02	0.84	.62
4. Ich bin im naturwissenschaftlichen Unterricht zufrieden mit meiner Fähigkeit vor der Klasse zu sprechen.	2.91	0.94	.65
5. Es fällt mir im naturwissenschaftlichen Unterricht leicht, Probleme zu lösen.	2.78	0.80	.69
6. Im naturwissenschaftlichen Unterricht fallen mir gute Zensuren zu.	2.98	0.78	.68
7. Ich gehöre im naturwissenschaftlichen Unterricht zu den guten Schülerinnen und Schülern.	2.88	0.82	.72
8. Ich habe ein gutes Gefühl, was meine Arbeit im naturwissenschaftlichen Unterricht angeht.	3.05	0.88	.70

Antwortkategorien: 1 = stimmt gar nicht, 2 = stimmt eher nicht, 3 = stimmt eher, 4 = stimmt genau

Dass die drei zentralen Konstrukte zu allen drei Messzeitpunkten sowie in den beiden Domänen Musik und Naturwissenschaften eine gute interne Konsistenz aufweisen, zeigt die folgende Tabelle 8.09. Sowohl für die Skala zum musikalischen als auch zum naturwissenschaftlichen Fachinteresse können für die drei Messzeitpunkte *Cronbachs* α -Werte $\geq .88$ bestätigt werden. Die interne Konsistenz für die Skalen des musikalischen und naturwissenschaftlichen Sachinteresses liegt im Längsschnitt zwischen *Cronbachs* $\alpha = .85$ und *Cronbachs* $\alpha = .93$ und ist damit ebenfalls als sehr gut zu bezeichnen. Auch für das fachspezifische Fähigkeitsselbstkonzept lässt sich eine gute interne Konsistenz im Längsschnitt feststellen (*Cronbachs* $\alpha \geq .88$).

Tabelle 8.09: Interne Konsistenz (*Cronbachs a*) der zentralen Konstrukte zu den drei Messzeitpunkten (MZP)

	Itemanzahl	α 1. MZP	α 2. MZP	α 3. MZP
<i>Fachinteresse</i>				
Musik	4	.92	.93	.92
Naturwissenschaften	4	.88	.92	.91
<i>Sachinteresse</i>				
Musik	4	.85	.90	.93
Naturwissenschaften	4	.87	.89	.89
<i>Fähigkeitsselbstkonzept</i>				
Musik	8	.91	.92	.93
Naturwissenschaften	8	.88	.91	.90

Weitere Hintergrundmerkmale

Der *soziale Status* der Familien wurde mittels eines Sozialindex erfasst (vgl. Pietsch, Bensen & Bos, 2006), der auf 49 dichotomisierten Items aus dem Fragebogen der Schülerinnen und Schüler sowie dem Elternfragebogen basiert. Von den Schülerinnen und Schülern wurden beispielsweise Angaben zur Anzahl der Bücher im Haushalt, zu weiteren Besitztümern (z.B. ein Klavier) oder zu kulturellen Aktivitäten mit den Eltern (z.B. Theater- oder Konzertbesuche) einbezogen. Die Elternangaben bezogen sich unter anderem auf das Bildungsniveau, das soziale Kapital (Ressourcengenerator; van der Gaag & Snijders, 2005), die Armutsgefährdung (vgl. Bardone & Guio, 2005) sowie Einkommensreichtum. Der Sozialindex adressiert somit – in Anlehnung an Bourdieu (1983) – das ökonomische, kulturelle und soziale Kapital der Familien. Der Sozialindex wurde auf Basis der *Item Response Theory* (IRT) unter Verwendung eines einparametrischen logistischen Modells (1PL) gebildet. Die 49 dichotomisierten Variablen wurden mit dem Package *TAM* (Robitzsch, Kiefer & Wu, 2017) in R (R Development Core Team, 2017) nach dem Raschmodell (Rasch, 1960) skaliert und zu einem Wert (Sozialindex) zusammengefasst. Der *weighted likelihood estimator* (WLE; Warm, 1989) ist auf $M = 300$ und $SD = 60$ normiert. Die WLE-Reliabilität des Sozialindex liegt bei .64 und ist demnach gerade noch akzeptabel.

8.2 Ordinales Skalenniveau – forschungsmethodische Aspekte

Die im vorherigen Kapitel 8.1.2 beschriebenen Skalen, die für die Analysen der vorliegenden Arbeit verwendet werden, weisen (mit Ausnahme des gebildeten Sozialindex) aufgrund ihres vierstufigen Antwortformates ein ordinales Skalenniveau auf. Um das ordinale Skalenniveau in den folgenden Analysen adäquat zu berücksichtigen, werden in diesem Kapitel methodische Überlegungen zum Umgang mit ordinalen Daten aufgegriffen und die sich daraus ergebenden forschungspraktischen Konsequenzen näher beleuchtet.

Insbesondere vor dem Hintergrund der anhaltenden Popularität des Einsatzes von Likert-Skalen zur Messung diverser latenter Konstrukte sind die Berücksichtigung des ordinalen Skalenniveaus sowie entsprechende Modellierungsansätze von besonderer Relevanz. Zahlreiche Publikationen haben sich bereits diesem Thema angenommen (vgl. u.a. Jia & Wu, 2019; Li, 2016a, 2016b; Liu et al., 2017; Rhemtulla, Brosseau-Liard & Savalei, 2012; Robitzsch, 2020; Svetina, Rutkowski & Rutkowski, 2020). Dabei sind zwei Forschungslinien vorherrschend: Einerseits wird in der Forschungspraxis statt der (eigentlich) ordinalen Datenstruktur eine kontinuierliche Datenstruktur angenommen und entsprechende Schätzverfahren verwendet. Der überwiegende Teil der Forschungsergebnisse empfiehlt hingegen, die ordinalen Daten auch als solche zu behandeln und die Analysemethoden entsprechend anzupassen (vgl. u.a. Liu et al., 2017; Svetina et al., 2020). Robitzsch (2020) merkt diesbezüglich an, es sei „[...] maybe improbable to reach a consensus about this issue“ (Robitzsch, 2020, S. 5) und weist darauf hin, dass methodische Empfehlungen dieser Art häufig auf Grundlage von Simulationsstudien ausgesprochen und mit fehlerhaften Schätzungen (z.B. der Faktorladungen in konfirmatorischen Faktorenanalysen) begründet werden. Rhemtulla et al. (2012) heben hervor, das ordinale Skalenniveau insbesondere dann zu berücksichtigen, wenn für die eingesetzten Skalen maximal fünf Antwortkategorien vorliegen. Chen, Wu, Garnier-Villarreal, Kite und Jia (2020) sowie Svetina et al. (2020) betonen die Relevanz im Rahmen von konfirmatorischen Faktorenanalysen sowie in Messinvarianztestungen das ordinale Skalenniveau angemessen zu berücksichtigen und somit andernfalls möglicherweise irrtümlich getroffene Annahmen einer multivariaten Normalverteilung kontinuierlicher Daten zu vermeiden. Im Rahmen von Mehrgruppenanalysen kann die Missachtung des ordinalen Skalenniveaus zudem weitgreifende Folgen für die Modellparameterschätzung, die Beurteilung der Modellgüte sowie die Interpretation von gruppenspezifischen Unterschieden haben (Lubke & Muthén, 2004; Svetina et al., 2020). Lubke und Muthén (2004) konnten diesbezüglich anhand einer Simulationsstudie zeigen, dass aussagekräftige Gruppenvergleiche auf latenter Ebene durch die fehlende Berücksichtigung des ordinalen Skalenniveaus erschwert werden, da beispielsweise Messinvarianzannahmen durch die fehlende Schätzung beziehungsweise Gleichsetzung von Schwellenwerten verletzt sein können (zu den Spezifika

der Messinvarianzanalyse mit ordinalen Daten s. Kapitel 8.4.4). Forschungspraktisch können sich hinter schlechten Model-Fits (s. Kapitel 8.4.2) solcher Modelle demnach sowohl fehlende Messinvarianz, Unterschiede in den geschätzten Schwellenwerten verschiedener Variablen zwischen den Gruppen oder aber die Missachtung des ordinalen Skalenniveaus und damit einhergehenden Verletzungen der Normalverteilungsannahmen verbergen. Die Unterschätzung der Faktorvarianz erschwert zudem die Interpretation von Gruppenunterschieden in Bezug auf die Faktormittelwerte. Im Rahmen längsschnittlicher Multigruppenmodelle ist außerdem keine klare Unterscheidung zwischen Messinvarianz zwischen Gruppen und Veränderungen in den Schwellenwerten über die Zeit möglich, da die Schwellenwerte in kontinuierlichen Modellen nicht berücksichtigt werden (Lubke & Muthén, 2004). Dies kann zu Verzerrungen in der Interpretation von Gruppenunterschieden führen (s. Kapitel 8.4.4).

Die Berücksichtigung des ordinalen Skalenniveaus im Rahmen der Strukturgleichungsmodellierung bringt einige forschungspraktische Konsequenzen mit sich, allen voran die Auswahl des passenden Modellschätzers. Die Auswahl des Schätzers hat wiederum Auswirkungen auf die im Modellierungsprozess getroffenen Annahmen zum Umgang mit fehlenden Werten oder die Berechnung sowie Interpretation von Model-Fit-Kriterien, wie dem χ^2 -Differenzentest (s. Kapitel 8.3 und 8.4.2).

Um dem ordinal-kategorialen Skalenniveau, der Stichprobengröße und fehlenden Normalverteilungsannahmen Rechnung zu tragen, wird für die vorliegenden Analysen der robuste *Mean and Variance adjusted diagonally Weighted Least Squares* (WLSMV)-Schätzer (Li, 2016a; Muthén, Muthén & Asparouhov, 2015) verwendet. Dieser eignet sich insbesondere für nicht-normalverteilte, ordinale Daten und wird standardmäßig in *Mplus 7.11* (Muthén & Muthén, 2012) für die Analyse ordinaler Daten verwendet (Muthén et al., 2015; Svetina et al., 2020; Wang & Wang, 2020).

Der WLSMV gehört zu den *Weighted Least Square* (WLS)-Schätzern (Methode der kleinsten Quadrate). Unter Verwendung eines WLS- beziehungsweise WLSMV-Schätzers umfasst eine Modellschätzung folgende drei Schritte:

First, univariate information of each variable in the sample is used to obtain the maximum likelihood estimates of the sample implied thresholds. Second, polychoric correlations between each pair of the observed indicators are calculated by treating the thresholds obtained in step 1 as fixed [...]. Third, the estimated thresholds and polychoric correlations are used to form the discrepancy function which is minimized to obtain the estimates for the model parameters (Chen et al., 2020, S. 90).

WLS und WLSMV nutzen die gleiche asymptotische Varianz-Kovarianz-Matrix; der WLSMV jedoch nur für die Schätzung der Standardfehler und Modellgütekriterien (Wang & Wang, 2020). Zudem ist der WLSMV gegenüber dem WLS hinsichtlich der Genauigkeit der Schätzung deutlich überlegen, sofern kleine Stichproben und/oder viele Variablen in die Modelle aufgenommen werden (Beauducel &

Herzberg, 2006; Li, 2016a). Unter Verwendung des WLS kann es zudem zu Verzerrungen in der Modellparameterschätzung kommen, da eine entsprechende Korrektur der Teststatistik (Mittelwerte und Varianzen) fehlt (Chen et al., 2020; Li, 2016a). Der Vergleich des WLSMV-Schätzers mit anderen Schätzern, die auf dem *Maximum Likelihood* oder *Bayesian* basieren, wurde bereits in einigen Publikationen unter besonderer Berücksichtigung des ordinalen Skalenniveaus vorgenommen (vgl. u.a. Beauducel & Herzberg, 2006; Chen et al., 2020; Li, 2016a; Liang & Yang, 2014). Gegenüber dem WLSMV-Schätzer stellt Li (2016a) heraus, dass unter Verwendung des *Robust Maximum Likelihood*-Schätzers (MLR), der eine multivariate Normalverteilung der beobachteten Variablen annimmt, in konfirmatorischen Faktorenanalysen mit erheblichen Unterschätzungen der Faktorladungen gerechnet werden muss – und zwar unabhängig von der Stichprobengröße oder der Anzahl der Antwortkategorien ordinaler Variablen. Durch die Verwendung des WLSMV können demnach die Faktorladungen korrekter geschätzt werden (vgl. Beauducel & Herzberg, 2006). Darüber hinaus wird der MLR-Schätzer bis zu einer Obergrenze von maximal vier Faktoren pro Modell empfohlen (Muthén et al., 2015) – eine Restriktion, die den komplexen Analysemodellen der vorliegenden Arbeit nicht gerecht wird. Nur mittels des WLSMV-Schätzers (nicht aber durch den MLR-Schätzer) kann zudem auf die THETA-Parametrisierung zurückgegriffen werden, um im Rahmen der Messinvarianzprüfung die Residualvarianzen der Parameter zur Prüfung von strikter Invarianz zwischen den Messzeitpunkten sowie im Mehrgruppenvergleich fixieren zu können (s. u.a. Kapitel 8.4.4). Hinsichtlich der adäquaten Berücksichtigung der Mehrebenenstruktur der Daten erlaubt der WLSMV-Schätzer das in *Mplus 7.11* (Muthén & Muthén, 2012) implementierte Verfahren für geclusterte Stichproben (*type = complex*) zu nutzen.

Bei sehr kleinen Stichproben von $n \leq 200$ kann es per WLSMV gegebenenfalls zu Verzerrungen der geschätzten Standardfehler (Muthén et al., 2015) sowie zu einer moderaten Überschätzung der Interfaktorkorrelationen (Li, 2016a) kommen. Zudem weisen Muthén et al. (2015) sowie Li (2016a) darauf hin, dass in kleinen Stichproben von $n \leq 200$ in WLSMV-Schätzungen eine Verschlechterung der Modell-Fits möglich ist. Zum Beispiel kann es zu einer überproportionalen Ablehnung der Modellpassung im Rahmen von χ^2 -Tests (Li, 2016a) kommen – insbesondere dann, wenn Variablen mit zwei oder drei Antwortkategorien verwendet werden (Beauducel & Herzberg, 2006). Dies wird bei der Interpretation der in dieser Arbeit präsentierten Ergebnisse berücksichtigt und neben dem χ^2 -Test ebenfalls alternative Modellgütekriterien zur Bewertung der Modell-Daten-Passung herangezogen (vgl. Kapitel 8.4.2).

8.3 Umgang mit fehlenden Werten

Grundsätzlich werden drei Herangehensweisen zum Umgang mit fehlenden Werten unterschieden: Neben den klassischen Verfahren, wie dem fall- oder paarweisen Ausschluss, können fehlende Werte mittels imputationsbasierter Verfahren ersetzt oder im Rahmen von modellbasierten Verfahren in der Modellschätzung berücksichtigt werden (Lüdtke, Robitzsch, Trautwein & Köller, 2007). Bei der Wahl des entsprechenden Verfahrens bedarf es zunächst der Feststellung, welcher Mechanismus den fehlenden Werten zu Grunde liegt. Enders (2010) definiert den Mechanismus fehlender Werte als möglichen Zusammenhang zwischen den beobachteten Variablen und der Wahrscheinlichkeit für das Auftreten fehlender Werte. Die Klassifizierung der Mechanismen fehlender Werte geht auf Rubin (1976) zurück und ist in der Literatur weit verbreitet. Demnach können drei Mechanismen unterschieden werden: *Missing Completely At Random* (MCAR) liegt vor, wenn die fehlenden Werte einer Variable sowohl unabhängig von den Werten der Variable an sich als auch unabhängig von den Werten der anderen gemessenen Variablen sind. Die fehlenden Werte sind demnach zufällig entstanden und „Personen mit fehlenden Werten stellen eine Zufallsstichprobe dar“ (Lüdtke et al., 2007, S. 104). MCAR kann beispielsweise durch das Forschungsdesign bedingt sein, in dem nicht alle Befragten alle Fragen gleichermaßen beantworten (z.B. Multi-Matrix-Design in einem Kompetenztest) oder entstehen, wenn einzelne Schülerinnen und Schüler am Erhebungstag (unvorhergesehen) nicht anwesend waren und den Fragebogen somit nicht ausfüllen konnten (Enders, 2010).

Etwas weniger restriktiv sind die Annahmen bei *Missing At Random* (MAR). Von MAR wird ausgegangen, wenn die Wahrscheinlichkeit für das Fehlen der Werte mit anderen beobachteten Variablen in Zusammenhang steht, nicht aber mit der Variable selbst, auf der die fehlenden Werte auftreten. Fehlende Werte auf einer Variable können folglich nicht durch die Variable an sich erklärt werden, sofern für die übrigen Variablen des Datensatzes kontrolliert wird. MAR geht entsprechend von einer systematischen Beziehung zwischen einer oder mehreren beobachteten Variablen und der Wahrscheinlichkeit für fehlende Werte aus (Enders, 2010). Im Gegensatz dazu wird bei *Missing Not At Random* (MNAR) angenommen, dass die fehlenden Werte einer Variable durch die Variable selbst bedingt sind, auch unter Kontrolle von anderen Variablen (Enders, 2010).

Little (2013) weist darauf hin, dass in den meisten Studien alle drei Mechanismen fehlender Werte vorliegen. Die methodischen Empfehlungen zum Umgang mit fehlenden Werten mittels imputationsbasierter oder modellbasierter Verfahren beziehen sich jedoch auf den MAR-Mechanismus beziehungsweise setzen diesen voraus. Im Forschungsprozess ist der MAR-Mechanismus zwar nicht identifizierbar, da keine statistischen Tests zur Verfügung stehen, um MAR zu belegen. Dennoch ist der MAR-Mechanismus forschungspraktisch vielfach plausibel (Enders, 2010).

Bei der multiplen Imputation werden fehlende Werte auf Basis beobachteter Informationen meist in durchschnittlich fünf bis zehn Imputationen durch geschätzte Werte ersetzt (van Buuren & Groothuis-Oudhoorn, 2011). Die Imputation der fehlenden Werte und die Spezifikation des Analysemodells findet dabei in zwei getrennten Analyseschritten statt. Multiple Imputationsmodelle werden als besonders flexibel beschrieben. Jedoch kann es zu Konflikten zwischen den Annahmen des Imputationsmodells und den Annahmen des Analysemodells kommen (von Hippel, 2015). Da die Forschungslage zur Kombination von multipler Imputation und des für ordinale Daten besonders robusten WLSMV-Schätzers aktuell noch sehr begrenzt ist (vgl. Chen et al., 2020; Teman, 2012), wurde auf die Imputation der fehlenden Werte in der vorliegenden Arbeit verzichtet.

Im modellbasierten Ansatz werden die Modellparameter geschätzt, in dem alle verfügbaren Daten (auch über Personen mit fehlenden Werten) genutzt werden (Lüdtke et al., 2007). Fehlende Werte werden folglich nicht ersetzt. Stattdessen wird für jeden Fall die Likelihood unter Berücksichtigung der beobachteten Werte geschätzt. Im Kontext der Strukturgleichungsmodellierung wird dieses Verfahren auch als *Full Information Maximum Likelihood* Schätzung (FIML) beschrieben (Arbuckle, 1996). In *Mplus 7.11* (Muthén & Muthén, 2012) werden standardmäßig alle verfügbaren Informationen für die jeweiligen Analysen verwendet. WLS-Schätzer, zu denen auch der in den vorliegenden Analysen genutzte WLSMV-Schätzer gehört, nutzen im ersten Schritt der Modellschätzung univariate *Full Information Maximum Likelihood* Schätzungen (Asparouhov & Muthén, 2010; s. Kapitel 8.2). Gleichwohl ist der WLSMV-Schätzer nicht per se „a full information estimation method“ (Chen et al., 2020, S. 91), sondern kann in Ausnahmefällen auf paarweisen Fallausschluss (*pairwise deletion*) zurückgreifen, sofern das Analysemodell keine exogenen Kovariaten enthält. Der paarweise Fallausschluss ist auch als „available-case analysis“ bekannt (Enders, 2010, S. 41) und versucht (z.B. im Gegensatz zur *listwise deletion*) den Verlust von Daten möglichst zu gering zu halten. Asparouhov und Muthén (2010) begründen das Vorgehen mit den Ergebnissen einer Simulationsstudie, in der sie herausstellten, dass mittels des paarweisen Fallausschlusses exakte Schätzungen der Modellparameter und der Standardfehler möglich und die Schätzungen mit paarweisem Fallausschluss wesentlich effizienter (und weniger verzerrt) sind als mittels fallweisem Ausschluss. Lüdtke et al. (2007) sowie Enders (2010) weisen allerdings darauf hin, dass es durch den paarweisen Fallausschluss zu einer Über- beziehungsweise Unterschätzung der Standardfehler kommen kann (Enders, 2010; Lüdtke et al., 2007).

Ebenfalls auf Basis von Simulationsstudien konnten Chen et al. (2020) bestätigen, dass im Rahmen von Messinvarianzanalysen die Schätzungen der Faktorladungen unter Verwendung des WLSMV-Schätzers nur geringfügig durch fehlende Werte beeinflusst werden und sich die Faktorladungen in ihrer Genauigkeit nicht unterscheiden beziehungsweise sogar weniger verzerrt sind als mittels des *Robust Maximum Likelihood*-Schätzers (MLR; vgl. Li, 2016a). In Bezug auf die Schätzung der Standardfehler weisen Chen et al. (2020) darauf hin, dass diese mittels WLSMV

akkurat geschätzt werden können, sofern die Anzahl der fehlenden Werte nicht allzu groß ist. Bei sehr kleinen Stichproben und sehr vielen fehlenden Werten kann der WLSMV-Schätzer gegebenenfalls zu Verzerrungen der Standardfehler führen (vgl. Enders, 2010). Allerdings ist unter Verwendung anderer *Full Information Maximum Likelihood*-Schätzer die Verzerrung der Standardfehler weitaus größer als unter Verwendung des WLSMV, da diese keine Anpassung an die fehlende Normalverteilung der ordinalen Daten vornehmen (Chen et al., 2020). Durch den Einsatz des WLSMV-Schätzers können zudem Fehler 1. Art begünstigt werden, sofern fehlende Daten vorliegen (Chen et al., 2020; Jia & Wu, 2019). Daher kann es zum Beispiel im Rahmen von Messinvarianztestungen in der Interpretation des χ^2 -Differenzentests fälschlicherweise zur Ablehnung eines Modells im Modellvergleich kommen (s. Kapitel 8.4.2). Allerdings ist in der für die Analysen verwendeten Programmversion *Mplus 7.11* (Muthén & Muthén, 2012) eine Korrektur des χ^2 -Differenzentests unter Verwendung des WLSMV-Schätzers implementiert (DIFFTEST-Option; s. dazu Kapitel 8.4.2) und wurde für die Modellvergleiche im Rahmen der Messinvarianztestungen entsprechend genutzt (Asparouhov & Muthén, 2006; vgl. Kapitel 8.4.4, 9.1.2 und 9.1.3).

Da Fälle mit vollständig fehlenden Werten keine Informationen für die zu analysierenden Variablen enthalten, wurden in den verschiedenen Modellen die entsprechenden Fälle von den Analysen ausgeschlossen (Geiser, 2010). In den vorliegenden Analysen wurden im Rahmen der konfirmatorischen Faktorenanalysen (s. Kapitel 9.1.1) und in den Messinvarianzprüfungen im Längsschnitt sowie im Multigruppenvergleich (s. Kapitel 9.1.2 und 9.1.3) zwischen einem und maximal 15 Fälle von den Analysen ausgeschlossen, die auf den in den jeweiligen Messmodellen betrachteten Variablen vollständig fehlende Werte aufwiesen. In den latenten Cross-Lagged-Panel-Modellen konnten maximal acht Fälle nicht in die Analysen einbezogen werden, da diese auf den verwendeten Variablen vollständig fehlende Werte aufwiesen (s. Kapitel 9.3). Die Anzahl ausgeschlossener Fälle ist somit als gering zu bezeichnen.

8.4 Methodisches Vorgehen

Die Strukturgleichungsmodellierung bildet den übergeordneten methodischen Rahmen der in dieser Arbeit durchgeführten Analysen und wird daher im Folgenden einführend beschrieben (s. Kapitel 8.4.1). Zur Beurteilung der analysierten Messbeziehungsweise Strukturmodelle kann auf verschiedene Modellgütekriterien (Fit-Indizes) zurückgegriffen werden, deren Berechnung und methodische Annahmen in Kapitel 8.4.2 näher erläutert werden. Zur Beantwortung der Forschungsfragen werden die folgenden Analyseverfahren angewendet und in den nächsten Abschnitten detailliert vorgestellt: Zunächst werden mittels konfirmatorischer Faktorenanalysen die zentralen latenten Konstrukte (domänenspezifisches Fach- und Sachinte-

resse sowie domänenspezifisches Fähigkeitsselbstkonzept) modelliert und statistisch geprüft (s. Kapitel 8.4.3). Um die Vergleichbarkeit der Konstrukte im Längsschnitt sowie zwischen verschiedenen Gruppen zu gewährleisten, wird in einem zweiten Schritt die Messinvarianz der sechs zentralen Konstrukte zwischen den Messzeitpunkten sowie zwischen den Vergleichsgruppen geprüft (s. Kapitel 8.4.4). In einem dritten Schritt werden schließlich die Längsschnittbeziehungen der sechs zentralen Konstrukte in latenten Cross-Lagged-Panel-Modellen untersucht (s. Kapitel 8.4.5). Die Modelle werden jeweils domänenspezifisch sowie für Profilklassen und Klassen ohne Profil getrennt spezifiziert, um die längsschnittliche Stabilität sowie die Zusammenhänge der zentralen Konstrukte zu analysieren und dabei den Effekt des Klassenprofils herauszustellen. In Kapitel 8.4.6 wird das analytische Vorgehen sowie die für die einzelnen Analyseschritte verwendete Software beschrieben.

8.4.1 Strukturgleichungsmodellierung

In der sozialwissenschaftlichen Forschung sind zentrale Untersuchungsmerkmale meist nicht direkt beobachtbar. Im Gegensatz zu den direkt messbaren (manifesten) Variablen (z.B. Alter, Körpergröße sowie Indikatorvariablen für latente Konstrukte) müssen die latenten Variablen stattdessen durch mehrere Indikatoren entsprechend operationalisiert werden (Bortz & Döring, 2006). Im sogenannten *Messmodell* wird das Zusammenhangsmuster zwischen den manifesten Variablen (Indikatorvariablen) und den latenten Variablen untersucht und damit die theoretisch angenommene faktorielle Struktur geprüft. Kann das Messmodell entsprechend identifiziert werden, wird in einem zweiten Schritt das *Strukturmodell* aufgestellt, das die Beziehung zwischen den latenten Variablen spezifiziert. Die Verknüpfung von Mess- und Strukturmodellen wird auch als *Strukturgleichungsmodellierung* (Structural Equation Modeling; SEM) bezeichnet (Christ & Schlüter, 2012; Wang & Wang, 2020). Mittels der Strukturgleichungsmodellierung können folglich sowohl theoretische Annahmen hinsichtlich der Messbeziehungen (Messmodell) als auch der strukturellen Beziehungen latenter Variablen (Strukturmodell) unter entsprechender Berücksichtigung von Messfehlern geprüft werden (vgl. Kline, 2010; Wang & Wang, 2020). Strukturgleichungsmodelle bieten außerdem diverse methodische Vorteile: Neben der simultanen Schätzung mehrerer abhängiger Variablen können sowohl direkte als auch indirekte Effekte modelliert und die Modellgüte der Mess- beziehungsweise Strukturmodelle anhand entsprechender Fit-Indizes beurteilt werden (vgl. Kapitel 8.4.2). Darüber hinaus können im Strukturgleichungsansatz Messinvarianzannahmen im Längsschnitt oder im Mehrgruppenvergleich berücksichtigt werden (vgl. Kapitel 8.4.4).

Die Strukturgleichungsmodellierung basiert auf der Analyse einer Kovarianz- oder Korrelationsmatrix und lässt sich im Grundsatz mit folgenden drei Gleichungen beschreiben:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (8.01)$$

$$Y = \Lambda_y \eta + \varepsilon \quad (8.02)$$

$$X = \Lambda_x \xi + \delta \quad (8.03)$$

Die erste Gleichung (8.01) definiert das Strukturmodell, die zweite Gleichung (8.02) das Messmodell der abhängigen Variablen und die dritte Gleichung (8.03) das Messmodell der unabhängigen Variablen. Dabei bezeichnet η die latente abhängige Variable und ξ die latente unabhängige Variable. B und Γ stehen für die Koeffizientenmatrix der latenten abhängigen beziehungsweise unabhängigen Variablen und bilden die Effekte der abhängigen Variablen auf andere abhängige Variablen (B) beziehungsweise die Effekte der unabhängigen auf die abhängigen Variablen (Γ) ab. ζ bezeichnet den Residualvektor. Die zweite (8.02) und dritte Gleichung (8.03) dient zur Spezifizierung der Messmodelle der latenten Variablen durch die beobachteten manifesten Variablen. Dabei werden die beobachteten abhängigen Variablen Y beziehungsweise unabhängigen Variablen X über die Faktorladungen Λ mit den latenten Variablen η beziehungsweise ξ verbunden. ε und δ sind die Messfehler der manifesten Variablen Y und X (Wang & Wang, 2020).

Zur Beurteilung der analysierten Mess- beziehungsweise Strukturmodelle kann auf eine Vielzahl unterschiedlicher Modellgütekriterien (Fit-Indizes) zurückgegriffen werden, die im Folgenden näher beschrieben werden.

8.4.2 Model-Fit

Für die Beurteilung der Modellgüte wird die Passung zwischen einem Mess- beziehungsweise Strukturmodell und den verwendeten Daten geprüft. Eine gute Modell-Daten-Passung (auch als Model-Fit bezeichnet) besteht folglich, wenn sich die im Modell spezifizierte Varianz-Kovarianz-Matrix nicht signifikant von der Kovarianzmatrix der beobachteten Daten unterscheidet. Ist dies der Fall, so kann angenommen werden, dass die im Modell theoretisch angenommenen Beziehungen zwischen den Variablen tatsächlich in den Daten zu finden sind. Zur Bestimmung der Modellgüte haben sich folgende Fit-Indizes in der Forschungsliteratur etabliert (vgl. Kline, 2010; Wang & Wang, 2020):

χ^2 -Teststatistik

Der χ^2 -Test ist eines der grundlegendsten und ältesten Kriterien für die Überprüfung der Modellgüte und wird nach folgender Gleichung berechnet:

$$\chi^2 = f_{ML}(N - 1) \quad (8.04)$$

Dabei steht f_{ML} für das Minimum der Fitfunktion des spezifizierten Modells und N für die Stichprobengröße (Wang & Wang, 2020). Wird der χ^2 -Wert signifikant, so muss die Nullhypothese abgelehnt werden, da keine Übereinstimmung zwischen der im Modell spezifizierten Kovarianzmatrix und der Kovarianzmatrix in der Population gefunden werden kann und das Modell demnach die Daten nur ungenügend widerspiegelt (Christ & Schlüter, 2012). Ein χ^2 -Wert von null zeigt einen perfekten Model-Fit an, während hohe χ^2 -Werte hingegen für einen schlechteren Model-Fit stehen. Im Modellvergleich (z.B. im Rahmen von Messinvarianzanalysen) sind nicht signifikante χ^2 -Werte ein Zeichen dafür, dass sich die beiden Modelle nicht signifikant voneinander unterscheiden.

Wang und Wang (2020) weisen darauf hin, dass der χ^2 -Test nur mit Einschränkungen zur Modellevaluation eingesetzt werden kann, da der χ^2 -Test stark abhängig von der Stichprobengröße ist (vgl. Svetina et al., 2020). Auch Christ und Schlüter (2012) sowie Kline (2010) betonen einerseits die Limitationen bei der Interpretation des χ^2 -Tests, verweisen andererseits aber auch auf die anhaltende Popularität der χ^2 -Teststatistik für den Vergleich unterschiedlich spezifizierter Modelle (z.B. im Rahmen von Messinvarianzanalysen) mittels des χ^2 -Differenztests.

Da unter Verwendung des WLSMV-Schätzers die Differenz der χ^2 -Werte zweier verschachtelter Modelle (z.B. konfigurale vs. metrische Messinvarianz) nicht der χ^2 -Verteilung entspricht, kann für den WLSMV-Schätzer nicht der in *Mplus 7.11* standardmäßig implementierte χ^2 -Differenztest verwendet werden. Stattdessen wird für den Vergleich verschiedener Messinvarianzmodelle die in *Mplus 7.11* implementierte DIFFTEST-Option genutzt, um den χ^2 -Differenztest entsprechend adäquat zu berechnen (Asparouhov & Muthén, 2006). Für einen korrekt geschätzten χ^2 -Differenztest wird in einem zweistufigen Verfahren das H0-Analysemodell mit einem weniger restriktiveren H1-Alternativmodell, in dem das H0-Modell verschachtelt ist, verglichen. Im ersten Schritt wird das weniger restriktivere H1-Modell geschätzt und die benötigten Ableitungen in einer externen .dat-Datei gespeichert. Anschließend wird im zweiten Schritt das restriktivere H0-Modell geschätzt und der χ^2 -Differenztest unter Berücksichtigung der Ableitungen aus beiden Modellen berechnet (Muthén & Muthén, 2012). Unter Verwendung des WLSMV wird der χ^2 -Differenztest folgendermaßen berechnet:

$$\Delta\chi^2_{WLSMV} = N \times m \times (F_{DWLS}(\theta_A) - F_{DWLS}(\theta_B)) + l \quad (8.05)$$

Dabei nähern sich mittels der Parameter m und l der Mittelwert und die Varianz von $\Delta\chi^2_{WLSMV}$ der χ^2 -Verteilung an, wobei die Freiheitsgrade (df) der Differenz der Anzahl der Parameter zwischen Modell A und Modell B entsprechen. $F_{DWLS}(\theta_A)$ und $F_{DWLS}(\theta_B)$ entsprechen den Anpassungsfunktionen der Modelle A und B, sind aber nicht mit der χ^2 -Verteilung der beiden Modelle gleichzusetzen (Chen et al., 2020).

Unter Verwendung des WLSMV-Schätzers kann es im Rahmen der χ^2 -Teststatistik vermehrt zu Fehlern 1. Art kommen und damit insbesondere in kleineren Stichproben tendenziell eine Ablehnung des Modells bedingen (Li, 2016a). Die χ^2 -Teststatistik wird daher nur unter Vorbehalt als Modellgütekriterium zur Beurteilung von Strukturgleichungsmodellen herangezogen, um Fehlinterpretationen zu vermeiden (Wang & Wang, 2020). Für die Analyse von Strukturgleichungsmodellen kann ergänzend auf eine Vielzahl alternativer Modellgütekriterien zurückgegriffen werden, die im Folgenden vorgestellt werden.

CFI und TLI

Neben dem *Comparative-Fit-Index* (CFI; Bentler, 1990) gehört auch der *Tucker-Lewis-Index* (TLI; Tucker & Lewis, 1973) zu den inkrementellen Fit-Indizes. Sowohl der CFI als auch der TLI vergleichen die Datenpassung des spezifizierten Modells mit der Datenpassung eines Unabhängigkeitsmodells (Null-Modells), in dem angenommen wird, dass zwischen den manifesten Variablen keine Zusammenhänge (d.h. Kovarianzen) bestehen (Christ & Schlüter, 2012; Geiser, 2010; Wang & Wang, 2020).

CFI und TLI sind normiert und können Werte zwischen null und eins annehmen, wobei ein höherer Wert eine bessere Modell-Daten-Passung anzeigt (Christ & Schlüter, 2012; Geiser, 2010). Nach Hu und Bentler (1999) kann bei CFI- und TLI-Werten von mindestens $\geq .90$ (idealerweise aber von $\geq .95$) von einem guten Model-Fit ausgegangen werden (vgl. Little, 2013).

Die Berechnung des CFI basiert auf dem *non-centrality parameter* d ($\chi^2 - df$), wobei df die Freiheitsgrade des Modells anzeigen. Der CFI wird folgendermaßen bestimmt:

$$CFI = \frac{d_{null} - d_{specified}}{d_{null}} \quad (8.06)$$

d_{null} bezeichnet den *non-centrality parameter* des Null-Modells

$d_{specified}$ bezeichnet den *non-centrality parameter* des spezifizierten Modells

Im Falle einer idealen Passung zwischen den Daten und dem spezifizierten Modell ist $d_{\text{specified}}$ nahe null und damit der CFI = 1. Der CFI eignet sich auch für kleine Stichprobengrößen (Wang & Wang, 2020).

Im Gegensatz zum CFI wird bei der Berechnung des TLI die Komplexität des Modells über die Freiheitsgrade (df) berücksichtigt. Sparsamere Modelle werden dabei mit einem höheren TLI belohnt. Im direkten Vergleich zum CFI fällt der TLI meist etwas niedriger aus. Der TLI wird nach folgender Gleichung bestimmt:

$$\text{TLI} = \frac{\left(\frac{\chi^2_{\text{null}}}{df_{\text{null}}} - \frac{\chi^2_{\text{specified}}}{df_{\text{specified}}} \right)}{\left(\frac{\chi^2_{\text{null}}}{df_{\text{null}}} - 1 \right)} \quad (8.07)$$

$\frac{\chi^2_{\text{null}}}{df_{\text{null}}}$ und $\frac{\chi^2_{\text{specified}}}{df_{\text{specified}}}$ beschreiben das Verhältnis zwischen der χ^2 -Teststatistik und den Freiheitsgraden (df) des Null-Modells beziehungsweise des spezifizierten Modells (Wang & Wang, 2020).

Im Rahmen von Messinvarianzanalysen sollten im Modellvergleich nach Chen (2007) sowie Cheung und Rensvold (2002) die Veränderungen des CFI (ΔCFI) nicht größer als .010 sein, um Messinvarianz zu bestätigen. In Abhängigkeit der jeweiligen Substichproben sollte im Mehrgruppen-CFA-Ansatz außerdem der $\Delta\text{CFI} \leq .010$ und der $\Delta\text{RMSEA} \leq .015$ sein, wenn die Substichprobengröße mindestens $n = 300$ Personen beträgt. Für Substichproben kleiner als $n = 300$ Personen empfiehlt Chen (2007) $\Delta\text{CFI} \leq .005$ und $\Delta\text{RMSEA} \leq .010$ zuzulassen, um Messinvarianz zwischen den Gruppen zu bestätigen.

RMSEA

Neben dem CFI und dem TLI ist der *Root Mean Square of Approximation* (RMSEA; Steiger, 1998) einer der Fit-Indizes, der am häufigsten zur Beurteilung der Modellgüte genutzt wird. Der RMSEA wird auch als „badness-of-fit-index“ (Christ & Schlüter, 2012, S. 39) bezeichnet, da dieser angibt, wie groß die Abweichung zwischen dem spezifizierten Modell und der Grundgesamtheit ist. Kleinere RMSEA-Werte kennzeichnen folglich einen besseren Model-Fit. Die Berechnung des RMSEA erfolgt auf Grundlage des sogenannten *non-centrality parameters*. Zudem wird die Stichprobengröße reskaliert, wie der Term $(\chi^2_s - df_s) / N$ zeigt. Dem RMSEA liegt folgende Gleichung zugrunde:

$$\text{RMSEA} = \sqrt{\frac{(\chi^2_s - df_s) / N}{df_s}} = \sqrt{\frac{(\chi^2_s - df_s) - 1}{N}} \quad (8.08)$$

Der Wertebereich des RMSEA liegt zwischen null und eins, wobei ein RMSEA-Wert von null einer perfekten Passung zwischen spezifiziertem Modell und der Grundgesamtheit entspricht. Der RMSEA-Wert sollte idealerweise $\leq .05$ sein, um einen guten Model-Fit anzuzeigen. RMSEA-Werte $\leq .08$ werden als zufriedenstellend bewertet (Wang & Wang, 2020). Hu und Bentler (1999) empfehlen einen RMSEA-Wert von $\leq .06$ als *Cutoff-Kriterium*.

SRMR

Ein weiterer Indikator zur Beurteilung der Modellgüte ist der *Standardized Root Mean Square Residual* (SRMR). Der SRMR „ist ein standardisiertes Maß zur Gesamtbewertung der Residuen (Stichprobenkennwerte minus modellimplizierte Kennwerte)“ (Geiser, 2010, S. 61). Die Verwendung des für ordinale Daten robusten WLSMV-Schätzers schließt die Berechnung des SRMR in *Mplus 7.11* (Muthén & Muthén, 2012) jedoch aus. Für die Beurteilung der Modellgüte der analysierten Mess- beziehungsweise Strukturmodelle werden folglich der χ^2 -Test, CFI, TLI und RMSEA als Fit-Indizes genutzt.

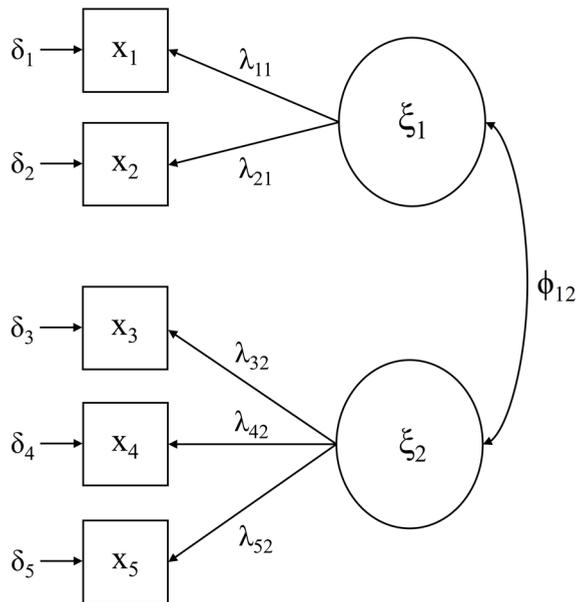
8.4.3 Konfirmatorische Faktorenanalyse

Wie in Kapitel 8.4.1 beschrieben, werden im Rahmen der Strukturgleichungsmodellierung in einem ersten Schritt zunächst die sogenannten Messmodelle spezifiziert, in denen das Zusammenhangsmuster zwischen den manifesten Variablen und den latenten Variablen modelliert wird. Um die theoretisch hergeleitete faktorielle Struktur, das heißt neben der Anzahl der latenten Variablen auch das spezifische Ladungsmuster der Indikatoren (manifesten Variablen) auf die latenten Konstrukte empirisch zu überprüfen, wird die *konfirmatorische Faktorenanalyse* (Confirmatory Factor Analysis; CFA) genutzt. Das im Rahmen der konfirmatorischen Faktorenanalyse definierte Messmodell gibt Aufschluss darüber, durch welche latenten Variablen die Varianz und Kovarianz der manifesten Variablen erklärt werden kann (Wang & Wang, 2020). Auf Basis der theoretischen Annahmen aus dem Messmodell können die entsprechenden Modellparameter geschätzt und anhand verschiedener Gütekriterien (Fit-Indizes) der sogenannte Model-Fit, das heißt die Passung zwischen dem Messmodell und den Daten, überprüft werden (Christ & Schlüter, 2012). Die konfirmatorische Faktorenanalyse wird anhand folgender Gleichung berechnet:

$$x_j = \lambda_j \xi + \delta_j \quad (j = 1, \dots, n) \quad (8.09)$$

Wobei x_j die Indikatorvariablen, λ_j die Faktorladungen, ξ die latente Variable und δ_j den Messfehler darstellen (Backhaus, Erichson, Plinke & Weiber, 2018).

Abbildung 8.02: Messmodell – konfirmatorische Faktorenanalyse



Wang & Wang, 2020, Abbildung 1.2(a), modifiziert

Wie Abbildung 8.02 zeigt, stellen die Faktorladungen λ_j die Verbindung zwischen den manifesten Variablen x_j und den latenten Variablen ξ her. Φ steht für die Kovarianz zwischen den latenten Variablen. Ist $\Phi = 0$, so sind die latenten Variablen unkorreliert (Wang & Wang, 2020). Für die Berechnung von konfirmatorischen Faktorenanalysen mit ordinalen Indikatoren sind einige Besonderheiten zu beachten. Diese werden im folgenden Abschnitt näher beleuchtet.

CFA-Modelle mit ordinalen Indikatoren

Wie in Kapitel 8.2 beschrieben, stellen ordinal-kategoriale Indikatoren besondere Anforderungen an die verwendeten Analyseverfahren, insbesondere wenn Likert-Skalen mit weniger als fünf Antwortkategorien vorliegen und das Skalenniveau adäquat berücksichtigt werden soll (Rhemtulla et al., 2012; Wang & Wang, 2020). Unter Verwendung des WLSMV-Schätzers wird die Beziehung zwischen den Indikatorvariablen und den Faktoren über eine Probit-Regression geschätzt (Muthén & Muthén, 2012). Die Berechnung von CFA-Modellen mit ordinalen Indikatoren erfolgt in einem mehrstufigen Verfahren. In einem ersten Schritt werden die Schwellenwerte der ordinalen Variablen unter Verwendung des *Maximum Likelihood* geschätzt. Anschließend werden die latenten Korrelationen geschätzt – im Falle von rein ordinalen Indikatoren sind diese als *Polychoric-Correlations* zu bezeichnen. Drittens wird die asymptotische Varianz-Kovarianz-Matrix der latenten Korrelationen

geschätzt. Viertens wird die inverse Varianz-Kovarianz-Matrix als Gewichtung des *Weighted Least Squares* (WLS)-Schätzers (bzw. des WLSMV-Schätzers) zur Bestimmung der Parameter genutzt (Wang & Wang, 2020). Diese vier Schritte werden in *Mplus 7.11* (Muthén & Muthén, 2012) automatisch umgesetzt, sofern das ordinale Skalenniveau entsprechend im *Variable-Command* definiert wurde.

Für die vorliegenden Analysen werden jeweils für die zentralen latenten Konstrukte (domänenspezifisches Fach- und Sachinteresse sowie domänenspezifisches Fähigkeitsselbstkonzept) konfirmatorische Faktorenanalysen unter Berücksichtigung des ordinalen Skalenniveaus und unter Verwendung des WLSMV-Schätzers berechnet, um die faktorielle Validität der verwendeten Konstrukte zu prüfen. Dabei wird die Analysetichprobe (s. Kapitel 8.1.1) verwendet und die CFA-Modelle für die drei Messzeitpunkte berechnet. In einem zweiten Schritt wird anschließend die Vergleichbarkeit der zentralen Konstrukte zwischen den verschiedenen Subpopulationen (musikalische bzw. naturwissenschaftliche Profilklassen und Vergleichsklassen ohne Profil) sowie in der längsschnittlichen Betrachtung im Rahmen von Messinvarianzanalysen geprüft. Das Verfahren der Messinvarianztestung wird im nächsten Kapitel ebenfalls unter besonderer Berücksichtigung des ordinalen Datenniveaus näher beschrieben. Dabei werden zum einen verschiedene methodische Ansätze zur Prüfung von Messäquivalenz gegenübergestellt und zum anderen Unterschiede in den methodischen Herangehensweisen zwischen Messinvarianztestungen mit ordinal-kategorialen Daten und mit kontinuierlichen Daten herausgestellt.

8.4.4 Messinvarianz

Fragen der Vergleichbarkeit und damit der Messinvarianz (synonym auch als Messäquivalenz bezeichnet) sind immer dann von Bedeutung, wenn unterschiedliche Kontextbedingungen (z.B. Kulturen, Länder, Gruppenzugehörigkeit, Messzeitraum, Alter, Geschlecht) die untersuchten Zusammenhänge zwischen den latenten Variablen und deren Indikatoren beeinflussen können. Daher wird Messinvarianz insbesondere auch in der Tradition der interkulturellen Forschung (Rippl & Seipel, 2015), im Kontext der international vergleichenden Large-Scale Assessments (vgl. u.a. He, Buchholz & Klieme, 2017; Rutkowski & Svetina, 2014, 2017; Schulte, Nonte & Schwippert, 2013) sowie unter Aspekten der Konstruktstabilität im Rahmen von Längsschnittanalysen diskutiert (Meredith & Horn, 2001; Widaman, Ferrer & Conger, 2010; Wirth, 2008).

Um Unterschiede zwischen Gruppen sowie Entwicklungsverläufe über die Zeit abbilden zu können, wird eine Vergleichbarkeit der verwendeten Konstrukte zwischen diesen Gruppen beziehungsweise über die Zeit vorausgesetzt (Liu et al., 2017). Dahinter steht die Annahme, dass die eingesetzten Messinstrumente in unterschiedlichen Subpopulationen ebenso wie im Zeitverlauf invariant sind. Fehlende Messinvarianz kann zu einer Verzerrung der Ergebnisse führen und damit

Fehlinterpretationen bedingen, da unklar bleibt, ob es sich bei den Beobachtungen um systematische Unterschiede oder aber methodische Artefakte handelt (Liu et al., 2017; Rippl & Seipel, 1997; Temme & Hildebrandt, 2009).

Die eingesetzten Messinstrumente sollten sich sowohl hinsichtlich der zugrundeliegenden theoretischen Überlegungen (Konstruktäquivalenz) und ihrer Operationalisierung (operationale Äquivalenz) als auch hinsichtlich des verwendeten Messmodells (Messäquivalenz oder auch Messinvarianz) nicht unterscheiden (Schulte et al., 2013).

Konstruktäquivalenz ist gegeben, wenn sich die Konstrukte in ihrer inhaltlichen Bedeutung nicht zwischen den Untersuchungsgruppen oder Messzeitpunkten unterscheiden. Von *operationaler Äquivalenz* kann ausgegangen werden, wenn die Operationalisierung der Konstrukte in allen Gruppen oder zu allen Messzeitpunkten vergleichbar ist, das heißt, jeweils die gleichen Indikatoren verwendet wurden (Rippl & Seipel, 1997; Schulte et al., 2013).

Messinvarianz kann mittels verschiedener statistischer Verfahren geprüft werden. Neben dem sogenannten *Differential Item Functioning* (DIF), das im Kontext der *Item Response Theory* (IRT) zu verorten ist, hat sich die Verwendung von *konfirmatorischen Faktorenanalysen* (Confirmatory Factor Analysis; CFA), die der Strukturgleichungsmodellierung zuzuordnen sind, etabliert. Zahlreiche Publikationen haben sich bereits dem methodischen Vergleich beider Ansätze angenommen (s. u.a. Kim & Yoon, 2011; Meade & Lautenschlager, 2004; Nonte, Knigge & Stubbe, 2019; Penfield & Lam, 2000; Schwab & Helm, 2015; Stark, Chernyshenko & Drasgow, 2006).

Im Gegensatz zum linearen CFA-Ansatz wird im alternativen DIF-Ansatz von einem logistischen, nicht zwingend linearen Zusammenhang zwischen dem Antwortverhalten der Befragten und dem latenten Konstrukt ausgegangen (Nonte et al., 2019; Schwab & Helm, 2015). Im IRT-Kontext wird fehlende Messäquivalenz als *Different Functioning* bezeichnet. Tritt dies auf Itemebene auf, wird von *Differential Item Functioning* (DIF) gesprochen (Stark et al., 2006). Die Vorzüge des nicht linearen DIF-Ansatzes, die unter anderem in der Schätzung der Antwortkategorien (Schwellenwerte) liegen, lassen sich jedoch auch per CFA-Ansatz umsetzen, sofern das ordinale Skalenniveau der Daten entsprechend berücksichtigt wird (vgl. u.a. Kim & Yoon, 2011). Nonte et al. (2019) weisen daher auf zahlreiche Vorteile des CFA-Ansatzes hin: Gegenüber des DIF-Ansatzes „[verfolgen] Messinvarianzanalysen im Rahmen von CFAs [...] einen globaleren Ansatz, der eine Ausweitung der jeweiligen Modelle im Rahmen des Strukturgleichungsansatzes ermöglicht und Beziehungen zwischen Faktoren zulässt“ (Nonte et al., 2019, S. 6). Für die Beurteilung der Modellgüte kann mittels des CFA-Ansatzes zudem auf eine Vielzahl von Fit-Indizes (s. Kapitel 8.4.2) zurückgegriffen werden. Darüber hinaus bietet die CFA für die Messinvarianzprüfung im Mehrgruppenvergleich, das heißt der Messinvarianzprüfung zwischen mehreren Substichproben, Vorteile gegenüber des IRT-Ansatzes, der hingegen „hinsichtlich Einzelitems oder einzelner Tests mehr psychometrische Informationen als CFAs [liefert]“ (Nonte et al., 2019, S. 6; vgl. Meade & Lautenschlager, 2004).

Auf Grundlage der beschriebenen Vorteile des CFA-Ansatzes, insbesondere der Möglichkeit, die mittels CFAs spezifizierten Messmodelle im Rahmen der Strukturgleichungsmodellierung zu erweitern und darüber hinaus die Beziehungen zwischen den Faktoren zu untersuchen, wird in der vorliegenden Arbeit der CFA-Ansatz für die Messinvarianztestung genutzt. Da die Überprüfung der Messinvarianz die Grundlage für weiterführende Analysen bildet (zum Beispiel im Rahmen der Cross-Lagged-Panel-Modelle), bietet die Strukturgleichungsmodellierung, bestehend aus Mess- und Strukturmodell (s. Kapitel 8.4.1), einen umfassenden und integrierenden Ansatz und stellt somit ein geeignetes analytisches Vorgehen dar.

Im Folgenden werden zunächst die Grundlagen der Messinvarianzprüfung mittels des CFA-Ansatzes dargestellt und anschließend die Besonderheiten der Messinvarianzprüfung für ordinale Daten näher beleuchtet. Dabei wird in einem ersten Schritt das Vorgehen einer Messinvarianztestung mit kontinuierlichen Indikatoren vorgestellt, um anschließend kontrastierend die nötigen Modellanpassungen für ordinale Daten herausstellen zu können.

Ebenen der Messinvarianz

Für die Überprüfung von Messinvarianz bei kontinuierlichen Indikatoren identifizieren Widaman et al. (2010) auf Basis vorheriger Arbeiten von Jöreskog (1971), Meredith (1993) sowie Widaman und Reise (1997) vier Stufen faktorieller Invarianz zwischen Gruppen. Der Ansatz nach Widaman und Reise (1997) ist zudem auf die Messinvarianztestung für verschiedene Messzeitpunkte übertragbar. Die vier Stufen faktorieller Invarianz unterliegen einer hierarchischen Struktur und unterscheiden sich hinsichtlich der Restriktivität der Modellannahmen.

Folgende vier Stufen werden nach Widaman et al. (2010) für die Messinvarianzprüfung mit kontinuierlichen Indikatoren unterschieden:

- Die *konfigurale Invarianz* (auch Basismodell genannt) ist die am wenigsten restriktive Form der Messinvarianz. Zwischen den Gruppen oder Messzeitpunkten besteht eine invariante Faktorstruktur, das heißt, die Anzahl der Faktoren und das Ladungsmuster (d.h. die Indikatoren pro Faktor) sind zwischen den Vergleichsgruppen oder über die Messzeitpunkte hinweg gleich.
- *Metrische faktorielle Invarianz* (auch schwache Invarianz genannt) ist gegenüber der konfiguralen Invarianz restriktiver und besteht, wenn zusätzlich zu den Grundannahmen der konfiguralen Invarianz auch die Faktorladungen zwischen den Gruppen beziehungsweise über die Messzeitpunkte invariant sind.
- *Skalare faktorielle Invarianz* (auch strenge Invarianz genannt) wird angenommen, wenn zusätzlich zu den Annahmen der konfiguralen sowie metrischen Invarianz (invariante Faktorladungen) keine Unterschiede in den Intercepts der manifesten Variablen zwischen den Gruppen oder Messzeitpunkten bestehen.

- *Strikte faktorielle Varianz* prüft, ob zusätzlich zu den invarianten Faktorladungen und Intercepts auch die Residualvarianzen zwischen Gruppen beziehungsweise zwischen verschiedenen Messzeitpunkten invariant sind.

Im Rahmen von konfirmatorischen Faktorenanalysen lassen sich die verschiedenen Messinvarianzebenen schrittweise prüfen, indem die Messmodelle mit zunehmenden Restriktionen spezifiziert und hinsichtlich möglicher Änderungen in den Gütekriterien (χ^2 -Test, χ^2 -Differenztest sowie anhand absoluter sowie inkrementeller Fit-Indizes; s. Kapitel 8.4.2) geprüft werden. Als Voraussetzung für den Vergleich von latenten Mittelwerten zwischen verschiedenen Gruppen oder Messzeitpunkten gilt die Bestätigung von skalarer faktorieller Messinvarianz. Sollen hingegen vornehmlich Beziehungen zwischen latenten Konstrukten untersucht werden, kann der Nachweis von metrischer Invarianz ausreichend sein (Brown, 2015; Christ & Schlüter, 2012). Svetina et al. (2020) sowie Meredith (1993) weisen zudem darauf hin, dass strikte faktorielle Invarianz in der Forschungspraxis nur selten nachgewiesen werden kann.

Wie in Kapitel 8.2 beschrieben, kann die Annahme eines kontinuierlichen Skalenniveaus bei ordinalen Daten (z.B. Likert-Skalen mit vier oder fünf Antwortoptionen) jedoch möglicherweise falsche Schlussfolgerungen implizieren (vgl. u.a. Lubke & Muthén, 2004; Svetina & Rutkowski, 2017; Svetina et al., 2020). Für die vorliegenden Analysen soll das ordinale Skalenniveau der Daten daher auch bei der Prüfung von Messinvarianz adäquat beachtet werden. Welche Modifikationen im analytischen Vorgehen durch die Berücksichtigung des ordinalen Datenniveaus bei der Prüfung von Messinvarianz zwischen verschiedenen Messzeitpunkten beziehungsweise unterschiedlichen Gruppen nötig sind, wird im Folgenden detailliert aufgeschlüsselt.

Messinvarianz im Längsschnitt mit ordinalen Indikatoren

Um Entwicklungen über die Zeit abbilden zu können, wird vorausgesetzt, dass die jeweils betrachteten Konstrukte über die verschiedenen Messzeitpunkte auch invariant gemessen werden. In Anlehnung an das von Liu et al. (2017) empfohlene Vorgehen, wird für die Prüfung längsschnittlicher Messinvarianz mit ordinal-kategorialen Indikatoren ein Modell entwickelt, das möglichst analog zu Modellen der Messinvarianztestung mit kontinuierlichen Indikatoren (konfigurale, metrische, skalare und strikte Messinvarianz) aufgebaut ist (vgl. Meredith, 1993; Widaman et al., 2010). Entsprechend wird ebenfalls ein hierarchisches Modell entwickelt, das mit jeweils zunehmenden Restriktionen vier Ebenen der Messinvarianz testet. Im (1) *Basismodell* (entspricht der konfiguralen Messinvarianz) wird geprüft, ob die gleiche Faktorstruktur über die verschiedenen Messzeitpunkte angenommen werden kann, das heißt, neben der Anzahl der spezifizierten Faktoren auch die Indikatoren pro Faktoren über die Zeit identisch sind. Weist das Basismodell einen guten Model-Fit auf,

werden die weiteren Stufen der (2) *invarianten Faktorladungen* (metrische Messinvarianz), (3) *invarianten Thresholds* (skalare Messinvarianz) und (4) *invarianten Residualvarianzen* (strikte Messinvarianz) überprüft. Die folgenden Bedingungen dienen zur Identifizierung des (1) *Basismodells*:

Zu Identifikationszwecken wird pro Faktor jeweils die erste Faktorladung auf eins fixiert und die übrigen Faktorladungen werden frei geschätzt. Pro Faktor dient die erste manifeste Variable folglich als Marker-Variable. Für jeden Indikator wird außerdem jeweils der erste Schwellenwert (Threshold) zwischen den Messzeitpunkten gleichgesetzt sowie für die drei Marker-Variablen der erste und zweite Schwellenwert über die Zeit fixiert. Für den ersten Messzeitpunkt wird der Faktormittelwert auf null fixiert; die übrigen zwei Faktormittelwerte zu Messzeitpunkt zwei und drei werden frei geschätzt. Ebenfalls frei geschätzt werden die Faktorvarianzen über alle drei Messzeitpunkte. Für den ersten Messzeitpunkt werden die Residualvarianzen der manifesten Variablen auf eins fixiert; die Residualvarianzen der Variablen des zweiten und dritten Messzeitpunkts können frei geschätzt werden. Durch die Autokorrelation der Residuen eines manifesten Indikators über die Zeit kann außerdem die Korrelation zwischen den Messfehlern im Längsschnitt berücksichtigt werden.

Im zweiten Schritt wird das Basismodell gegen das Modell der *invarianten Faktorladungen* getestet. Zu den für das Basismodell aufgeführten Bedingungen wird eine zusätzliche Restriktion eingeführt: Es werden nun die Faktorladungen über die drei Messzeitpunkte fixiert. Für das dritte Modell der *invarianten Thresholds* werden zusätzlich zu den Faktorladungen auch die Schwellenwerte über die Messzeitpunkte konstant gehalten. Äquivalent zum Modell der strikten Messinvarianz für kontinuierliche Daten werden für die vierte Stufe der Messinvarianz mit ordinalen Daten (*invariante Residualvarianzen*) zusätzlich zu den Faktorladungen und Schwellenwerte auch die Residualvarianzen über die drei Messzeitpunkte konstant gehalten (Liu et al., 2017).

Die Messinvarianzanalysen werden für die drei zentralen Konstrukte (Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept) jeweils domänenspezifisch über die drei Messzeitpunkte modelliert. Dazu werden die jeweiligen Messmodelle mit zunehmenden Restriktionen spezifiziert und hinsichtlich möglicher Änderungen in den Gütekriterien (χ^2 -Test, χ^2 -Differenztest sowie anhand absoluter sowie inkrementeller Fit-Indizes; s. Kapitel 8.4.2) geprüft (vgl. Chen, 2007; Cheung & Rensvold, 2002). Die Mehrebenenstruktur der Daten wird mittels des in *Mplus 7.11* implementierten Verfahrens für geclusterte Stichproben (*type = complex*) berücksichtigt. Statt auf die DELTA-Parametrisierung, die standardmäßig in *Mplus 7.11* verwendet wird, wird auf die THETA-Parametrisierung zurückgegriffen, um die Residualvarianzen der Parameter im vierten Schritt der Messinvarianzprüfung (strikte Messinvarianz) zwischen den Messzeitpunkten fixieren zu können.

Messinvarianz im Gruppenvergleich mit ordinalen Indikatoren

Obwohl sich ordinale Messinvarianzmodelle im Allgemeinen bereits gut etabliert haben (vgl. u.a. Jöreskog, 1993; Millsap, 2011; Muthén & Asparouhov, 2002), sind diese für den Mehrgruppenvergleich mit einer Reihe besonderer Anforderungen sowie vielfältigen Empfehlungen zur konkreten methodischen Umsetzungen verbunden (Millsap & Yun-Tein, 2004; Svetina et al., 2020; Wu & Estabrook, 2016).

Die verschiedenen Stufen der Messinvarianz lassen sich im Rahmen von multiplen Gruppenvergleichen mittels konfirmatorischer Faktorenanalysen auch für ordinal-kategoriale Daten prüfen, indem schrittweise in der Modellierung der Messmodelle verschiedene Gleichheitsrestriktionen angenommen werden. Millsap und Yun-Tein (2004) empfehlen für die Prüfung von Messinvarianz zwischen Gruppen folgendes Vorgehen: Für den Mehrgruppenvergleich werden zunächst für jede Substichprobe (Gruppe) separate Faktorladungen, Schwellenwerte (Thresholds) und Varianz-Kovarianz-Matrizen geschätzt. Konkret werden für die Schätzung des Basismodells:

- die Faktorladungen in allen Gruppen frei geschätzt; mit Ausnahme der jeweils ersten Ladung pro Faktor, die auf eins fixiert wird (Marker-Variable).
- die Schwellenwerte der manifesten Variablen frei geschätzt; mit Ausnahme der Schwellenwerte der Marker-Variablen sowie der jeweils der ersten Antwortkategorie, die zwischen den Gruppen fixiert werden.
- die Faktormittelwerte in der ersten Gruppe auf null fixiert und in den übrigen Gruppen frei geschätzt.
- die Residualvarianzen der manifesten Variablen in der ersten Gruppe auf eins fixiert, während diese in den anderen Gruppen frei geschätzt werden.
- die Faktorvarianzen in allen Gruppen frei geschätzt.

Im zweiten Schritt (entspricht metrischer Messinvarianz) wird zunächst die Invarianz der Faktorladungen zwischen den Gruppen getestet, indem zusätzlich zur bereits fixierten ersten Faktorladung im Basismodell nun auch die übrigen Faktorladungen zwischen den Gruppen fixiert werden. Die Faktormittelwerte der ersten Gruppe bleiben auf null fixiert und werden in den anderen Gruppen frei geschätzt. Auch für die Residualvarianzen bleibt die Fixierung in der ersten Gruppe bestehen, während diese in den übrigen Gruppen frei geschätzt werden. Die Faktorvarianzen bleiben in allen Gruppen unrestringiert.

Im dritten Schritt (entspricht skalarer Messinvarianz) werden zusätzlich zu den Faktorladungen die Schwellenwerte (Thresholds) zwischen den Gruppen gleichgesetzt und im vierten Schritt (entspricht strikter Messinvarianz) schließlich die Residualvarianzen zwischen den Gruppen fixiert (Millsap & Yun-Tein, 2004).

Die Messinvarianzanalysen im Mehrgruppenvergleich werden für die drei zentralen Konstrukte jeweils domänenspezifisch unter Berücksichtigung der Mehrebenenstruktur (*type = complex*) sowie unter Verwendung der THETA-Parametrisierung

berechnet. Dazu werden die jeweiligen Messmodelle mit zunehmenden Restriktionen spezifiziert und hinsichtlich möglicher Änderungen in den Gütekriterien (χ^2 -Test, χ^2 -Differenztest sowie anhand absoluter sowie inkrementeller Fit-Indizes; s. Kapitel 8.4.2) verglichen.

8.4.5 Cross-Lagged-Panel-Modelle

Statistische Analysen von Paneldaten im Rahmen der Strukturgleichungsmodellierung können mittels verschiedener Ansätze durchgeführt werden. Neben latenten Wachstumskurvenmodellen haben sich insbesondere *Cross-Lagged-Panel-Modelle* (CLP) etabliert, die zu den autoregressiven Modellen gehören (Newsom, 2015). In autoregressiven Modellen wird angenommen, dass Beobachtungen eines Attributs zu einem vorherigen Messzeitpunkt die Beobachtungen desselben Attributs zu einem späteren Zeitpunkt vorhersagen können (*autoregressive Beziehung*). Zudem können vorherigen Beobachtungen anderer Merkmale die Beobachtungen des Attributs zu einem späteren Zeitpunkt erklären (*Cross-Lagged-Beziehung*; Christ & Schlüter, 2012). Die Grundidee des Cross-Lagged-Panel-Designs stammt von dem Soziologen Paul Lazarsfeld (1940).

Für die Modellierung von CLP-Modellen sind mindestens zwei, besser aber drei Messzeitpunkte nötig, um die Modell-Daten-Passung und damit die empirische Aussagekraft zu prüfen (Kühnel & Mays, 2019). Reinders (2006) formuliert fünf Grundannahmen, die bei der Modellierung und Interpretation von CLP-Modellen zu bedenken sind:

- *Zeitgleiche Messung der Variablen eines Messzeitpunkts*: Die im CLP-Modell verwendeten Variablen sollten zu den jeweiligen Messzeitpunkten jeweils simultan erhoben werden.
- *Time-Lag*: Die im CLP-Modell verwendeten Variablen sollten in spezifischen, theoretisch begründbaren Messintervallen erhoben werden, sodass Aussagen über die untersuchten Zusammenhänge sich auf ebendiese Zeitspanne beziehen.
- *Konstante Merkmale vs. variable Merkmale*: Da im Rahmen von CLP-Modellen längsschnittliche Veränderungen abgebildet werden sollen, können konstante, weitgehend unveränderbare Merkmale (wie z.B. das Geschlecht) nicht modelliert werden, da für Konstanten keine Varianzaufklärung der Veränderung zu beobachten ist.
- *Reliabilität der Messung*: Die Zuverlässigkeit der Messung muss sowohl für die einzelnen Messzeitpunkte als auch im Zeitverlauf sowie im Gruppenvergleich gewährleistet sein, um sicherzustellen, dass es sich bei Veränderungen der Merkmale nicht um methodische Artefakte handelt. Um die Reliabilität

der Messung sicherzustellen, sollte die Messinvarianz aller verwendeten Konstrukte sowohl über die Zeit als auch zwischen den Vergleichsgruppen geprüft werden (vgl. Kapitel 8.4.4).

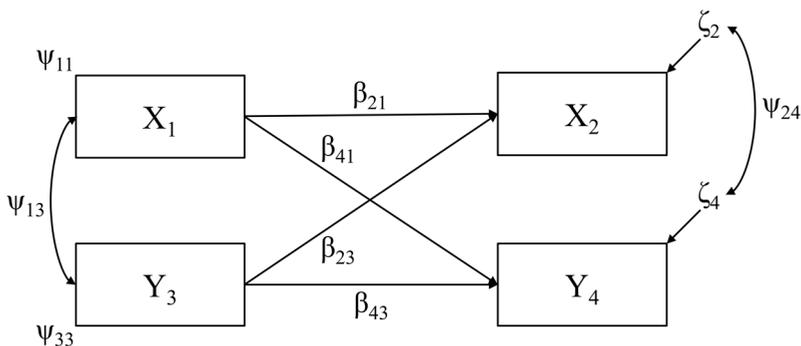
- *Drittvariablen-Effekte*: Da die untersuchten Zusammenhänge zwischen Merkmalen stets nur Ausschnitte aller möglichen theoretischen Zusammenhänge darstellen, ist nicht auszuschließen, dass weitere Variablen mit den analysierten Merkmalen korrelieren und gegebenenfalls sogar erklärungs mächtiger für die untersuchten Zusammenhänge sind als die betrachteten Variablen im CLP-Modell.

Die ersten vier Grundannahmen sind durch das Design der Studie ProBiNi gegeben beziehungsweise werden in den CLP-Modellen entsprechend modelliert. Eventuelle Drittvariablen-Effekte werden bei der Interpretation der untersuchten CLP-Modelle diskutiert.

Die Abbildung 8.03 veranschaulicht schematisch den Aufbau eines manifesten CLP-Modells von zwei Merkmalen (X und Y), die zu zwei Messzeitpunkten (X_1 und Y_3 zum ersten Messzeitpunkt sowie X_2 und Y_4 zum zweiten Messzeitpunkt) beobachtet wurden. Die Variable X_1 wurde wiederholt zu Messzeitpunkt zwei erhoben (X_2). Entsprechendes gilt für die Variable Y_4 , die eine wiederholte Messung der Variable Y_3 darstellt. Über die autoregressiven Pfade (Beziehung zwischen X_1 und X_2 sowie zwischen Y_3 und Y_4) wird die Stabilität eines Merkmals über die Zeit modelliert, da davon ausgegangen wird, dass die Messung eines Merkmals zum zweiten Messzeitpunkt durch die Ausprägung desselben Merkmals zum ersten Messzeitpunkt vorhergesagt werden kann (Kennzeichnung der autoregressiven Beziehungen β_{21} und β_{43}). Eine hohe Stabilität eines Merkmals über die Zeit (d.h. eine hohe Autokorrelation) geht mit einer geringen Veränderungsvarianz des Merkmals einher und bedingt daher, dass auch andere Merkmale nur in geringem Maße zur Aufklärung der Varianz des Merkmals beitragen können (Reinders, 2006). Die Kernelemente der CLP-Modelle stellen die kreuzversetzten Beziehungen zwischen den verschiedenen Merkmalen (X und Y) zu unterschiedlichen Messzeitpunkten dar und geben Auskunft über die Längsschnittbeziehung zwischen den verschiedenen Merkmalen. Die Zusammenhänge zwischen Y_4 und X_1 und sowie zwischen X_2 und Y_3 werden als Cross-Lagged-Beziehungen (oder Kreuzpfade) bezeichnet (Kennzeichnung β_{41} und β_{23}). Anhand einzelner signifikanter Kreuzpfade kann eine inhaltliche Interpretation der längsschnittlichen Beziehung zwischen den jeweiligen Merkmalen herausgestellt werden. Wäre beispielsweise der Pfad β_{41} signifikant, der Pfad β_{23} aber nicht signifikant, so ließe sich daraus ableiten, dass Y_4 mit X_1 assoziiert ist, X_2 und Y_3 jedoch in keinem statistisch bedeutsamen Zusammenhang stehen. Ist hingegen keiner der modellierten Cross-Lagged-Effekte signifikant (oder sind alle Kreuzpfade signifikant), so können inhaltliche Aussagen zur längsschnittlichen Zusammenhangsstruktur anhand der internen Relation der Cross-Lagged-Beziehun-

gen abgeleitet werden (bspw. β_{41} fällt substantiell höher aus als β_{23}), sofern entsprechend standardisierte Werte zur Interpretation genutzt werden (Newsom, 2015; Reinders, 2006). Neben den autoregressiven und den Cross-Lagged-Beziehungen wird zwischen den Merkmalen eines Messzeitpunkts (X_1 und Y_3 sowie X_2 und Y_4) ein wechselseitiger Zusammenhang kontrolliert (ψ_{13}). ψ_{13} ist demnach die Kovarianz zwischen den zwei unabhängigen Variablen X_1 und Y_3 . ψ_{24} gibt die Kovarianz der Messfehler ζ_2 und ζ_4 der abhängigen Variablen X_2 und Y_4 an (Newsom, 2015).

Abbildung 8.03: Bivariates manifestes Cross-Lagged-Panel-Modell mit zwei Messzeitpunkten



Newsom, 2015, Abbildung 5.1, modifiziert

Die Beziehungen zwischen X_1 , X_2 , Y_3 und Y_4 können mittels folgender Gleichungen beschrieben werden, die zwei Regressionsmodellen entsprechen (Newsom, 2015):

$$X_2 = \beta_{21} X_1 + \beta_{23} Y_3 + \zeta_2 \quad (8.10)$$

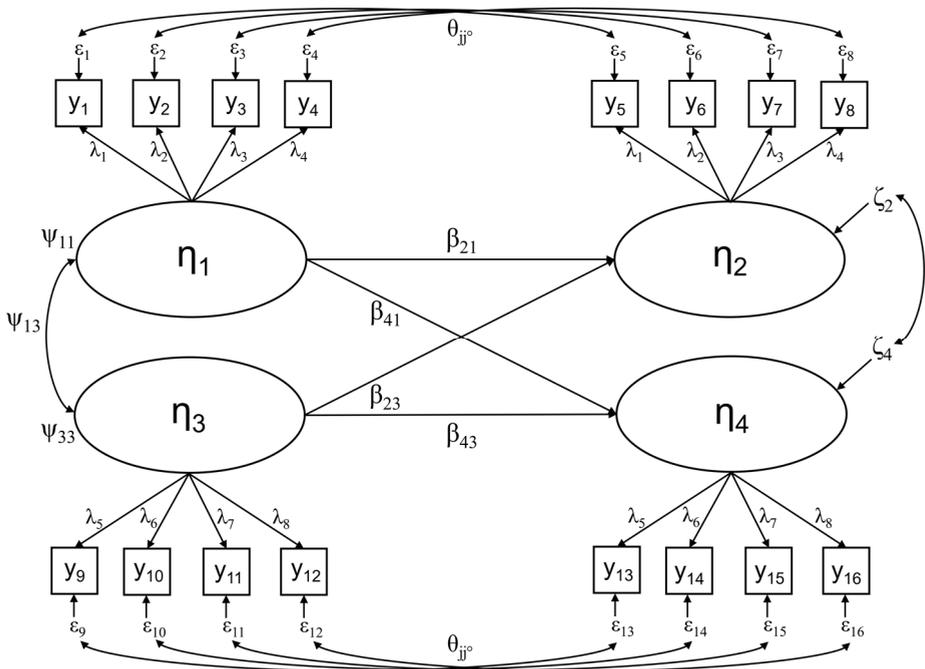
$$Y_4 = \beta_{43} Y_3 + \beta_{41} X_1 + \zeta_4 \quad (8.11)$$

CLP-Modelle können sowohl manifest als auch latent modelliert werden. Werden latente Variablen mit mehreren manifesten Indikatorvariablen verwendet, können Messfehler entsprechend berücksichtigt und somit die autoregressiven und Cross-Lagged-Beziehungen genauer geschätzt werden. Mittels der latenten Modellierung kann außerdem die Messinvarianz der Konstrukte im Längsschnitt sowie im Multi-gruppenvergleich kontrolliert werden (s. Kapitel 8.4.4). Latente CLP-Modelle erlauben es darüber hinaus, mögliche Messfehlerverzerrungen in der Schätzung der autoregressiven und Cross-Lagged-Beziehungen zu vermeiden, in dem die Korrelation der Messfehler eines Indikators über die verschiedenen Messzeitpunkte modelliert wird (Christ & Schlüter, 2012; Newsom, 2015). Die Nichtbeachtung der Auto-

korrelation der Indikatorresiduen kann zu einer Verzerrung der Stabilitätsannahmen eines Konstrukts führen sowie schlechtere Werte in den Modellgütekriterien bedingen und damit gegebenenfalls zur Ablehnung des jeweiligen Modells führen (Christ & Schlüter, 2012).

Die Abbildung 8.04 zeigt den schematischen Aufbau eines latent modellierten CLP-Modells mit zwei latenten Faktoren (η_1 und η_3), die zum zweiten Messzeitpunkt wiederholt (η_2 und η_4) und jeweils anhand von vier manifesten Indikatoren (y_x) gemessen werden.

Abbildung 8.04: Bivariates latentes Cross-Lagged-Panel-Modell mit zwei Messzeitpunkten



Newsom, 2015, Abbildung 5.2, modifiziert

Anmerkung: η_2 stellt eine wiederholte Messung desselben Konstrukts η_1 zum zweiten Messzeitpunkts dar und η_4 stellt eine wiederholte Messung desselben Konstrukts η_3 zum zweiten Messzeitpunkts dar.

Dabei stehen η_2 und η_4 für die zwei latenten Variablen des zweiten Messzeitpunkts, die jeweils über die autoregressiven Pfade β_{21} und β_{43} mit den latenten Variablen des jeweils gleichen Konstrukts des ersten Messzeitpunkts (η_1 und η_3) verbunden sind. β_{23} und β_{41} stehen für die Cross-Lagged-Beziehungen zwischen den latenten Variablen η_2 und η_3 beziehungsweise η_4 und η_1 .

Die autoregressiven und Cross-Lagged-Beziehungen zwischen den latenten Variablen können anhand folgender Strukturgleichungen beschrieben werden:

$$\eta_2 = \beta_{21} \eta_1 + \beta_{23} \eta_3 + \zeta_2 \quad (8.12)$$

$$\eta_4 = \beta_{43} \eta_3 + \beta_{41} \eta_1 + \zeta_4 \quad (8.13)$$

Für die latenten Variablen eines Konstrukts (η_1 und η_2 bzw. η_3 und η_4) werden invariante Faktorladungen zwischen den beiden Messzeitpunkten angenommen (metrische Messinvarianz; s. Bezeichnung der λ_x in Abbildung 8.04; Newsom, 2015). Zusätzlich können für die Prüfung skalarer Messinvarianz die Schwellenwerte (unter Berücksichtigung der ordinalen Datenstruktur; ansonsten der Intercepts) eines Konstrukts im Längsschnitt (und gegebenenfalls Gruppenvergleich) fixiert werden. Ebenfalls modelliert werden die Autokorrelationen der Indikatorresiduen (ε_x) eines Konstrukts über die Zeit, um Verzerrungen in den Stabilitätsannahmen zu vermeiden (Newsom, 2015).

8.4.6 Software und analytisches Vorgehen

Für die Datenaufbereitung (Bildung der Analytestichprobe (s. Kapitel 8.1.1), Rekodierung der Variablen sowie Erzeugung des .dat-files zum Einlesen der Daten in *Mplus 7.11*; Muthén & Muthén, 2012) wurde die Software *SPSS 27* (IBM Corp, 2021) genutzt. Darüber hinaus wurden die deskriptiven Analysen zur Beschreibung der Gesamt- und Analytestichprobe (s. Kapitel 8.1.1) sowie die Berechnung der Itemkennwerte zur Vorstellung der Erhebungsinstrumente (s. Kapitel 8.1.2) in *SPSS 27* durchgeführt. Ebenfalls in *SPSS 27* analysiert wurden die Mittelwertvergleiche zur Darstellung der deskriptiven Befunde des durchschnittlichen domänenspezifischen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts zu den drei Messzeitpunkten sowie zwischen den Profilklassen und deren Parallelklassen ohne Profil (s. Kapitel 9.2).

Die Strukturgleichungsmodellierung wurde in *Mplus 7.11* (Muthén & Muthén, 2012) umgesetzt. Zunächst wurde die Konstruktvalidität und die Messinvarianz der für diese Arbeit zentralen Erhebungsinstrumente im Längsschnitt sowie im Mehrgruppenvergleich im Rahmen von konfirmatorischen Faktorenanalysen untersucht (s. Kapitel 9.1). Daran anknüpfend wurde die längsschnittliche Beziehungsstruktur zwischen dem Interesse (Fach- bzw. Sachinteresse) und dem domänenspezifischen Fähigkeitsselbstkonzept anhand von latenten Cross-Lagged-Panel-Modellen (CLP) umgesetzt (s. Kapitel 9.3).

Für die konfirmatorischen Faktorenanalysen, Messinvarianzanalysen im Längsschnitt und im Multigruppenvergleich sowie die latenten Cross-Lagged-Panel-Modelle wurde der für ordinale Daten besonders robuste WLSMV-Schätzer (Li, 2016a) verwendet und die Mehrebenenstruktur der Daten mittels des in *Mplus 7.11* (Muthén & Muthén, 2012) implementierten Verfahrens für geclusterte Stichproben (*type = complex*) berücksichtigt.

9 Empirische Analysen

Im folgenden Kapitel werden die Ergebnisse der empirischen Analysen in der Reihenfolge der in Kapitel 7 formulierten Forschungsfragen vorgestellt. Zunächst werden die sechs zentralen Konstrukte des naturwissenschaftlichen sowie musikalischen Fach- und Sachinteresses sowie des fachspezifischen Fähigkeitsselbstkonzepts hinsichtlich ihrer Faktorenstruktur geprüft (Kapitel 9.1.1) und anschließend mittels Messinvarianzanalysen die Vergleichbarkeit der Konstrukte im Längsschnitt sowie im Multigruppenvergleich bestätigt (Kapitel 9.1.2 und 9.1.3). Unter Annahme entsprechender Messinvarianz können anschließend deskriptive Befunde zur Entwicklung des durchschnittlichen domänenspezifischen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts präsentiert und dabei Unterschiede zwischen den MINT- und Musik-Klassen sowie deren Parallelklassen ohne Profil beleuchtet werden (Kapitel 9.2). Die Ergebnisse der latenten Cross-Lagged-Panel-Modelle geben Aufschluss über die längsschnittliche Stabilität sowie die Zusammenhänge des Fach- beziehungsweise Sachinteresses und des Fähigkeitsselbstkonzepts. Die Modelle werden für die beiden Domänen Musik und Naturwissenschaften getrennt spezifiziert und zunächst für die Gesamtstichprobe berichtet (Kapitel 9.3.1 und 9.3.2). Durch die anschließende getrennte Modellierung für die Profilklassen und die Klassen ohne Profil kann der Effekt des jeweiligen Klassenprofils auf die Längsschnittbeziehung der Konstrukte herausgestellt werden (Kapitel 9.3.3 und 9.3.4).

9.1 Die Konstrukte: Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept

9.1.1 Überprüfung der Faktorenstruktur

Bezugnehmend auf die erste Forschungsfrage wird die Konstruktvalidität der zentralen Messinstrumente zur Erfassung des domänenspezifischen Fach- und Sachinteresses sowie des domänenspezifischen Fähigkeitsselbstkonzepts mittels konfirmatorischer Faktorenanalysen unter Verwendung des robusten WLSMV-Schätzers (Li, 2016a) sowie unter Berücksichtigung der Mehrebenenstruktur in *Mplus 7.11* (Muthén & Muthén, 2012) statistisch geprüft. Getrennt nach den Domänen Musik und Naturwissenschaften wird pro Messzeitpunkt ein latenter Faktor des jeweiligen Konstrukts (Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept) spezifiziert. Die drei Faktorenmodelle eines domänenspezifischen Konstrukts werden simultan geschätzt und die Modellgüte der Messmodelle anhand der Fit-Indizes CFI, TLI und RMSEA beurteilt.

Tabelle 9.01: Model-Fit der konfirmatorischen Faktorenanalysen der zentralen Konstrukte in den Domänen Musik und Naturwissenschaften

	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>
<i>Fachinteresse</i>			
Musik	.993	.990	.049
Naturwissenschaften	.994	.992	.045
<i>Sachinteresse</i>			
Musik	.996	.994	.033
Naturwissenschaften	.996	.995	.027
<i>Fähigkeitsselbstkonzept</i>			
Musik	.982	.980	.041
Naturwissenschaften	.976	.974	.037

In den Messmodellen konnten $n \geq 441$ Fälle berücksichtigt werden.

Wie aus der Tabelle 9.01 abgelesen werden kann, zeigen die Fit-Indizes für alle zentralen Konstrukte eine sehr gute Modell-Daten-Passung. Die inkrementellen Fit-Werte CFI und TLI liegen für die betrachteten Konstrukte oberhalb der von Hu und Bentler (1999) sowie Little (2013) empfohlenen *Cutoff*-Grenze von $CFI/TLI \geq .95$, wobei die TLI-Werte erwartungsgemäß etwas niedriger sind als die CFI-Werte, da im TLI die Komplexität des Modells über die Freiheitsgrade berücksichtigt wird. Die RMSEA-Werte liegen für alle Konstrukte unterhalb der *Cutoff*-Grenze von $RMSEA \leq .05$ (vgl. Wang & Wang, 2020). Somit können die betrachteten

Messmodelle der sechs Konstrukte (domänenspezifisches Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept) angenommen und für weitere Analysen verwendet werden.

Zusammenfassend kann für die erste Forschungsfrage festgehalten werden, dass die zentralen Konstrukte des domänenspezifischen Fach- beziehungsweise Sachinteresses sowie des Fähigkeitsselbstkonzepts eine angemessene Konstruktvalidität aufweisen. Ob sich die Konstrukte auch im Längsschnitt sowie im Gruppenvergleich als messinvariant erweisen, wird im Folgenden geprüft und damit die Grundlage für weiterführende Analysen gelegt. Zunächst werden die Ergebnisse der Messinvarianztestung im Längsschnitt berichtet (Kapitel 9.1.2) und anschließend die Ergebnisse des Mehrgruppenvergleichs dargestellt (Kapitel 9.1.3).

9.1.2 Messinvarianz im Längsschnitt

Um Entwicklungsverläufe abbilden zu können, wird eine Vergleichbarkeit der verwendeten Konstrukte über die Zeit vorausgesetzt (Liu et al., 2017). Dazu sollte bei der Prüfung auf Messinvarianz mindestens metrische, für den Vergleich von Mittelwerten jedoch skalare, Messinvarianz vorliegen (Brown, 2015; Christ & Schlüter, 2012). Über die verschiedenen Messzeitpunkte könnte folglich von invarianten Faktorladungen (metrische Messinvarianz) sowie ergänzend invarianten Schwellenwerten (skalare Messinvarianz) ausgegangen werden (vgl. Kapitel 8.4.4).

Zur Beantwortung der zweiten Forschungsfrage wird für die Prüfung der Messinvarianz über die drei Messzeitpunkte hinweg getrennt nach den Domänen Musik und Naturwissenschaften pro Konstrukt und pro Messzeitpunkt je ein latenter Faktor spezifiziert. In Anlehnung an das von Liu et al. (2017) empfohlene Vorgehen werden entsprechende Messinvarianzmodelle mit ordinal-kategorialen Indikatoren spezifiziert und der für ordinale Daten besonders robuste WLSMV-Schätzer (Li, 2016a) verwendet. Im Basismodell sind die Faktorladungen, Schwellenwerte und Residualvarianzen der manifesten Variablen über die drei Messzeitpunkte unrestrictiert. Anschließend werden zunehmend Gleichheitsrestriktionen in die Modellierung aufgenommen, um die weiteren Stufen der Messinvarianz im Längsschnitt zu testen. Für den Vergleich der unterschiedlich restriktiven Messinvarianzmodelle wird die in *Mplus 7.11* implementierte DIFFTEST-Option genutzt, um den χ^2 -Differenztest unter Verwendung des WLSMV-Schätzers adäquat zu berechnen (Asparouhov & Muthén, 2006). Die Mehrebenenstruktur der Daten wird mittels des in *Mplus 7.11* implementierten Verfahrens für geclusterte Stichproben (*type = complex*) berücksichtigt. Zudem wird auf die THETA-Parametrisierung zurückgegriffen, um die Residualvarianzen der Parameter im vierten Schritt der Messinvarianzprüfung (strikte Invarianz) zwischen den Messzeitpunkten fixieren zu können. Außerdem wird die Korrelation zwischen den Messfehlern eines Indikators über die drei Messzeitpunkte berücksichtigt (Autokorrelation der Indikatorresiduen).

Sowohl für das Konstrukt Fachinteresse in Musik als auch für das Konstrukt Fachinteresse in den Naturwissenschaften kann anhand der inkrementellen Fit-Werte (CFI, TLI) sowie der absoluten Fit-Werte (χ^2 , RMSEA) mindestens skalare Messinvarianz bestätigt werden (s. Tabelle 9.02).

Tabelle 9.02: Model-Fit der Messinvarianzanalysen (Längsschnitt) des Konstrukts Fachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)

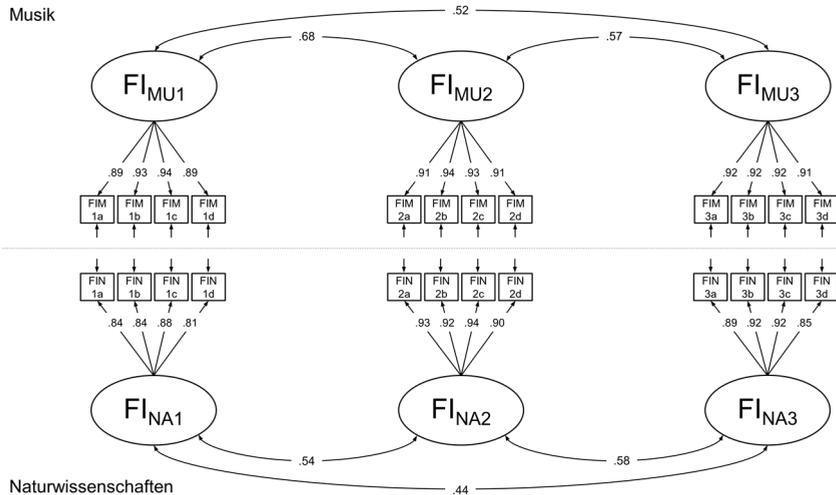
		χ^2	df	p	$\Delta\chi^2$	df	p	CFI	ΔCFI	TLI	RMSEA
Konfigurale Messinvarianz (Basismodell)	Musik	101.462	39	.000	-	-	-	.993	-	.988	.054
	Nawi	84.779	39	.000	-	-	-	.995	-	.991	.046
Metrische Messinvarianz (invariante λ)	Musik	112.866	45	.000	19.085	6	.004	.993	.000	.989	.052
	Nawi	90.349	45	.000	5.691	6	.459	.995	.000	.993	.043
Skalare Messinvarianz (invariante λ ; ν)	Musik	131.866	59	.000	34.404	20	.024	.992	.001	.991	.047
	Nawi	110.103	59	.000	31.148	20	.053	.994	.001	.994	.040
Strikte Messinvarianz (invariante λ ; ν ; δ)	Musik	139.617	67	.000	52.571	28	.003	.992	.000	.992	.044
	Nawi	121.881	67	.000	50.509	28	.006	.994	.000	.994	.039

λ = Faktorladungen; ν = Schwellenwerte; δ = Residualvarianzen

CFI und TLI liegen in allen getesteten Messinvarianzmodellen oberhalb der empfohlenen *Cutoff*-Werte von $CFI/TLI \geq .95$ (Little, 2013). Der ΔCFI , der die Differenz des Modells zum jeweils weniger restriktiven Modell angibt, sollte nach Chen (2007) sowie Cheung und Rensvold (2002) nicht größer als $\Delta CFI = .010$ sein, um Messinvarianz anzuzeigen. Auch diese Annahme kann auf Basis der Modellvergleiche bestätigt werden. Der χ^2 -Differenzentest wurde mittels der in *Mplus 7.11* implementierten DIFFTEST-Option für den WLSMV-Schätzer entsprechend adäquat berechnet und unterstreicht zusätzlich die Messinvarianzannahme des Fachinteresses in Musik und den Naturwissenschaften über die drei Messzeitpunkte hinweg. Die Modelle skalarer Messinvarianz unterscheiden sich folglich nicht signifikant von den jeweils weniger restriktiven Modellen.

Das Gesamtmodell skalarer Messinvarianz, das heißt unter Annahme gleicher Faktorladungen und Schwellenwerte über die drei Messzeitpunkte, wird für das musikalische und naturwissenschaftliche Fachinteresse in der folgenden Abbildung 9.01 veranschaulicht. Wie in Abbildung 9.01 deutlich wird, zeigen sich für die Konstrukte des musikalischen und naturwissenschaftlichen Fachinteresses unter Annahme skalarer Messinvarianz über die drei Messzeitpunkte hinweg hohe Faktorladungen zwischen den vier manifesten Items und den pro Messzeitpunkt spezifizierten Faktoren. Darüber hinaus weisen die Konstrukte im Längsschnitt mittlere Korrelationen auf.

Abbildung 9.01: Konfirmatorische Faktorenanalyse unter Annahme skalarer Messinvarianz im Längsschnitt des Konstrukts Fachinteresse (FI) in den Domänen Musik (MU) und Naturwissenschaften (NA)



Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. Alle Werte sind signifikant ($p \leq .01$); autokorrierte Residuen.

Fachinteresse Musik: CFI = .992; TLI = .991; RMSEA = .047

Fachinteresse Naturwissenschaften: CFI = .994; TLI = .994; RMSEA = .040

Für das Konstrukt Sachinteresse werden ebenfalls getrennt nach den Domänen Musik und Naturwissenschaften schrittweise restringierte Messinvarianzmodelle getestet. Anhand der Modellgütekriterien kann ebenfalls ein sehr guter Modell-Fit bestätigt werden und zwar sowohl für das Basismodell als auch für die schrittweise restringierten Modelle mit fixierten Parametern (s. Tabelle 9.03).

Die RMSEA-Werte liegen in allen Modellen deutlich unter $RMSEA \leq .05$ und CFI sowie TLI in den Modellen metrischer und skalarer Messinvarianz sowohl für das musikalische als auch naturwissenschaftliche Sachinteresse nahe einem Wert von 1, was für eine nahezu perfekte Modell-Daten-Passung spricht und sich auch in den Modellvergleichen in Werten von $\Delta CFI \leq .010$ deutlich zeigt (vgl. Chen, 2007; Cheung & Rensvold, 2002). Einzig die Ergebnisse des χ^2 -Differenzentests könnten darauf hinweisen, dass sich die Modelle signifikant voneinander unterscheiden. Vor dem Hintergrund der in Kapitel 8.4.2 beschriebenen Einschränkungen der χ^2 -Teststatistik zur Beurteilung der Modellgüte (u.a. Stichprobensensitivität) sollten diese Ergebnisse aber nicht überbewertet und stattdessen ebenfalls auf die übrigen Fit-Indizes zurückgegriffen werden, um Fehlinterpretationen in der Messinvarianztestung zu vermeiden (vgl. Wang & Wang, 2020). Ausgehend von diesen

Befunden kann für das Konstrukt Sachinteresse in den Domänen Musik und Naturwissenschaften angenommen werden, dass Messäquivalenz über die drei Messzeitpunkte hinweg vorliegt und somit latente Mittelwertvergleiche zulässig sind.

Tabelle 9.03: Model-Fit der Messinvarianzanalysen (Längsschnitt) des Konstrukts Sachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)

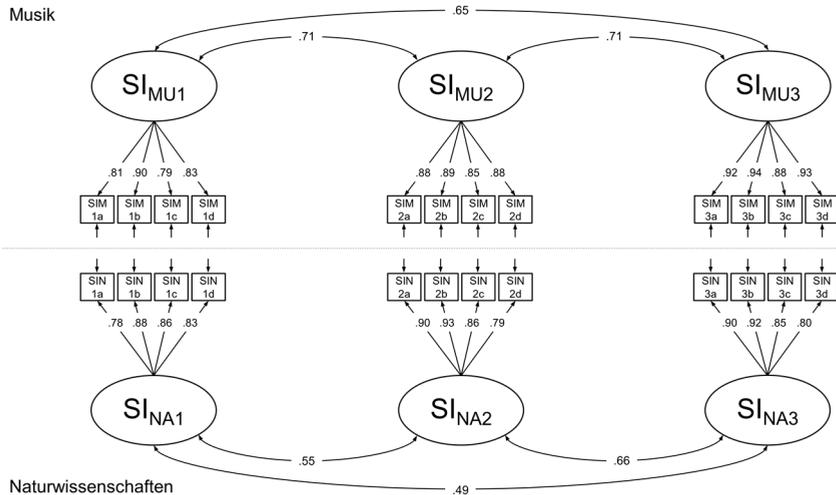
		χ^2	df	p	$\Delta\chi^2$	df	p	CFI	ΔCFI	TLI	RMSEA
Konfigurale Messinvarianz (Basismodell)	Musik	54.024	39	.055	-	-	-	.998	-	.996	.026
	Nawi	46.131	39	.201	-	-	-	.999	-	.998	.018
Metrische Messinvarianz (invariante λ)	Musik	60.780	45	.058	7.780	6	.255	.998	.000	.997	.025
	Nawi	58.008	45	.092	26.385	6	.000	.998	.001	.997	.023
Skalare Messinvarianz (invariante λ ; ν)	Musik	86.854	59	.011	51.315	20	.000	.996	.002	.995	.029
	Nawi	71.527	59	.127	30.769	20	.058	.998	.000	.998	.020
Strikte Messinvarianz (invariante λ ; ν ; δ)	Musik	93.914	67	.017	46.515	28	.015	.996	.000	.996	.027
	Nawi	115.593	67	.000	87.271	28	.000	.991	.007	.992	.036

λ = Faktorladungen; ν = Schwellenwerte; δ = Residualvarianzen

In Abbildung 9.02 ist das Gesamtmodell skalarer Messinvarianz des Konstrukts Sachinteresse dargestellt. Dabei wird das musikalische Sachinteresse dem naturwissenschaftlichen Sachinteresse gegenübergestellt.

Wie in Abbildung 9.02 deutlich wird, sind auch für das Konstrukt Sachinteresse unter Annahme skalarer Messinvarianz hohe Faktorladungen zwischen den manifesten und latenten Variablen zu beobachten, die im direkten Vergleich zum Faktorenmodell des domänenspezifischen Fachinteresses jedoch etwas niedriger ausfallen (vgl. Abbildung 9.01). Zwischen den Faktoren des musikalischen Sachinteresses zeigen sich im Längsschnitt hohe signifikante Korrelationen, zwischen den Faktoren des naturwissenschaftlichen Sachinteresses hingegen etwas niedrigere Zusammenhänge.

Abbildung 9.02: Konfirmatorische Faktorenanalyse unter Annahme skalarer Messinvarianz im Längsschnitt des Konstrukts Sachinteresse (SI) in den Domänen Musik (MU) und Naturwissenschaften (NA)



Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. Alle Werte sind signifikant ($p \leq .01$); autokorrelierte Residuen.

Sachinteresse Musik: CFI = .996; TLI = .995; RMSEA = .029

Sachinteresse Naturwissenschaften: CFI = .998; TLI = .998; RMSEA = .020

Auch das Konstrukt des musikalischen beziehungsweise naturwissenschaftlichen Fähigkeitsselbstkonzepts wird hinsichtlich der messäquivalenten Erfassung zu den drei Messzeitpunkten geprüft, in dem schrittweise in der Modellierung der Messmodelle verschiedene Gleichheitsrestriktionen angenommen werden. Die Ergebnisse der Messinvarianzprüfung des Konstrukts des Fähigkeitsselbstkonzepts sind in der folgenden Tabelle 9.04 dargestellt.

Sowohl für das musikalische als auch für das naturwissenschaftliche Fähigkeitsselbstkonzept kann auf Basis der inkrementellen und absoluten Fit-Indizes mindestens skalare Messinvarianz angenommen werden. Im Modellvergleich liegt die Differenz der CFI-Werte deutlich unter der empfohlenen Grenze von $\Delta CFI \leq .010$ (Chen, 2007; Cheung & Rensvold, 2002), sodass sich die Modelle mit zunehmenden Gleichheitsrestriktionen hinsichtlich ihrer Modellgüte nicht wesentlich von den jeweils weniger restriktiven Modellen unterscheiden. Der χ^2 -Differenzentest wird signifikant, sofern die Faktorladungen und die Schwellenwerte (skalare Messinvarianz) sowie Residualvarianzen (strikte Messinvarianz) über die drei Messzeitpunkte fixiert werden. Wie in Kapitel 8.4.2 beschrieben, kann es unter Verwendung des WLSMV-Schätzers im Rahmen der χ^2 -Teststatistik vermehrt zu Fehlern 1. Art kommen und damit insbesondere in kleineren Stichproben tendenziell eine Ablehnung des Mo-

dells bedingt werden (Beauducel & Herzberg, 2006; Li, 2016a). Unter Berücksichtigung der übrigen Modellgütekriterien (CFI, Δ CFI, TLI und RMSEA) wird daher angenommen, dass auch das Konstrukt des domänenspezifischen Fähigkeitsselbstkonzepts im Längsschnitt messinvariant und daher vergleichbar ist.

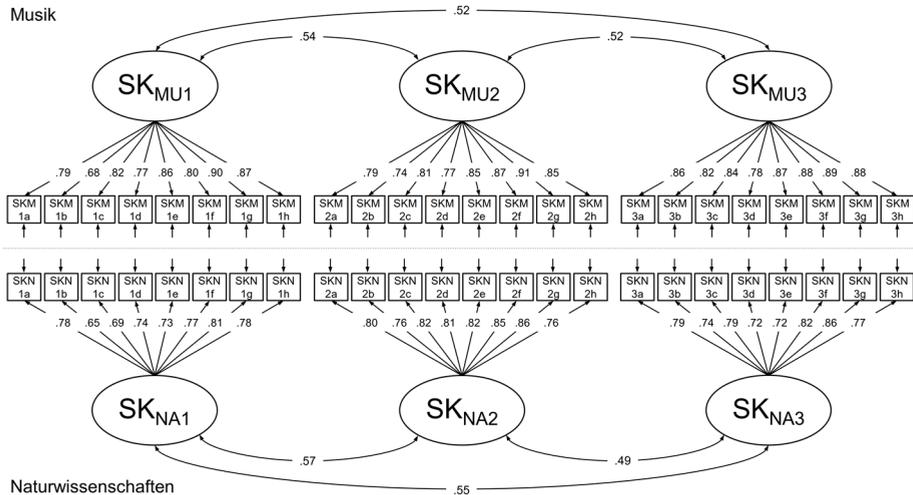
Tabelle 9.04: Model-Fit der Messinvarianzanalysen (Längsschnitt) des Konstrukts Fähigkeitsselbstkonzept in den Domänen Musik und Naturwissenschaften (Nawi)

		χ^2	df	p	$\Delta\chi^2$	df	p	CFI	Δ CFI	TLI	RMSEA
Konfigurale Messinvarianz (Basismodell)	Musik	433.387	225	.000	-	-	-	.983	-	.979	.041
	Nawi	383.121	225	.000	-	-	-	.980	-	.975	.036
Metrische Messinvarianz (invariante λ)	Musik	446.177	239	.000	16.407	14	.289	.983	.000	.981	.040
	Nawi	390.741	239	.000	17.843	14	.214	.980	.000	.977	.034
Skalare Messinvarianz (invariante λ ; ν)	Musik	480.800	269	.000	76.051	44	.002	.983	.000	.982	.038
	Nawi	445.538	269	.000	107.341	44	.000	.977	.003	.977	.035
Strikte Messinvarianz (invariante λ ; ν ; δ)	Musik	492.391	285	.000	111.992	60	.000	.983	.000	.984	.037
	Nawi	455.722	285	.000	121.208	60	.000	.978	.001	.979	.033

λ = Faktorladungen; ν = Schwellenwerte; δ = Residualvarianzen

Abbildung 9.03 veranschaulicht das längsschnittliche Gesamtmodell skalarer Messinvarianz für das Konstrukt des musikalischen und naturwissenschaftlichen Fähigkeitsselbstkonzepts. Zwischen den über die drei Messzeitpunkte hinweg analog formulierten manifesten Items und den drei spezifizierten Faktoren zeigen sich ebenfalls hohe Faktorladungen für das Konstrukt des domänenspezifischen Fähigkeitsselbstkonzepts. Die Korrelationen zwischen den Faktoren liegen sowohl für das musikalische als auch naturwissenschaftliche Fähigkeitsselbstkonzept im mittleren Bereich, sind aber durchweg signifikant.

Abbildung 9.03: Konfirmatorische Faktorenanalyse unter Annahme skalarer Messinvarianz im Längsschnitt des Konstrukts Fähigkeitsselbstkonzept (SK) in den Domänen Musik (MU) und Naturwissenschaften (NA)



Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. Alle Werte sind signifikant ($p \leq .01$); autokorrelierte Residuen.

Fähigkeitsselbstkonzept Musik: CFI = .983; TLI = .982; RMSEA = .038

Fähigkeitsselbstkonzept Naturwissenschaften: CFI = .977; TLI = .977; RMSEA = .035

9.1.3 Messinvarianz im Gruppenvergleich

Fragen der Vergleichbarkeit und damit der Messinvarianz sind neben längsschnittlichen Analysen auch insbesondere dann von Bedeutung, wenn Aussagen über verschiedene Gruppen getroffen werden sollen, die sich hinsichtlich ihrer Kontextbedingungen unterscheiden. Als eine solche Kontextbedingung kann auch die Gruppenzugehörigkeit zu einer Profilkategorie mit musikalischem oder naturwissenschaftlichem Schwerpunkt beziehungsweise einer Klasse ohne ein solches Profil angesehen werden. Um zu gewährleisten, dass es sich bei den Beobachtungen nicht um systematische Unterschiede handelt, die durch die reine Gruppenzugehörigkeit bedingt sind, werden die zentralen Konstrukte (domänenspezifisches Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept) im Folgenden hinsichtlich ihrer Vergleichbarkeit zwischen den verschiedenen Gruppen geprüft (s. Forschungsfrage 3). In Anlehnung an das von Millsap und Yun-Tein (2004) empfohlene Vorgehen werden zunehmend restriktive Messinvarianzmodelle im multiplen Gruppenvergleich spezifiziert und die Modelle anhand der Änderungen in den Modellgütekriterien miteinander verglichen (s. Kapitel 8.4.4). Dazu wird der für ordinale Daten robuste WLSMV-Schätzer (Li, 2016a), die THETA-Parametrisierung sowie die in *Mplus*

7.11 implementierte DIFFTEST-Option (Asparouhov & Muthén, 2006) genutzt. Die Mehrebenenstruktur der Daten wird mittels des in *Mplus 7.11* implementierten Verfahrens für geclusterte Stichproben (*type = complex*) berücksichtigt. Im Folgenden werden die Messinvarianzanalysen des multiplen Gruppenvergleichs für die sechs zentralen Konstrukte vorgestellt.

Die erste Analyse prüft, inwiefern das Konstrukt des domänenspezifischen Fachinteresses messinvariant zwischen den Musik- beziehungsweise MINT-Klassen sowie Klassen ohne Profil ist (s. Tabelle 9.05). Dazu wird, getrennt nach den Domänen Musik und Naturwissenschaften, pro Messzeitpunkt ein latenter Faktor spezifiziert und die Messmodelle anschließend unter Berücksichtigung verschiedener Restriktionen zwischen den Gruppen verglichen.

Die Fit-Indizes des Basismodells sprechen für eine sehr gute Modell-Daten-Passung. Sowohl für das musikalische als auch für das naturwissenschaftliche Fachinteresse liegen die Fit-Indizes CFI und TLI über den empfohlenen *Cutoff*-Grenzen $CFI/TLI \geq .95$ beziehungsweise der RMSEA unter dem empfohlenen Wert $RMSEA \leq .05$ (vgl. Little, 2013). Auch durch die Fixierung der Faktorladungen sowie der Schwellenwerte zwischen den Gruppen kann keine Verschlechterung in der Modellgüte beobachtet werden: Die Differenz der CFI-Werte beträgt zwischen den Modellen lediglich maximal $\Delta CFI = .002$.

Tabelle 9.05: Model-Fit der Messinvarianzanalysen (Gruppenvergleich) des Konstrukts Fachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)

		χ^2	<i>df</i>	<i>p</i>	$\Delta\chi^2$	<i>df</i>	<i>p</i>	<i>CFI</i>	ΔCFI	<i>TLI</i>	<i>RMSEA</i>
Konfigurale Messinvarianz (Basismodell)											
	Musik	184.751	111	.000	-	-	-	.993	-	.991	.049
	Nawi	173.817	111	.000	-	-	-	.993	-	.991	.045
Metrische Messinvarianz (invariante λ)											
	Musik	187.707	114	.000	4.902	3	.179	.993	.000	.992	.048
	Nawi	174.905	114	.000	2.147	3	.542	.993	.000	.992	.044
Skalare Messinvarianz (invariante λ ; ν)											
	Musik	206.512	132	.000	23.888	21	.299	.993	.000	.993	.045
	Nawi	188.982	132	.001	13.899	21	.874	.994	.001	.994	.040
Strikte Messinvarianz (invariante λ ; ν ; δ)											
	Musik	214.727	144	.000	40.887	33	.163	.993	.000	.994	.042
	Nawi	216.910	144	.000	61.292	33	.002	.992	.002	.992	.043

λ = Faktorladungen; ν = Schwellenwerte; δ = Residualvarianzen

Der χ^2 -Differenzentest, der mittels der in *Mplus 7.11* implementierten DIFFTEST-Option für den WLSMV-Schätzer entsprechend adäquat berechnet wurde, ist sowohl in den Modellen der metrischen Messinvarianz als auch der skalaren Messinvarianz nicht signifikant, sodass sich auch nach der Fixierung der Faktorladungen und der Schwellenwerte die Modelle nicht signifikant unterscheiden. Für das Konstrukt Fachinteresse kann somit mindestens skalare Messinvarianz zwischen den Gruppen bestätigt werden.

In einem zweiten Schritt wird getestet, inwiefern das domänenspezifische Sachinteresse zu den drei Messzeitpunkten messäquivalent in den Musik- beziehungsweise MINT-Profilklassen sowie in den Klassen ohne Profil erfasst werden kann. Wie die folgende Tabelle 9.06 zeigt, liegen die Fit-Indizes für die Modelle skalarer Invarianz im Rahmen der angestrebten *Cutoff*-Grenzen (s. Kapitel 8.4.2). Weder anhand des ΔCFI noch mittels des χ^2 -Differenzentest können signifikante Verschlechterungen des Modells invarianter Faktorladungen und Schwellenwerte gegenüber den weniger restringierten Modellen festgestellt werden, sodass skalare Messinvarianz im Gruppenvergleich für musikalische und naturwissenschaftliche Sachinteresse bestätigt werden kann.

Tabelle 9.06: Model-Fit der Messinvarianzanalysen (Gruppenvergleich) des Konstrukts Sachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)

		χ^2	<i>df</i>	<i>p</i>	$\Delta\chi^2$	<i>df</i>	<i>p</i>	<i>CFI</i>	ΔCFI	<i>TLI</i>	<i>RMSEA</i>
Konfigurale Messinvarianz (Basismodell)											
	Musik	144.797	111	.017	-	-	-	.996	-	.995	.033
	Nawi	160.309	111	.002	-	-	-	.992	-	.990	.040
Metrische Messinvarianz (invariante λ)											
	Musik	149.011	114	.015	11.771	3	.008	.995	.001	.995	.033
	Nawi	163.902	114	.002	6.785	3	.079	.992	.000	.990	.040
Skalare Messinvarianz (invariante λ ; ν)											
	Musik	169.727	132	.015	34.556	21	.032	.995	.000	.995	.032
	Nawi	176.174	132	.006	20.263	21	.505	.993	.001	.993	.035
Strikte Messinvarianz (invariante λ ; ν ; δ)											
	Musik	188.384	144	.008	56.189	33	.007	.994	.001	.995	.033
	Nawi	193.109	144	.004	45.649	33	.070	.992	.001	.993	.035

λ = Faktorladungen; ν = Schwellenwerte; δ = Residualvarianzen

In einem dritten Schritt wird für das musikalische und naturwissenschaftliche Fähigkeitsselbstkonzept geprüft, ob von einer messinvarianten Erfassung der Konstrukte in den Klassen mit Musik- oder MINT-Profil und in den Klassen ohne Profil zu den drei Messzeitpunkten ausgegangen werden kann (s. Tabelle 9.07). Die Modellgütekriterien sprechen auch nach der Fixierung der entsprechenden Parameter (Faktorladungen und Schwellenwerte) klar für eine äquivalente Messung des Fähigkeitsselbstkonzepts in den verschiedenen Gruppen. Für das Fähigkeitsselbstkonzept in den Naturwissenschaften sind die Werte des ΔCFI in den verschiedenen

Messinvarianzmodellen leicht höher als die entsprechenden ΔCFI -Werte des musikalischen Fähigkeitsselbstkonzepts, liegen aber dennoch weit unter der *Cutoff*-Grenze von $\Delta CFI \leq .010$. Zudem sprechen die nicht signifikanten Werte des χ^2 -Differenzentests dafür, dass sich die Modelle nicht wesentlich vom Basismodell beziehungsweise vom weniger restringierten Modell metrischer Messinvarianz unterscheiden und bestätigen somit mindestens skalare Messinvarianz für das Konstrukt des musikalischen und naturwissenschaftlichen Fähigkeitsselbstkonzepts.

Tabelle 9.07: Model-Fit der Messinvarianzanalysen (Gruppenvergleich) des Konstrukts Fähigkeitsselbstkonzept in den Domänen Musik und Naturwissenschaften (Nawi)

		χ^2	df	p	$\Delta\chi^2$	df	p	CFI	ΔCFI	TLI	RMSEA
Konfigurale Messinvarianz (Basismodell)	Musik	700.852	507	.000	-	-	-	.983	-	.982	.038
	Nawi	691.855	507	.000	-	-	-	.973	-	.971	.037
Metrische Messinvarianz (invariante λ)	Musik	716.134	522	.000	25.448	15	.044	.983	.000	.982	.037
	Nawi	698.699	522	.000	17.972	15	.264	.975	.002	.973	.035
Skalare Messinvarianz (invariante λ ; ν)	Musik	756.040	564	.000	75.953	57	.047	.983	.000	.984	.035
	Nawi	733.442	564	.000	60.759	57	.342	.976	.001	.976	.033
Strikte Messinvarianz (invariante λ ; ν ; δ)	Musik	774.223	587	.000	112.322	80	.010	.984	.001	.985	.034
	Nawi	751.791	587	.000	99.228	80	.072	.976	.000	.978	.032

λ = Faktorladungen; ν = Schwellenwerte; δ = Residualvarianzen

Zusammenfassend kann – bezugnehmend auf die zweite Forschungsfrage – folglich mindestens skalare Messinvarianz der sechs zentralen Konstrukte im Längsschnitt sowie – bezugnehmend auf die dritte Forschungsfrage – mindestens skalare Messinvarianz im Mehrgruppenvergleich bestätigt werden. Somit sind die Voraussetzungen für den Vergleich der Konstrukte zwischen der fünften und siebten Jahrgangsstufe sowie zwischen den verschiedenen Gruppen gegeben (vgl. Brown, 2015; Christ & Schlüter, 2012).

9.2 Deskriptive Befunde

Nachdem die für diese Arbeit zentralen Konstrukte des domänenspezifischen Fachbeziehungsweise Sachinteresses und des Fähigkeitsselbstkonzepts hinsichtlich ihrer faktoriellen Struktur geprüft sowie skalare Messinvarianz der Konstrukte sowohl im Längsschnitt als auch im multiplen Gruppenvergleich bestätigt wurde, können für die im Folgenden präsentierten deskriptiven Befunde Mittelwertvergleiche herangezogen werden. Diese ermöglichen einen ersten Einblick in die Entwicklung der

zentralen Konstrukte, die jeweils getrennt für die verschiedenen Subgruppen (Musik- bzw. MINT-Klassen sowie deren direkte Parallelklassen ohne Profil) berichtet werden, um zusätzlich Gruppenunterschiede herausstellen zu können. Mittels der im Folgenden vorgestellten Analysen wird die vierte Forschungsfrage untersucht und die entsprechenden *Hypothesen 4.1a* bis *4.6b* geprüft.

Zunächst wird die Entwicklung des durchschnittlichen musikalischen und naturwissenschaftlichen Fachinteresses zwischen der fünften und siebten Jahrgangsstufe (erster bis dritter Messzeitpunkt) in den verschiedenen Subgruppen näher beleuchtet (Kapitel 9.2.1), bevor anschließend auf die Entwicklung des mittleren domänenspezifischen Sachinteresses eingegangen wird (Kapitel 9.2.2). In einem dritten Schritt werden die deskriptiven Befunde zur Entwicklung des Fähigkeitsselbstkonzepts zu Beginn der Sekundarstufe I in den Domänen Musik und Naturwissenschaften präsentiert und dabei erneut Gruppenunterschiede zwischen Schülerinnen und Schülern in Profilklassen und Klassen ohne Profil herausgestellt (Kapitel 9.2.3).

9.2.1 Entwicklung des Fachinteresses in Profilklassen und Klassen ohne Profil

Tabelle 9.08 und Tabelle 9.09 stellen die deskriptiven Befunde des durchschnittlichen musikalischen beziehungsweise naturwissenschaftlichen Fachinteresses im Längsschnitt dar. Dabei werden die vier Subgruppen der Profilklassen mit musikalischem beziehungsweise naturwissenschaftlichem Schwerpunkt sowie deren direkte Parallelklassen ohne Profil (Vergleichsklassen) gegenübergestellt.

Tabelle 9.08: Deskriptive Befunde – Entwicklung des durchschnittlichen musikalischen Fachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen)

	1. MZP		2. MZP		3. MZP	
	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)
Musik-Klassen	3.6	(0.05)	3.1 ▶	(0.07)	3.0 ▶	(0.07)
Musik-Vergleichsklassen	▼ 2.7	(0.08)	▼ 2.2 ▶	(0.09)	▼ 2.6	(0.08)
MINT-Klassen	▼ 2.4	(0.08)	▼ 2.3	(0.09)	▼ 2.4	(0.09)
MINT-Vergleichsklassen	▼ 2.4	(0.08)	▼ 2.1 ▶	(0.08)	▼ 2.2 ▶	(0.09)

▶ Wert im Vergleich zur Referenzkategorie „1. MZP“ signifikant verschieden ($p \leq .05$).

▼ Wert im Vergleich zur Referenzkategorie „Musik-Klassen“ signifikant niedriger ($p \leq .05$).

Erwartungskonform ist das musikalische Fachinteresse in den Musik-Klassen besonders hoch ausgeprägt. Hervorzuheben ist der hohe Ausgangswert im musikalischen Fachinteresse der Schülerinnen und Schüler in den Musik-Klassen zu Beginn

der fünften Jahrgangsstufe (erster Messzeitpunkt), der die These untermauert, dass Musik-Profilklassen vor allem besonders interessierte Kinder adressieren (vgl. Kapitel 5.3). Zwischen dem ersten und zweiten Messzeitpunkt verringert sich das musikalische Fachinteresse in den Musik-Klassen zwar zunächst signifikant, stabilisiert sich jedoch zum dritten Messzeitpunkt auf einem hohen Niveau und liegt zu allen drei Messzeitpunkten deutlich über dem theoretischen Mittelwert von 2.5. *Hypothese 4.1a* kann demnach empirisch bekräftigt werden. Im Vergleich zu den übrigen Untersuchungsgruppen weisen die Schülerinnen und Schüler in den Musik-Klassen deutlich mehr musikalische Fachinteresse auf als ihre Mitschülerinnen und Mitschüler. Der Unterschied ist – hypothesenkonform – zu allen drei Messzeitpunkten zwischen den Musik-Klassen und den Musik-Vergleichsklassen ohne Profil aber auch im Vergleich zu den MINT-Klassen sowie zu den MINT-Vergleichsklassen signifikant (*Hypothese 4.1b*).

Tabelle 9.09: Deskriptive Befunde – Entwicklung des durchschnittlichen naturwissenschaftlichen Fachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen)

	1. MZP		2. MZP		3. MZP	
	M	(SE)	M	(SE)	M	(SE)
MINT-Klassen	3.0	(0.06)	3.1	(0.08)	2.5 ▶	(0.09)
MINT-Vergleichsklassen	▼ 2.7	(0.07)	2.9	(0.09)	2.3 ▶	(0.08)
Musik-Klassen	3.0	(0.06)	3.0	(0.07)	2.7 ▶	(0.07)
Musik-Vergleichsklassen	▼ 2.7	(0.07)	▼ 2.7	(0.09)	2.5 ▶	(0.08)

▶ Wert im Vergleich zur Referenzkategorie „1. MZP“ signifikant verschieden ($p \leq .05$).

▼ Wert im Vergleich zur Referenzkategorie „MINT-Klassen“ signifikant niedriger ($p \leq .05$).

Wie in Tabelle 9.09 ersichtlich, sind die Schülerinnen und Schüler in den MINT-Klassen erwartungsgemäß besonders am naturwissenschaftlichen Unterricht interessiert. Das naturwissenschaftliche Fachinteresse liegt zum ersten und zweiten Messzeitpunkt in den MINT-Profilklassen über dem theoretischen Mittelwert von 2.5 und ist somit als überdurchschnittlich zu bezeichnen. Zwischen der fünften und sechsten Jahrgangsstufe ist nominell sogar eine leichte Steigerung des naturwissenschaftlichen Fachinteresses in den MINT-Klassen feststellbar. Das überdurchschnittlich hohe Interesse am Naturwissenschaftsunterricht kann jedoch in den MINT-Klassen (aber auch in den übrigen Untersuchungsgruppen) nicht bis zur siebten Jahrgangsstufe aufrechterhalten werden, sodass in allen vier Subgruppen das durchschnittliche naturwissenschaftliche Fachinteresse zum dritten Messzeitpunkt signifikant niedriger ausfällt als zur Beginn der Sekundarstufe I (erster Messzeitpunkt). Zum dritten Messzeitpunkt ist daher in den MINT-Klassen lediglich ein

durchschnittliches naturwissenschaftliches Fachinteresse feststellbar, sodass die empirischen Befunde für eine Ablehnung der *Hypothese 4.2a* sprechen.

Verglichen mit den Schülerinnen und Schülern, die keine MINT-Klasse besuchen, fällt auf, dass das naturwissenschaftliche Fachinteresse der MINT-Klassen nominell häufig über dem naturwissenschaftlichen Fachinteresse der anderen Untersuchungsgruppen liegt. Allerdings sind die Unterschiede in den Ausprägungen des durchschnittlichen naturwissenschaftlichen Fachinteresses zwischen den MINT-Klassen und den übrigen Untersuchungsgruppen nur partiell signifikant. Besonders in den Musik-Klassen ist ein ebenfalls hohes naturwissenschaftliches Fachinteresse feststellbar, das zumindest zwischen der fünften und sechsten Jahrgangsstufe (erster und zweiter Messzeitpunkt) stabil auf einem hohen Niveau verbleibt. Der Unterschied zwischen den MINT-Klassen und den Musik-Klassen ist zu keinem Messzeitpunkt signifikant. Verglichen mit den MINT-Vergleichsklassen ohne Profil zeigen die Schülerinnen und Schüler in den MINT-Klassen nur zum ersten Messzeitpunkt signifikant mehr naturwissenschaftliches Fachinteresse als ihre Mitschülerinnen und Mitschüler. Zum zweiten und dritten Messzeitpunkt sind die Mittelwertunterschiede zwischen den MINT-Klassen und deren direkten Parallelklassen ohne Profil hingegen nicht signifikant. Für *Hypothese 4.2b* finden sich empirisch folglich nur bedingt Hinweise – der überwiegende Teil der Ergebnisse spricht für die Ablehnung der *Hypothese 4.2b*.

9.2.2 Entwicklung des Sachinteresses in Profilklassen und Klassen ohne Profil

Tabelle 9.10 stellt die Mittelwertunterschiede des mittleren musikalischen Sachinteresses zu den drei Messzeitpunkten sowie für die Klassen mit musikalischem beziehungsweise naturwissenschaftlichem Profil und deren direkte Parallelklassen ohne Profil (Vergleichsklassen) dar. Auch hier zeigen – verglichen mit den übrigen Untersuchungsgruppen – erwartungskonform die Schülerinnen und Schüler in den Musik-Klassen zu allen drei Messzeitpunkten mit Abstand das höchste Interesse an musikalischen Themen. Erneut ist der hohe Ausgangswert des musikalischen Sachinteresses in den Musik-Klassen direkt nach dem Übergang in die Sekundarstufe I (erster Messzeitpunkt) auffällig.

Hinsichtlich der Entwicklung des Interesses an musikalischen Themen zeigt sich in den Musik-Klassen zwischen erstem und zweitem Messzeitpunkt eine leichte (aber signifikante) Verringerung des musikalischen Sachinteresses, das sich anschließend zum dritten Messzeitpunkt jedoch auf einem hohen Niveau stabilisiert, sodass zwischen dem ersten und dritten Messzeitpunkt keine signifikante Abnahme des musikalischen Sachinteresses feststellbar ist. Somit kann *Hypothese 4.3a* empirisch untermauert werden.

Tabelle 9.10: Deskriptive Befunde – Entwicklung des durchschnittlichen musikalischen Sachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen)

	1. MZP		2. MZP		3. MZP	
	M	(SE)	M	(SE)	M	(SE)
Musik-Klassen	3.4	(0.05)	3.2 ▶	(0.06)	3.3	(0.06)
Musik-Vergleichsklassen	▼ 2.7	(0.07)	▼ 2.3 ▶	(0.08)	▼ 2.5 ▶	(0.08)
MINT-Klassen	▼ 2.5	(0.08)	▼ 2.3 ▶	(0.09)	▼ 2.4	(0.10)
MINT-Vergleichsklassen	▼ 2.6	(0.08)	▼ 2.3 ▶	(0.08)	▼ 2.3 ▶	(0.09)

▶ Wert im Vergleich zur Referenzkategorie „1. MZP“ signifikant verschieden ($p \leq .05$).

▼ Wert im Vergleich zur Referenzkategorie „Musik-Klassen“ signifikant niedriger ($p \leq .05$).

Zudem sind – entsprechend der *Hypothese 4.3b* – die Mittelwertunterschiede zwischen den Musik-Klassen und deren direkten Parallelklassen ohne Profil zu allen drei Messzeitpunkten signifikant. Auch im Vergleich zu den MINT-Klassen sowie den MINT-Vergleichsklassen weisen die Schülerinnen und Schüler in den Musik-Klassen zwischen fünfter und siebter Jahrgangsstufe ein signifikant höheres Sachinteresse in Musik auf.

In Tabelle 9.11 sind die Ergebnisse der deskriptiven Analysen zum durchschnittlichen naturwissenschaftlichen Sachinteresse im Längsschnitt sowie im Mehrgruppenvergleich dargestellt. Im Vergleich zu Tabelle 9.09 fällt auf, dass die Schülerinnen und Schüler in allen Untersuchungsgruppen nominell weniger Interesse für naturwissenschaftliche Themen (Sachinteresse) als für den Naturwissenschaftsunterricht (Fachinteresse) aufbringen.

Tabelle 9.11: Deskriptive Befunde – Entwicklung des durchschnittlichen naturwissenschaftlichen Sachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen)

	1. MZP		2. MZP		3. MZP	
	M	(SE)	M	(SE)	M	(SE)
MINT-Klassen	2.8	(0.07)	2.6 ▶	(0.07)	2.4 ▶	(0.07)
MINT-Vergleichsklassen	▼ 2.3	(0.08)	▼ 2.3	(0.08)	2.3	(0.07)
Musik-Klassen	▼ 2.5	(0.08)	▼ 2.4 ▶	(0.08)	2.4 ▶	(0.07)
Musik-Vergleichsklassen	▼ 2.3	(0.07)	▼ 2.1	(0.08)	▼ 2.1	(0.07)

▶ Wert im Vergleich zur Referenzkategorie „1. MZP“ signifikant verschieden ($p \leq .05$).

▼ Wert im Vergleich zur Referenzkategorie „MINT-Klassen“ signifikant niedriger ($p \leq .05$).

Dennoch berichten die Schülerinnen und Schüler in den MINT-Klassen zu allen drei Messzeitpunkten gegenüber den anderen Untersuchungsgruppen über ein vergleichsweise hohes Interesse an naturwissenschaftlichen Themen. In der fünften und sechsten Jahrgangsstufe ist das naturwissenschaftliche Sachinteresse der Schülerinnen und Schüler in den MINT-Klassen signifikant höher ausgeprägt als in den übrigen Untersuchungsgruppen. Zum dritten Messzeitpunkt berichten lediglich die Schülerinnen und Schüler in den Musik-Vergleichsklassen über signifikant weniger naturwissenschaftliches Sachinteresse als die Schülerinnen und Schüler in den MINT-Klassen. Für den gesamten Untersuchungszeitraum von der fünften bis zur siebten Jahrgangsstufe kann folglich kein signifikanter Interessenvorsprung der MINT-Klassen gegenüber den Vergleichsgruppen festgestellt werden, sodass *Hypothese 4.4b* abgelehnt werden muss. Im Längsschnitt nimmt das naturwissenschaftliche Sachinteresse in den MINT-Klassen signifikant ab und unterschreitet zum dritten Messzeitpunkt den theoretischen Mittelwert von 2.5, sodass *Hypothese 4.4a* ebenfalls verworfen werden muss.

9.2.3 Entwicklung des Fähigkeitsselbstkonzepts in Profilklassen und Klassen ohne Profil

Die längsschnittliche Entwicklung des durchschnittlichen domänenspezifischen Fähigkeitsselbstkonzepts wird deskriptiv für die verschiedenen Subgruppen analysiert und in Tabelle 9.12 für das musikalische Fähigkeitsselbstkonzept sowie in Tabelle 9.13 für das naturwissenschaftliche Fähigkeitsselbstkonzept dargestellt.

Tabelle 9.12: Deskriptive Befunde – Entwicklung des durchschnittlichen musikalischen Fähigkeitsselbstkonzepts zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen)

	1. MZP		2. MZP		3. MZP	
	M	(SE)	M	(SE)	M	(SE)
Musik-Klassen	3.2	(0.05)	3.0 ▶	(0.06)	3.0 ▶	(0.06)
Musik-Vergleichsklassen	▼ 2.8	(0.05)	▼ 2.6 ▶	(0.08)	▼ 2.8	(0.06)
MINT-Klassen	▼ 2.6	(0.07)	▼ 2.7	(0.07)	▼ 2.7	(0.08)
MINT-Vergleichsklassen	▼ 2.6	(0.07)	▼ 2.5	(0.07)	▼ 2.6	(0.08)

▶ Wert im Vergleich zur Referenzkategorie „1. MZP“ signifikant verschieden ($p \leq .05$).

▼ Wert im Vergleich zur Referenzkategorie „Musik-Klassen“ signifikant niedriger ($p \leq .05$).

Erwartungskonform weisen die Schülerinnen und Schüler in den Musik-Klassen zu allen drei Messzeitpunkten das vergleichsweise höchste musikalische Fähigkeitsselbstkonzept auf. Zwischen der fünften und sechsten Jahrgangsstufe verlieren die

Schülerinnen und Schüler in den Musik-Klassen zwar signifikant an Vertrauen in die eigenen musikalischen Fähigkeiten, jedoch stabilisiert sich zwischen dem zweiten und dritten Messzeitpunkt das musikalische Fähigkeitsselbstkonzept in den Musik-Klassen auf einem hohen Niveau, sodass in der Gesamtschau die Ergebnisse für *Hypothese 4.5a* sprechen. Die Mittelwertunterschiede zwischen den Musik-Klassen und deren direkten Parallelklassen ohne Profil sind zu allen drei Messzeitpunkten hypothesenkonform signifikant und auch gegenüber den MINT-Klassen sowie den MINT-Vergleichsklassen berichten die Schülerinnen und Schüler in den Musik-Klassen über ein signifikant höheres musikalisches Fähigkeitsselbstkonzept (*Hypothese 4.5b*).

In Tabelle 9.13 sind die deskriptiven Befunde zur Entwicklung des mittleren naturwissenschaftlichen Fähigkeitsselbstkonzepts in den verschiedenen Untersuchungsgruppen dargestellt. In den MINT-Klassen (und auch in den übrigen Untersuchungsgruppen) liegt das durchschnittliche naturwissenschaftliche Fähigkeitsselbstkonzept zu allen drei Messzeitpunkten über dem theoretischen Mittelwert von 2.5. Im Zeitverlauf erweist sich das naturwissenschaftliche Fähigkeitsselbstkonzept in allen Untersuchungsgruppen als recht stabil. Nach einem leichten – aber nur in den Musik-Klassen signifikanten – Anstieg zum zweiten Messzeitpunkt, pendelt sich das naturwissenschaftliche Fähigkeitsselbstkonzept zum dritten Messzeitpunkt auf dem Niveau des ersten Messzeitpunkts ein. Lediglich in den MINT-Klassen ist eine leichte (aber signifikante) Verringerung des naturwissenschaftlichen Fähigkeitsselbstkonzepts zwischen fünfter und siebter Jahrgangsstufe feststellbar. Dennoch ist das naturwissenschaftliche Fähigkeitsselbstkonzept in den MINT-Klassen zu allen drei Messzeitpunkten als überdurchschnittlich zu bezeichnen, was für die *Hypothese 4.6a* spricht. Signifikante Mittelwertunterschiede zwischen den MINT-Klassen und den anderen Subgruppen sind allerdings nicht feststellbar, sodass *Hypothese 4.6b* verworfen werden muss.

Tabelle 9.13: Deskriptive Befunde – Entwicklung des durchschnittlichen naturwissenschaftlichen Fähigkeitsselbstkonzepts zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen)

	1. MZP		2. MZP		3. MZP	
	M	(SE)	M	(SE)	M	(SE)
MINT-Klassen	2.8	(0.05)	2.9	(0.06)	2.7 ▶	(0.06)
MINT-Vergleichsklassen	2.7	(0.05)	2.9	(0.07)	2.6	(0.07)
Musik-Klassen	2.8	(0.05)	3.0 ▶	(0.06)	2.8	(0.05)
Musik-Vergleichsklassen	2.7	(0.05)	2.8	(0.07)	2.7	(0.05)

▶ Wert im Vergleich zur Referenzkategorie „1. MZP“ signifikant verschieden ($p \leq .05$).

▼ Wert im Vergleich zur Referenzkategorie „MINT-Klassen“ signifikant niedriger ($p \leq .05$).

Zusammenfassend kann für die in diesem Kapitel 9.2 berichteten deskriptiven Befunde festgehalten werden, dass die Schülerinnen und Schüler in den Musik-Klassen erwartungskonform zu allen drei Messzeitpunkten ein signifikant höheres musikalisches Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept in Musik aufwiesen als die Schülerinnen und Schüler in den Parallelklassen ohne Profil sowie in den MINT-Profilklassen und den MINT-Vergleichsklassen. Nach einer signifikanten (aber meist leichten) Abnahme des musikalischen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts in den Musik-Klassen im Zeitraum zwischen fünfter und sechster Jahrgangsstufe pendelten sich die Mittelwerte der drei Konstrukte anschließend zur siebten Jahrgangsstufe auf einem hohen Niveau ein, sodass im gesamten Untersuchungszeitraum von einem überdurchschnittlichen musikalischen Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept in den Musik-Klassen gesprochen werden kann. Im direkten Vergleich zwischen dem musikalischen Fach- und Sachinteresse fiel auf, dass das Interesse an musikalischen Themen in den Musik-Klassen im Zeitverlauf nominell leicht höher ausgeprägt war als das Interesse am Musikunterricht. Besonders auffällig war zudem, dass direkt nach dem Übergang in die Sekundarstufe I hohe Werte im musikalischen Fach- und Sachinteresse in den Musik-Klassen gemessen werden konnten, was die These der Wahl von Profilangeboten durch besonders interessierte Schülerinnen und Schüler untermauert (vgl. Kapitel 5.3).

Für die MINT-Klassen ergab sich hingegen ein differenzielles Bild: Während das naturwissenschaftliche Fachinteresse und das naturwissenschaftliche Fähigkeitsselbstkonzept in den ersten beiden Jahren der Sekundarstufe I gesteigert werden konnte, verringerte sich das Interesse der Schülerinnen und Schüler an naturwissenschaftlichen Themen im gleichen Zeitraum. Die Mittelwertunterschiede zwischen dem ersten und dritten Messzeitpunkt fielen in den MINT-Klassen zwar so-

wohl für das naturwissenschaftliche Fach- und Sachinteresse als auch Fähigkeitsselbstkonzept signifikant aus, dennoch verblieb zumindest das naturwissenschaftliche Fähigkeitsselbstkonzept in der siebten Jahrgangsstufe auf einem überdurchschnittlichen Niveau. Im Vergleich zu den anderen Untersuchungsgruppen zeichnete die MINT-Klassen zudem das signifikant höchste naturwissenschaftliche Sachinteresse in der fünften und sechsten Jahrgangsstufe aus. Die übrigen Mittelwertunterschiede im Mehrgruppenvergleich waren nur partiell signifikant.

9.3 Cross-Lagged-Panel-Modelle

Um die längsschnittlichen Beziehungen der drei zentralen Konstrukte des domänenspezifischen Fach- beziehungsweise Sachinteresses sowie des Fähigkeitsselbstkonzepts zu analysieren, werden nachfolgend verschiedene latente Cross-Lagged-Panel-Modelle (CLP) spezifiziert, die sich hinsichtlich der betrachteten Konstrukte, der Domänen Musik und Naturwissenschaften sowie hinsichtlich der Untersuchungsgruppen unterscheiden. Um mögliche Unterschiede zwischen den verschiedenen Facetten der Wertkomponente des EVT-Modells (Fach- bzw. Sachinteresse; vgl. Kapitel 4.1) und ihrem Zusammenhang zum Fähigkeitsselbstkonzept (Erwartungskomponente) herausstellen zu können, werden zunächst die längsschnittlichen Zusammenhangsstrukturen zwischen dem Fachinteresse und dem Fähigkeitsselbstkonzept und in einem zweiten Modell zwischen dem Sachinteresse und dem Fähigkeitsselbstkonzept untersucht. Die jeweiligen Modelle werden domänenspezifisch, das heißt getrennt für die Domänen Musik und Naturwissenschaften, betrachtet. Die ersten vier Modelle (musikalisches/naturwissenschaftliches Fach- bzw. Sachinteresse und Fähigkeitsselbstkonzept) werden in einem ersten Schritt für die Gesamtstichprobe berichtet (s. Kapitel 9.3.1 und 9.3.2). Anschließend wird die Zusammenhangsstruktur des musikalischen Fach- beziehungsweise Sachinteresses und des Fähigkeitsselbstkonzepts getrennt für die Musik-Klassen und deren direkte Parallelklassen ohne Profil präsentiert (s. Kapitel 9.3.3). In der Domäne Naturwissenschaften werden in entsprechenden Zusammenhangsanalysen der drei Konstrukte die MINT-Klassen und deren direkte Parallelklassen ohne Profil gegenübergestellt (s. Kapitel 9.3.4). In allen nachfolgenden latenten CLP-Modellen wird die Mehrebenenstruktur der Daten mittels des in *Mplus 7.11* implementierten Verfahrens für geclusterte Stichproben berücksichtigt (*type = complex*) und darüber hinaus der für ordinale Daten robuste WLSMV-Schätzer (Li, 2016a) verwendet. Fälle mit vollständig fehlenden Werten auf den jeweils betrachteten Variablen werden von den Analysen ausgeschlossen (analysierte Fallzahl siehe jeweilige Fußnote der Abbildungen). Die Ergebnisse der nachfolgenden CLP-Modelle sind außerdem unter Annahme von skalarer Messinvarianz im Längsschnitt sowie (falls Subgruppen verglichen werden; vgl. Kapitel 9.3.3 und 9.3.4) auch unter Annahme von skalarer Messinvarianz im Gruppenvergleich zu interpretieren. Ebenfalls modelliert werden die Auto-

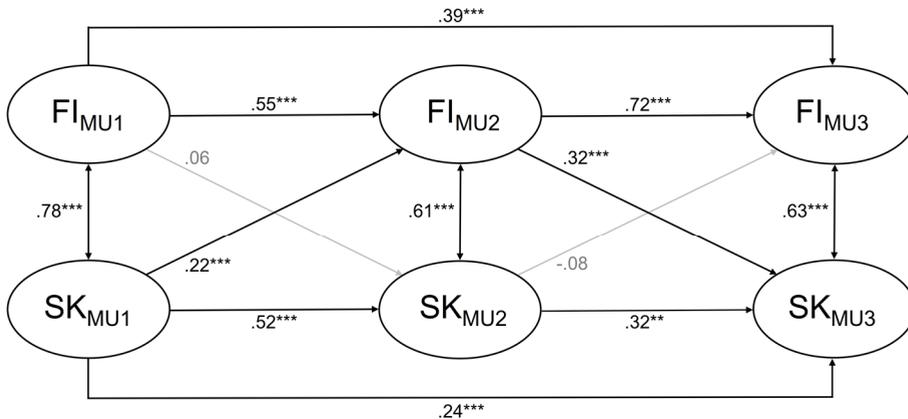
korrelationen der Indikatorresiduen eines Konstrukts über die Zeit, um die Korrelation der Messfehler zu berücksichtigen und somit Verzerrungen in den Stabilitätsannahmen zu vermeiden (Newsom, 2015).

9.3.1 Musikalisches Interesse und Fähigkeitsselbstkonzept

Zunächst wird die längsschnittliche Beziehung zwischen dem musikalischen Fachinteresse und dem Fähigkeitsselbstkonzept in Musik zwischen der fünften (erster Messzeitpunkt) und siebten Jahrgangsstufe (dritter Messzeitpunkt) analysiert (s. Abbildung 9.04). Alle Konstrukte werden latent modelliert, wobei – wie von Newsom (2015) empfohlen – die Autokorrelationen zwischen den Indikatorresiduen über die Zeit zugelassen werden, um die Korrelation der Messfehler entsprechend zu berücksichtigen und somit Verzerrungen in den Stabilitätsannahmen zu vermeiden. Die Fit-Indizes sprechen für einen ausgezeichneten Model-Fit des ersten latenten CLP-Modells ($CFI = .980$; $TLI = .980$; $RMSEA = .031$).

Unter Annahme invarianter Faktorladungen sowie Schwellenwerte innerhalb eines Konstrukts über die Zeit (skalare Messinvarianz) bestehen erwartungsgemäß (vgl. Kapitel 4.2) zwischen dem musikalischen Fachinteresse und dem Fähigkeitsselbstkonzept in Musik zu allen drei Messzeitpunkten hohe korrelative Zusammenhänge. Schülerinnen und Schüler, die an den Unterrichtsinhalten in Musik interessiert sind, trauen sich selbst auch viel in Musik zu und umgekehrt. Anhand der autoregressiven Pfade innerhalb eines Konstrukts können Aussagen über die zeitliche Stabilität getroffen werden. Die autoregressiven Pfade zwischen dem ersten und zweiten Messzeitpunkt sprechen sowohl für das musikalische Fachinteresse als auch musikalische Fähigkeitsselbstkonzept nur bedingt für eine Stabilität der Konstrukte. Für beide Merkmale liegen somit intraindividuelle Unterschiede in der Veränderung der Konstrukte innerhalb eines Schuljahres vor. Das musikalische Fachinteresse zum dritten Messzeitpunkt kann durch das Fachinteresse zum zweiten Messzeitpunkt hingegen erheblich besser vorhergesagt werden als dies für die Beziehung des musikalischen Fähigkeitsselbstkonzepts zwischen dem zweiten und dritten Messzeitpunkt der Fall ist. Das musikalische Fachinteresse wird im Zeitverlauf folglich stabiler während sich das musikalische Fähigkeitsselbstkonzept weniger stabil entwickelt.

Abbildung 9.04: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in der Gesamtstichprobe



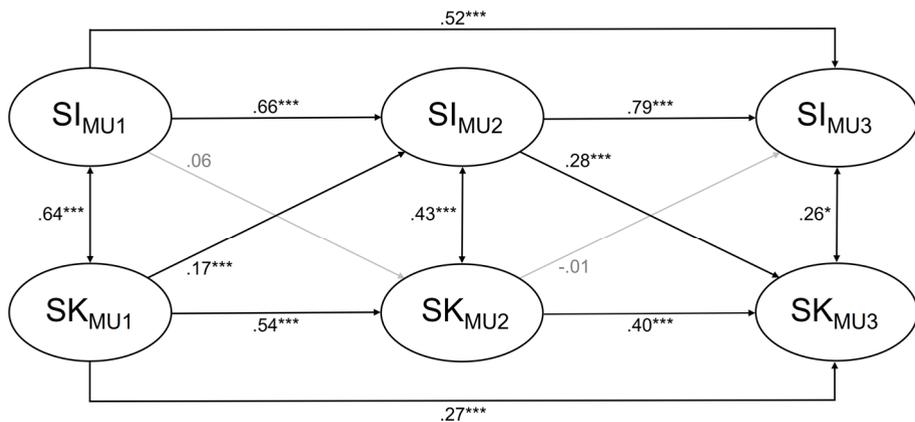
$n = 550$; $\chi^2(p) = 968.523 (.000)$; $df = 99$; $CFI = .980$; $TLI = .980$; $RMSEA = .031$

Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

Wird – unter Kontrolle der korrelativen sowie autoregressiven Zusammenhänge – zwischen zwei Messzeitpunkten nur einer der möglichen zwei Kreuzpfade signifikant, so verweist dieser auf einen bedeutsamen Cross-Lagged-Effekt zwischen den betrachteten Konstrukten. Zwischen dem ersten und zweiten Messzeitpunkt bildet der kreuzversetzte Pfad zwischen einem hohen musikalischen Fähigkeitsselbstkonzept in Jahrgangsstufe fünf und einem entsprechend hohen musikalischen Fachinteresse in der sechsten Jahrgangsstufe einen mittleren, signifikant positiven Effekt ab. Sich zu Beginn der fünften Jahrgangsstufe in Musik selbst viel zuzutrauen, geht somit mit einem höheren fachspezifischen Interesse am Unterrichtsfach Musik zu einem späteren Zeitpunkt einher. Zwischen fünfter und sechster Jahrgangsstufe kann folglich in der Domäne Musik für die Gesamtstichprobe die *Hypothese 5.1a* empirisch bekräftigt werden. Ein höheres musikalisches Fachinteresse in der sechsten Jahrgangsstufe steht hingegen mit einem hohen musikalischen Fähigkeitsselbstkonzept in der siebten Jahrgangsstufe in Zusammenhang. Dieser Befund ist auch vor dem Hintergrund einer geringen Stabilität des Fähigkeitsselbstkonzepts bei gleichzeitig vergleichsweise hoher Stabilität des musikalischen Fachinteresses zwischen sechster und siebter Jahrgangsstufe zu interpretieren, was eine höhere Veränderbarkeit des Fähigkeitsselbstkonzepts impliziert. Die aufgezeigten Befunde für die Zusammenhangsstruktur der Konstrukte zwischen sechster und siebter Jahrgangsstufe stützen *Hypothese 5.2a*.

In einem zweiten Schritt wird ein entsprechendes latentes CLP-Modell für den Zusammenhang des musikalischen Sachinteresses sowie des Fähigkeitsselbstkonzepts in Musik über die drei Messzeitpunkte für die Gesamtstichprobe modelliert (s. Abbildung 9.05). Auch für das zweite CLP-Modell weisen die Modellgütekriterien auf einen sehr guten Model-Fit hin ($CFI = .981$; $TLI = .981$; $RMSEA = .027$).

Abbildung 9.05: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in der Gesamtstichprobe



$n = 555$; $\chi^2(p) = 900.576 (.000)$; $df = 99$; $CFI = .981$; $TLI = .981$; $RMSEA = .027$

Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

Im Gegensatz zum vorherigen Modell (s. Abbildung 9.04) nimmt die korrelative Beziehung zwischen dem musikalischen Sachinteresse und dem Fähigkeitsselbstkonzept in Musik im Zeitverlauf erheblich ab. Während zu Beginn der fünften Jahrgangsstufe interessierte Schülerinnen und Schüler auch über ein hohes Fähigkeitsselbstkonzept in Musik verfügen, fällt der korrelative Zusammenhang zu Beginn der sechsten Klasse sowie zu Beginn der siebten Klassen deutlich geringer aus. Diese – im Vergleich zum Zusammenhang des musikalischen Fachinteresses und Fähigkeitsselbstkonzepts differentiellen – Befunde unterstreichen, dass die konkrete Definition der Wertkomponente als Interesse am Unterrichtsfach Musik (Fachinteresse) oder an musikalischen Themen (Sachinteresse) sowie deren Operationalisierung nicht unerheblich für die spätere Ergebnisinterpretation sind.

Wie im vorherigen Modell (s. Abbildung 9.04) liegen zwischen dem ersten und zweiten Messzeitpunkt relativ große intraindividuelle Unterschiede in der Entwicklung des musikalischen Sachinteresses beziehungsweise des Fähigkeitsselbstkonzepts vor, sodass zwischen der fünften und sechsten Jahrgangsstufe nur bedingt

von stabilen Konstrukten gesprochen werden kann. Wie bereits für das musikalische Fachinteresse berichtet, wird auch das musikalische Sachinteresse zwischen dem zweiten und dritten Messzeitpunkt stabiler, sodass das Interesse an musikalischen Themen in der siebten Jahrgangsstufe zu einem erheblichen Teil durch das vorhandene Sachinteresse in Jahrgang sechs vorhergesagt werden kann. Das musikalische Fähigkeitsselbstkonzept verliert hingegen im Zeitverlauf an Stabilität, sodass eine Veränderbarkeit des Konstrukts durch andere Merkmale wahrscheinlicher wird.

Analog zum vorherigen Modell (s. Abbildung 9.04) weisen die Kreuzpfade auch in diesem Modell (s. Abbildung 9.05) auf einen signifikant positiven Effekt zwischen dem musikalischen Fähigkeitsselbstkonzept zu Beginn der fünften Klasse und dem Interesse an musikalischen Themen in der sechsten Jahrgangsstufe hin. Schülerinnen und Schüler, die sich zu Beginn der Sekundarstufe I viel in Musik zutrauen, berichten somit ein Schuljahr später, in der sechsten Jahrgangsstufe, häufiger über ein hohes Interesse an musikalischen Themen. Dieser Befund untermauert *Hypothese 5.1b*. Ein hohes Sachinteresse in Musik in der sechsten Jahrgangsstufe ist hingegen mit einem höheren Fähigkeitsselbstkonzept in Klasse sieben assoziiert. In der Gesamtstichprobe kann somit für die Domäne Musik und für den Zeitraum zwischen der sechsten und siebten Jahrgangsstufe *Hypothese 5.2b* empirisch bekräftigt werden.

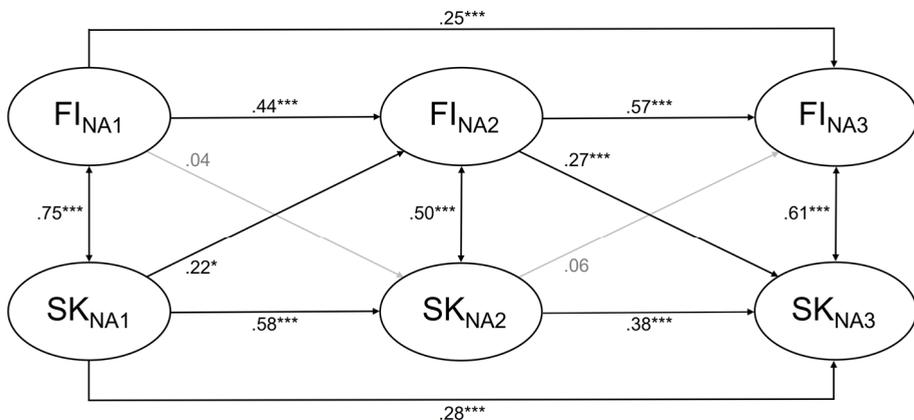
9.3.2 Naturwissenschaftliches Interesse und Fähigkeitsselbstkonzept

Analog zu den beiden vorherigen Modellen wird auch für die Domäne Naturwissenschaften je ein latentes CLP-Modell für die Beziehung des Fach- beziehungsweise Sachinteresses und des Fähigkeitsselbstkonzepts für die Gesamtstichprobe spezifiziert. Abbildung 9.06 gibt das für den Zusammenhang des naturwissenschaftlichen Fachinteresses und des Fähigkeitsselbstkonzepts in Naturwissenschaften aufgestellte Cross-Lagged-Panel-Modell wieder. Unter der Annahme skalarer Messinvarianz im Längsschnitt innerhalb eines Konstrukts liegen auch für dieses CLP-Modell die Modellgütekriterien oberhalb der empfohlenen Werte von CFI/TLI $\geq .95$ sowie unterhalb von RMSEA $\leq .05$ und sprechen somit für eine ausgezeichnete Modellgüte. Gemäß der in Kapitel 4.2 dargestellten Forschungsbefunde zum Zusammenhang der Erwartungs- und Wertkomponente zeigt sich auch anhand der vorliegenden Daten ein hoher wechselseitiger Zusammenhang zwischen dem naturwissenschaftlichen Fachinteresse und dem naturwissenschaftlichen Fähigkeitsselbstkonzept in der fünften Jahrgangsstufe. Ein hohes Interesse am naturwissenschaftlichen Unterricht geht daher auch mit einem hohen Fähigkeitsselbstkonzept in diesem Fachgebiet einher und umgekehrt.

Im Zeitverlauf bleiben diese korrelativen Zusammenhänge signifikant, jedoch wird der Effekt in der sechsten Klassenstufe zunächst etwas kleiner, bevor dieser in der siebten Jahrgangsstufe wieder ansteigt.

Für die autoregressiven Pfade des naturwissenschaftlichen Fachinteresses zeigen sich mittlere signifikante Effekte zwischen den drei Messzeitpunkten – wobei das Konstrukt Fachinteresse zwischen der sechsten und siebten Jahrgangsstufe an Stabilität gewinnt und somit ein höherer Anteil der Veränderung zum dritten Messzeitpunkt durch den zweiten Messzeitpunkt vorhergesagt werden kann als dies für die Veränderung zwischen dem ersten und zweiten Messzeitpunkt der Fall ist. Das Fähigkeitsselbstkonzept in den Naturwissenschaften ist hingegen zwischen dem ersten und zweiten Messzeitpunkt relativ stabil. Zwischen der sechsten und siebten Jahrgangsstufe verliert das Konstrukt Fähigkeitsselbstkonzept an Stabilität. Die Kreuzpfade stützen den in *Hypothese 5.1a* formulierten positiven Effekt zwischen dem naturwissenschaftlichen Fähigkeitsselbstkonzept in Klasse fünf und dem naturwissenschaftlichen Fachinteresse in Klasse sechs. Das naturwissenschaftliche Fähigkeitsselbstkonzept zum dritten Messzeitpunkt steht hingegen signifikant positiv mit dem naturwissenschaftlichen Fachinteresse zum zweiten Messzeitpunkt in Zusammenhang. Zwischen sechster und siebter Jahrgangsstufe sprechen die Ergebnisse für die Gesamtstichprobe in der Domäne Naturwissenschaften folglich für *Hypothese 5.2a*.

Abbildung 9.06: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in der Gesamtstichprobe

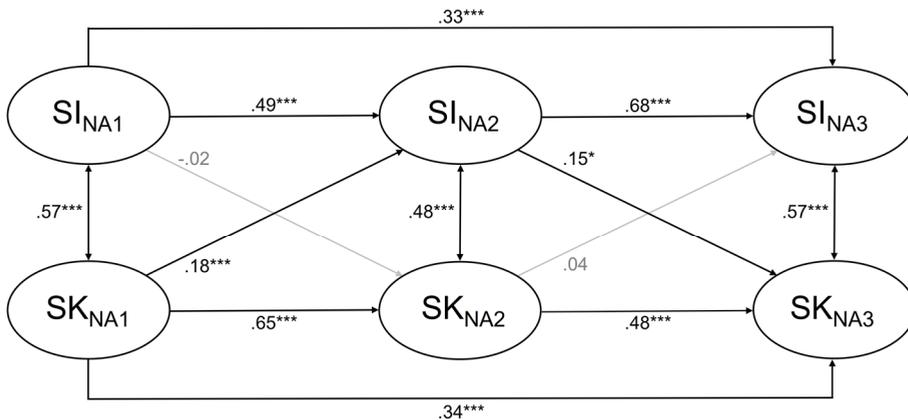


$n = 548$; $\chi^2(p) = 968.159 (.000)$; $df = 99$; $CFI = .973$; $TLI = .973$; $RMSEA = .031$

Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

Auch für die zweite Wertkomponente – das Sachinteresse – wird die längsschnittliche Beziehungsstruktur zum Fähigkeitsselbstkonzept in der Domäne Naturwissenschaften unter Annahme skalarer Messinvarianz in einem latenten CLP-Modell analysiert (CFI = .970; TLI = .971; RMSEA = .029). Die Ergebnisse sind in der folgenden Abbildung 9.07 dargestellt. Die korrelativen Zusammenhänge sind im Zeitverlauf durchweg signifikant positiv und verbleiben im Zeitverlauf auf einem ähnlichen, mittleren Niveau. Die autoregressiven Pfade sprechen nur bedingt für stabile Konstrukte, wobei sich wie in den vorherigen Modellen zwei Gemeinsamkeiten herausstellen lassen: Während das naturwissenschaftliche Sachinteresse im Zeitverlauf an Stabilität gewinnt, wird das fachspezifische Fähigkeitsselbstkonzept zwischen der sechsten und siebten Jahrgangsstufe instabiler. Dieser Befund spricht dafür, dass insbesondere zwischen der fünften und sechsten Klasse für das naturwissenschaftliche Sachinteresse intraindividuelle Unterschiede in der Entwicklung vorliegen und das Fähigkeitsselbstkonzept in den Naturwissenschaften im Zeitverlauf zunehmend durch weitere Faktoren veränderbar wird.

Abbildung 9.07: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in der Gesamtstichprobe



$n = 544$; $\chi^2(p) = 938.262 (.000)$; $df = 99$; $CFI = .970$; $TLI = .971$; $RMSEA = .029$

Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

Analog zu den bereits berichteten CLP-Modellen geht ein höheres naturwissenschaftliches Fähigkeitsselbstkonzept zu Beginn der fünften Klasse mit einem höheren naturwissenschaftlichen Sachinteresse in Klasse sechs einher und bekräftigt somit *Hypothese 5.1b*. Allerdings fällt dieser Effekt gegenüber dem Zusammenhang zwischen dem Fähigkeitsselbstkonzept und dem naturwissenschaftlichen Fachinteresse (s. Abbildung 9.06) geringer aus. Deutlich geringer als im vorherigen Modell

(s. Abbildung 9.06) ist auch der Kreuzpfad zwischen dem naturwissenschaftlichen Sachinteresse in der sechsten Klasse und dem naturwissenschaftlichen Fähigkeitsselbstkonzept in der siebten Klasse. Dennoch scheint das naturwissenschaftliche Sachinteresse in der sechsten Jahrgangsstufe positiv mit einem später berichteten Fähigkeitsselbstkonzept in der siebten Jahrgangsstufe assoziiert und unterstreicht folglich den in *Hypothese 5.2b* formulierten Zusammenhang.

Zusammenfassend lässt sich festhalten, dass in den für die Gesamtstichprobe spezifizierten latenten Cross-Lagged-Panel-Modellen unter Annahme skalarer längsschnittlicher Messinvarianz innerhalb eines Konstrukts unabhängig von der untersuchten Domäne (Musik oder Naturwissenschaften) mittlere bis hohe positive korrelative Zusammenhänge zwischen dem Fach- beziehungsweise Sachinteresse und dem fachspezifischen Fähigkeitsselbstkonzept festgestellt werden konnten. Die nach der Erwartungs-Wert Theorie angenommene positive wechselseitige Beziehung zwischen der Erwartungskomponente (Fähigkeitsselbstkonzept) und der Wertkomponente (Fach- bzw. Sachinteresse) konnte somit anhand der vorliegenden Daten bestätigt werden. Zudem wurde deutlich, dass sich die beobachteten korrelativen Beziehungen zwischen der Wert- und Erwartungskomponente durchaus in der Höhe der Korrelationskoeffizienten unterschieden, je nachdem, ob das Fach- oder das Sachinteresse als Wertkomponente im Modell betrachtet wurde. Die korrelativen Zusammenhänge zwischen dem domänenspezifischen Sachinteresse und dem Fähigkeitsselbstkonzept waren dabei stets niedriger ausgeprägt als die Zusammenhänge zwischen dem domänenspezifischen Fachinteresse und dem Fähigkeitsselbstkonzept.

Darüber hinaus wurde anhand der autoregressiven Pfade sichtbar, dass sich die Konstrukte des Fach- beziehungsweise Sachinteresses und auch des Fähigkeitsselbstkonzepts unabhängig von der untersuchten Domäne nur bedingt als stabile Konstrukte zwischen der fünften und siebten Jahrgangsstufe erwiesen. Dabei zeichneten sich zwei Tendenzen ab: Das Fach- beziehungsweise Sachinteresse wurde sowohl in Musik als auch den Naturwissenschaften über die Zeit stabiler, sodass ein größerer Anteil des Interesses am Musik- beziehungsweise Naturwissenschaftsunterricht oder an musikalischen/naturwissenschaftlichen Themen in der siebten Jahrgangsstufe durch das bestehende Interesse in der sechsten Jahrgangsstufe erklärt werden konnte als dies für den Zusammenhang zwischen der fünften und sechsten Jahrgangsstufe der Fall war. Das fachspezifische Fähigkeitsselbstkonzept verlor hingegen unabhängig von der untersuchten Domäne über die Zeit an Stabilität.

Hinsichtlich der analysierten Kreuzpfade sind die Ergebnisse insbesondere dann aufschlussreich, wenn unter Kontrolle der korrelativen sowie autoregressiven Zusammenhänge zwischen zwei Messzeitpunkten nur einer der möglichen Pfade signifikant wird, da sich daraus bedeutsame Effekte zwischen den zeitlich versetzt gemessenen Konstrukten ableiten lassen. Die in den Abbildungen 9.04, 9.05, 9.06 und 9.07 dargestellten latenten CLP-Modelle wiesen bezüglich der signifikanten Kreuz-

pfade ein homogenes Muster auf: Unabhängig von der Domäne (Musik oder Naturwissenschaften) sowie der analysierten Wertfacette (Fach- bzw. Sachinteresse) ging ein höheres domänenspezifisches Fähigkeitsselbstkonzept zu Beginn der fünften Jahrgangsstufe mit einem höheren Interesse am Unterrichtsfach Musik beziehungsweise Naturwissenschaften und einem höheren Interesse an musikalischen und naturwissenschaftlichen Themen in der sechsten Klasse einher. Zwischen sechster und siebter Jahrgangsstufe gewann hingegen das Fach- beziehungsweise Sachinteresse an Bedeutung: Ein höheres Fach- beziehungsweise Sachinteresse der sechsten Jahrgangsstufe war mit einer positiven Ausprägung des fachspezifischen Fähigkeitsselbstkonzepts in der siebten Jahrgangsstufe assoziiert – und zwar unabhängig von der betrachteten Domäne. Im direkten Modellvergleich waren für die Beziehungen des Fähigkeitsselbstkonzepts und des Fachinteresses im gesamten Untersuchungszeitraum sowohl in Musik als auch in den Naturwissenschaften stets größerer Effekte beobachtbar als im jeweils gleichen Zeitraum für den Zusammenhang zwischen Sachinteresse und Fähigkeitsselbstkonzept.

Daraus lässt sich ableiten, dass die Operationalisierung der Wertkomponente einen entscheidenden Faktor für die Ergebnisinterpretation darstellt und es nicht ausreicht, undifferenziert das ‚Interesse an Musik/Naturwissenschaften‘ zu erfassen. Dies fällt insbesondere auch dann ins Gewicht, wenn spezifische Lehr-Lern-Bedingungen (wie in Profilklassen) im Fokus der Analysen stehen, denen ein interessenförderndes Potential zugesprochen wird (s. Kapitel 5.6 und 5.7). Die in Kapitel 9.2 berichteten deskriptiven Befunde machten darüber hinaus deutlich, dass sowohl zwischen Musik- und MINT-Klassen als auch im Vergleich mit deren Parallelklassen ohne Profil erhebliche Unterschiede in den Ausprägungen und Entwicklungen des mittleren Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts bestehen. Daher werden im Folgenden erneut latente CLP-Modelle berichtet, die jedoch getrennt für Musik- und MINT-Profilklassen und deren Parallelklassen ohne Profil modelliert werden. Auch in den folgenden Kapiteln 9.3.3 und 9.3.4 wird zwischen dem Interesse am Musik- beziehungsweise Naturwissenschaftsunterricht (Fachinteresse) und dem Interesse an musikalischen/naturwissenschaftlichen Themen (Sachinteresse) differenziert, um Unterschiede in der Operationalisierung der Wertkomponente sichtbar zu machen.

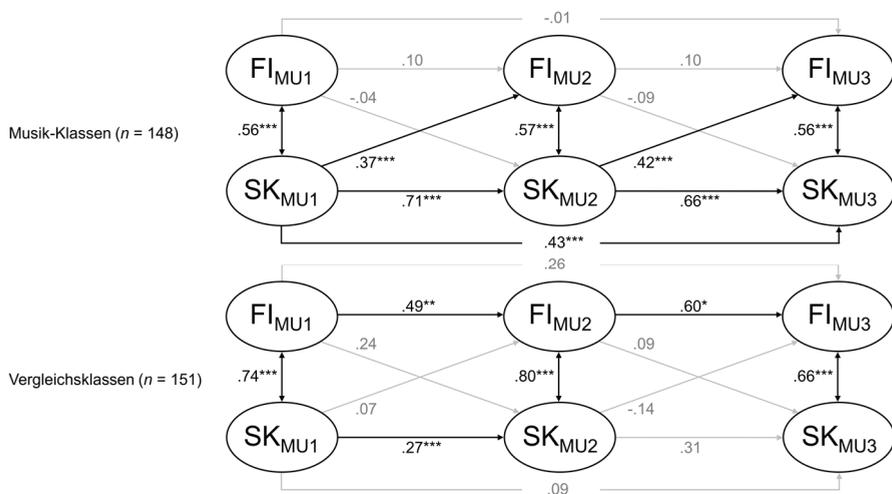
9.3.3 Musikalisches Interesse und Fähigkeitsselbstkonzept in Musik-Klassen

Um die spezifischen Rahmenbedingungen in Profilklassen (vgl. Kapitel 5.5, 5.6 und 5.7) in der Betrachtung der längsschnittlichen Zusammenhänge des Fach- beziehungsweise Sachinteresses sowie des Fähigkeitsselbstkonzepts zu berücksichtigen, wird – bezugnehmend auf die sechste Forschungsfrage – explorativ der Effekt des Klassenprofils auf die längsschnittliche Beziehungsstruktur der Konstrukte untersucht. Dazu werden im Folgenden pro Domäne (Musik/Naturwissenschaften) je

zwei latente CLP-Modelle vorgestellt, die getrennt für die Subgruppen Musik- beziehungsweise MINT-Klassen und deren direkte Parallelklassen ohne Profil im Multigruppenansatz modelliert werden. Dabei werden erneut die Zusammenhänge zwischen den zwei unterschiedlichen Wertkomponenten (Fach- bzw. Sachinteresse) und der Erwartungskomponente (Fähigkeitsselbstkonzept) getrennt betrachtet. In den vier folgenden latenten CLP-Modellen sind die Faktorladungen und Schwellenwerte innerhalb eines Konstrukts sowohl über die drei Messzeitpunkte als auch zwischen den Vergleichsgruppen invariant, sodass die folgenden Modelle unter der Annahme von skalarer Messinvarianz im Längsschnitt sowie im Gruppenvergleich interpretiert werden können. Ebenfalls berücksichtigt werden die Autokorrelationen der Indikatorresiduen eines Konstrukts über die Zeit, um die Korrelation der Messfehler zu berücksichtigen und somit Verzerrungen in den Stabilitätsannahmen zu vermeiden.

In Abbildung 9.08 ist die längsschnittliche Beziehungsstruktur zwischen dem musikalischen Fachinteresse und dem Fähigkeitsselbstkonzept in Musik dargestellt. Dabei werden die direkten und indirekten Pfade zwischen den Konstrukten getrennt für die Musik-Klassen und deren direkte Parallelklassen ohne Profil modelliert. Die Modellgütekriterien weisen auf einen guten Model-Fit hin (CFI = .951; TLI = .952; RMSEA = .044).

Abbildung 9.08: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in den Musik-Klassen und den Musik-Vergleichsklassen ohne Profil



$\chi^2(p) = 1637.015 (.000)$; $df = 199$; $CFI = .951$; $TLI = .952$; $RMSEA = .044$

Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

Im Vergleich zu dem in Kapitel 9.3.1 berichteten Modell für die Gesamtstichprobe (s. Abbildung 9.04) fällt auf, dass sich zwischen den Musik-Klassen und deren Parallelklassen ohne Profil erhebliche Unterschiede hinsichtlich der längsschnittlichen Beziehungsstruktur zwischen dem Fachinteresse und dem Fähigkeitsselbstkonzept ergeben, die im Modell der Gesamtstichprobe bisher verborgen blieben. Dies unterstreicht die Notwendigkeit einer getrennten Betrachtung der Profilklassen und ihrer Parallelklassen ohne Profil, um profilspezifische Effekte in der Entwicklung des Interesses sowie des Fähigkeitsselbstkonzepts herausstellen zu können.

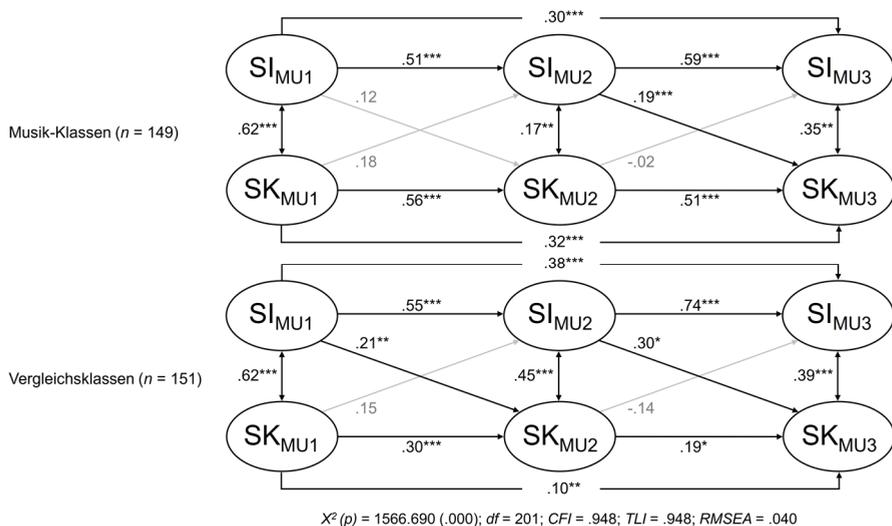
Sowohl im Modell der Musik-Klassen als auch im Modell der Vergleichsklassen ohne Profil zeigen sich zwischen den Konstrukten des musikalischen Fachinteresses und des Fähigkeitsselbstkonzepts zu allen drei Messzeitpunkten hohe bis mittlere signifikant positive korrelative Zusammenhänge. Schülerinnen und Schüler mit einem hohen musikalischen Fachinteresse weisen folglich auch ein hohes Fähigkeitsselbstkonzept in Musik auf und umgekehrt. Diese wechselseitigen Zusammenhänge sind im Modell der Vergleichsklassen ohne Profil im Längsschnitt stets etwas höher ausgeprägt als im Modell der Musik-Klassen. Die autoregressiven Pfade verweisen in den Musik-Klassen auf eine relative Stabilität des Konstrukts des musikalischen Fähigkeitsselbstkonzepts über die drei Messzeitpunkte. Hingegen ist das musikalische Fachinteresse in den Musik-Klassen zwischen der fünften und siebten Jahrgangsstufe als instabil – und somit die Entwicklung des Fachinteresses durch andere Faktoren vorhersagbar – zu bezeichnen. Das musikalische Fachinteresse in der sechsten Jahrgangsstufe steht eher mit dem Fähigkeitsselbstkonzept der fünften Jahrgangsstufe in Zusammenhang als mit dem Fachinteresse der fünften Jahrgangsstufe. Auch im weiteren Entwicklungsprozess kann das musikalische Fachinteresse der siebten Jahrgangsstufe eher durch das musikalische Fähigkeitsselbstkonzept als durch das musikalische Fachinteresse im vorherigen Schuljahr erklärt werden. Dieser Befund ist anhand der signifikanten Kreuzpfade zwischen den Konstrukten im Modell der Musik-Klassen ablesbar und unterstreicht den in *Hypothese 6.1a* angenommenen Zusammenhang. Im Modell der Vergleichsklassen ohne Profil finden sich hingegen keine signifikanten Kreuzpfade, sodass keine zeitversetzten Effekte zwischen den Konstrukten auszumachen sind und demnach keine der zur sechsten Forschungsfrage formulierten Hypothesen empirisch bekräftigt werden können.

In einem zweiten Schritt werden auch die Zusammenhänge zwischen der alternativen Wertkomponente (Sachinteresse) und dem Fähigkeitsselbstkonzept in der Domäne Musik in einem latenten CLP-Modell spezifiziert. Unter Annahme invarianter Faktorladungen und Schwellenwerte innerhalb eines Konstrukts im Längsschnitt sowie zwischen den Vergleichsgruppen weist das berechnete Modell akzeptable Fit-Werte auf (CFI = .948; TLI = .948; RMSEA = .040).

Wie in Abbildung 9.09 deutlich wird, zeigen sich in den Vergleichsklassen ohne Profil zwischen dem musikalischen Sachinteresse und dem Fähigkeitsselbstkonzept erwartungskonform mittlere bis hohe korrelative Zusammenhänge, wobei diese Effekte im Zeitverlauf kleiner werden. In den Musik-Klassen ist der wechselseitige

Zusammenhang zwischen den Konstrukten zu Beginn der fünften Klasse annähernd so groß wie in den Parallelklassen ohne Profil sowie wie im Modell der Gesamtstichprobe (s. Abbildung 9.05). Nach einem Schuljahr lässt sich jedoch nur noch ein schwacher signifikant positiver Zusammenhang zwischen dem musikalischen Sachinteresse und dem Fähigkeitsselbstkonzept feststellen, bevor sich der Zusammenhang in der siebten Jahrgangsstufe wieder intensiviert. Im Gegensatz zum musikalischen Fachinteresse im Modell der Musik-Klassen (s. Abbildung 9.08) weist das musikalische Sachinteresse in den Musik-Klassen eine mittlere Stabilität auf. Dies trifft auch auf das musikalische Sachinteresse in den Vergleichsklassen ohne Profil zu: In beiden Gruppen werden die Konstrukte über die Zeit stabiler, wobei in den Klassen ohne Profil der vergleichsweise größte Anteil der Veränderung im Sachinteresse in der siebten Jahrgangsstufe durch das Sachinteresse in der sechsten Jahrgangsstufe erklärt werden kann.

Abbildung 9.09: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in den Musik-Klassen und den Musik-Vergleichsklassen ohne Profil



Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

Im Gegensatz zu dem in Kapitel 9.3.1 dargestellten Modell der Gesamtstichprobe (s. Abbildung 9.05) weisen die Kreuzpfade im Modell der Vergleichsklassen ohne Profil sowohl zwischen fünfter und sechster Jahrgangsstufe als auch zwischen sechster und siebter Jahrgangsstufe auf einen signifikant positiven Effekt zwischen dem vorherigen musikalischen Sachinteresse und dem späteren Fähigkeitsselbstkonzept in Musik hin. Auch für die Musik-Klassen ist dieser Effekt beobachtbar –

jedoch nur zwischen dem zweiten und dritten Messzeitpunkt. Für die Vergleichsklassen ohne Profil sprechen die Befunde folglich für die Bekräftigung der *Hypothese 6.2b* im gesamten Untersuchungszeitraum. In den Musik-Klassen lässt lediglich der zwischen sechster und siebter Jahrgangsstufe signifikante Kreuzpfad eine empirische Untermauerung der *Hypothese 6.2b* zu.

Zusammenfassend lässt sich für die in diesem Kapitel berichteten latenten CLP-Modelle festhalten, dass sich sowohl zwischen den untersuchten Subgruppen (Musik-Klassen und direkte Parallelklassen ohne Profil) als auch im Vergleich zu den Modellen der Gesamtstichprobe (s. Kapitel 9.3.1) Unterschiede in den Zusammenhangsmustern des musikalischen Fach- beziehungsweise Sachinteresses und des Fähigkeitsselbstkonzepts in Musik herausstellen ließen: Unter Berücksichtigung der korrelativen und autoregressiven Effekte konnten in den Musik-Klassen signifikant positive kreuzversetzte Pfade zwischen dem vorherigen musikalischen Fähigkeitsselbstkonzept und dem späteren musikalische Fachinteresse der Schülerinnen und Schüler beobachtet werden. Diese Befunde waren konsistent für den gesamten Untersuchungszeitraum zwischen der fünften und siebten Jahrgangsstufe.

Für den Zusammenhang zwischen dem musikalischen Sachinteresse und dem domänenspezifischen Fähigkeitsselbstkonzept war hingegen ein gegenteiliges Ergebnis beobachtbar. Sowohl in den Musik-Klassen als auch in den Parallelklassen ohne Profil war das vorherige musikalische Sachinteresse mit einem späteren Fähigkeitsselbstkonzept in Musik assoziiert. Im direkten Vergleich der beiden Modelle (s. Abbildung 9.08 und 9.09) wurde somit wiederholt deutlich, dass das Interesse am Unterrichtsfach (Fachinteresse) und das Interesse an fachspezifischen Themen (Sachinteresse) keineswegs gleichzusetzende Facetten der Wertkomponente sind und sich in der Betrachtung der Zusammenhgangsstrukturen mit dem fachspezifischen Fähigkeitsselbstkonzept sogar konträre Tendenzen feststellen lassen.

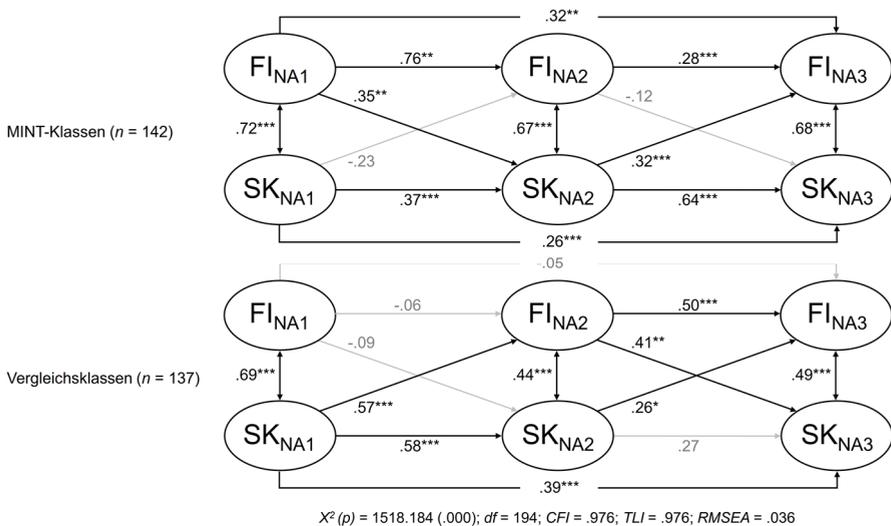
Auch hinsichtlich der Stabilität im Längsschnitt zeigten sich in den Musik-Klassen gegensätzliche Befunde zwischen dem musikalischen Fachinteresse und dem Sachinteresse. Während sich das musikalische Sachinteresse in den Musik-Klassen als relativ stabiles Konstrukt erwies, sprachen die nicht signifikanten autoregressiven Pfade des musikalischen Fachinteresses in den Musik-Klassen für eine fehlende Stabilität und somit eine höhere Veränderbarkeit des musikalischen Fachinteresses durch andere Faktoren.

9.3.4 Naturwissenschaftliches Interesse und Fähigkeitsselbstkonzept in MINT-Klassen

Analog zu den beiden vorherigen CLP-Modellen (s. Kapitel 9.3.3) werden für die MINT-Klassen sowie deren direkte Parallelklassen ohne Profil ebenfalls zwei latente CLP-Modelle für den Zusammenhang zwischen dem Fach- beziehungsweise Sachinteresse und dem Fähigkeitsselbstkonzept für die Domäne Naturwissenschaft-

ten spezifiziert. Das nachfolgend erste Modell betrachtet die längsschnittliche Beziehungsstruktur zwischen dem naturwissenschaftlichen Fachinteresse sowie dem domänenspezifischen Fähigkeitsselbstkonzept und stellt dabei MINT-Klassen und deren direkte Parallelklassen ohne Profil im Multigruppenvergleich gegenüber (s. Abbildung 9.10). Unter der Annahme skalarer Messinvarianz innerhalb eines Konstrukts im Längsschnitt und zwischen den Vergleichsgruppen weisen die Fit-Indizes auf eine ausgezeichnete Modellgüte hin (CFI = .976; TLI = .976; RMSEA = .036).

Abbildung 9.10: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in den MINT-Klassen und den MINT-Vergleichsklassen ohne Profil



Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

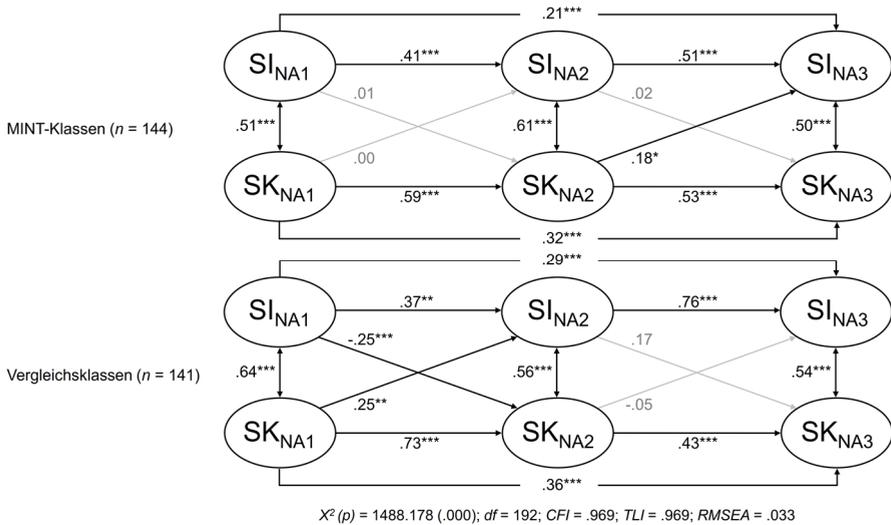
Sowohl für die MINT-Klassen als auch in den direkten Parallelklassen ohne Profil zeigen sich mittlere bis hohe positive korrelative Zusammenhänge zwischen dem naturwissenschaftlichen Fachinteresse und dem fachspezifischen Fähigkeitsselbstkonzept, die zu allen drei Messzeitpunkten signifikant sind. Hinsichtlich der Stabilität der beiden Konstrukte verweisen die autoregressiven Pfade in den MINT-Klassen und den Klassen ohne Profil auf gegensätzliche Befunde. Während das Fachinteresse in den MINT-Klassen im ersten Jahr nach dem Schulübertritt stabil ist, verliert es zwischen der sechsten und siebten Jahrgangsstufe an Stabilität. In den Parallelklassen ohne Profil zeigt sich erst zwischen dem zweiten und dritten Messzeitpunkt ein signifikanter autoregressiver Effekt und verweist somit auf eine erhöhte Veränderbarkeit des naturwissenschaftlichen Fachinteresses zwischen erstem und zweitem Messzeitpunkt sowie auf eine relative Stabilität des Konstrukts zwischen

sechster und siebter Jahrgangsstufe. Das naturwissenschaftliche Fähigkeitsselbstkonzept stabilisiert sich in den MINT-Klassen im Längsschnitt, während es in den Klassen ohne Profil im Zeitverlauf an Stabilität verliert.

In den Klassen ohne Profil lässt sich – äquivalent zum Modell der Gesamtstichprobe (s. Abbildung 9.06) – unter Kontrolle der korrelativen und autoregressiven Effekte ein signifikant positiver Kreuzpfad zwischen dem naturwissenschaftlichen Fähigkeitsselbstkonzept zu Beginn der fünften Klasse und dem Fachinteresse in der sechsten Jahrgangsstufe beobachten. Das naturwissenschaftliche Fähigkeitsselbstkonzept beim Übergang in die Sekundarstufe I hat somit – hypothesenkonform – einen positiven Effekt auf das Interesse, das die Schülerinnen und Schüler ein Schuljahr später dem naturwissenschaftlichen Unterricht entgegenbringen (*Hypothese 6.1a*). Zwischen sechster und siebter Jahrgangsstufe zeigt sich für die Vergleichsklassen ohne Profil ein signifikant positiver Effekt des naturwissenschaftlichen Fachinteresses auf das domänenspezifische Fähigkeitsselbstkonzept. Auch dieser Effekt ist bereits aus dem entsprechenden Modell in der Gesamtstichprobe bekannt (s. Abbildung 9.06). Durch den zusätzlichen kreuzversetzten Pfad zwischen dem naturwissenschaftlichen Fähigkeitsselbstkonzept der sechsten Jahrgangsstufe und dem naturwissenschaftlichen Fachinteresse der siebten Jahrgangsstufe wird die Interpretation der Zusammenhangsstruktur der beiden Konstrukte allerdings uneindeutig, auch wenn einer der Pfade in Relation zum entgegengesetzten Kreuzpfad substanziell höher ausfällt. Sowohl für *Hypothese 6.1a* als auch für *Hypothese 6.2a* finden sich folglich empirische Hinweise. Hingegen ist in den MINT-Klassen der erste Kreuzpfad im Zeitraum fünfte/sechste Jahrgangsstufe zwischen dem naturwissenschaftlichen Fachinteresse und dem Fähigkeitsselbstkonzept signifikant positiv. Schülerinnen und Schüler, die zu Beginn der fünften Klasse Interesse am Fach Naturwissenschaften aufweisen, trauen sich folglich ein Schuljahr auch mehr in den Naturwissenschaften zu. Dieser Befund ist vor dem Hintergrund der hohen Stabilität des naturwissenschaftlichen Fachinteresses bei gleichzeitig geringer Stabilität des Fähigkeitsselbstkonzepts in diesem Zeitraum zu interpretieren und betont die besondere Rolle des Fachinteresses zwischen fünfter und sechster Jahrgangsstufe in den MINT-Klassen. Entsprechend kann *Hypothese 6.2a* bekräftigt werden. Zwischen dem zweiten und dritten Messzeitpunkt besteht in den MINT-Klassen hingegen ein signifikant positiver Zusammenhang zwischen dem früheren naturwissenschaftlichen Fähigkeitsselbstkonzept und dem späteren naturwissenschaftlichen Fachinteresse. Dieser Befund spricht für das in *Hypothese 6.1a* formulierte Zusammenhangsmuster. Der Effekt des Fähigkeitsselbstkonzepts ist in Hinblick auf die in diesem Zeitraum geringe Stabilität des naturwissenschaftlichen Fachinteresses und der damit einhergehenden Varianzaufklärung des Fachinteresses durch andere Merkmale zu interpretieren.

Auch für die zweite Wertkomponente (Sachinteresse) wird in der Domäne Naturwissenschaften ein latentes CLP-Modell im Längsschnitt spezifiziert – wobei erneut die MINT-Klassen ihren direkten Parallelklassen ohne Profil gegenübergestellt werden (s. Abbildung 9.11). Die Fit-Indizes weisen unter Annahme skalarer Messinvarianz im Längsschnitt sowie im Multigruppenvergleich auf eine gute Modell-Daten-Passung hin (CFI = .969; TLI = .969; RMSEA = .033).

Abbildung 9.11: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in den MINT-Klassen und den MINT-Vergleichsklassen ohne Profil



Die Indizes 1, 2 und 3 stehen für die drei Messzeitpunkte. Die Pfadkoeffizienten sind standardisiert. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; grau markierte Pfade sind nicht signifikant.

Im Gegensatz zum vorherigen Modell zum Zusammenhang des naturwissenschaftlichen Fachinteresses und des Fähigkeitsselbstkonzepts (s. Abbildung 9.10) zeigen sich zwischen dem naturwissenschaftlichen Sachinteresse und dem fachspezifischen Fähigkeitsselbstkonzept zu Beginn der fünften Klasse in den MINT- und deren Vergleichsklassen etwas geringere wechselseitige Zusammenhänge. Sowohl in den Vergleichsklassen ohne Profil als auch in den MINT-Klassen bleiben die signifikant korrelativen Zusammenhänge jedoch zwischen den Konstrukten im Längsschnitt bestehen. Während die autoregressiven Pfade im vorherigen Modell (s. Abbildung 9.10) auf einen erheblichen Stabilitätsverlust des naturwissenschaftlichen Fachinteresses in den MINT-Klassen hindeuten, verbleibt das naturwissenschaftliche Sachinteresse in den MINT-Klassen im Zeitverlauf auf einem ähnlichen (mittleren) Stabilitätsniveau. Für das naturwissenschaftliche Fähigkeitsselbstkonzept in den MINT-Klassen kann zwischen den drei Messzeitpunkten nur ein Teil

der Veränderung durch den jeweils vorherigen Messzeitpunkt erklärt werden, sodass das Fähigkeitsselbstkonzept nur bedingt als stabiles Konstrukt zu bezeichnen ist. In den Klassen ohne Profil gewinnt das naturwissenschaftliche Sachinteresse über die Zeit an Stabilität während das Fähigkeitsselbstkonzept in den Naturwissenschaften im Zeitverlauf weniger stabil ausfällt.

Die Kreuzpfade zeigen unter Kontrolle der korrelativen und autoregressiven Effekte hingegen ein eher uneinheitliches Muster zwischen den MINT-Klassen und deren Parallelklassen ohne Profil: In den Vergleichsklassen ohne Profil können keine eindeutigen Cross-Lagged-Effekte zwischen dem naturwissenschaftlichen Sachinteresse und dem Fähigkeitsselbstkonzept abgeleitet werden, sodass keine der im siebten Kapitel aufgestellten Hypothesen empirisch zu untermauern ist. In den MINT-Klassen ist der einzige signifikante Kreuzpfad zwischen der sechsten und siebten Jahrgangsstufe zu beobachten. Der eher schwache Effekt verweist auf eine positive Beziehung zwischen dem naturwissenschaftlichen Fähigkeitsselbstkonzept der sechsten Jahrgangsstufe und dem Interesse an naturwissenschaftlichen Themen in der siebten Jahrgangsstufe (s. *Hypothese 6.1b*). Zudem lassen sich Überschneidungen zum vorherigen Modell (s. Abbildung 9.10) feststellen: Auch für das naturwissenschaftliche Fachinteresse konnte in den MINT-Klassen zwischen der sechsten und siebten Jahrgangsstufe ein signifikant positiver Zusammenhang zwischen dem vorherigen Fähigkeitsselbstkonzept und dem späteren Interesse ausgemacht werden.

Zusammenfassend lässt sich für die Domäne Naturwissenschaften festhalten, dass in den latenten CLP-Modellen zum Zusammenhang zwischen dem naturwissenschaftlichen Fach- beziehungsweise Sachinteresse und dem Fähigkeitsselbstkonzept sowohl deutliche Unterschiede zu den entsprechenden Modellen der Gesamtstichprobe (s. Abbildung 9.06 und Abbildung 9.07) als auch zwischen den MINT-Klassen und den Parallelklassen ohne Profil aufgezeigt werden konnten. Zudem zeigten sich in Abhängigkeit der jeweils betrachteten Wertkomponente (Fach- oder Sachinteresse) unterschiedliche Effekte.

In den MINT-Klassen spielte das naturwissenschaftliche Fachinteresse zu Beginn der Sekundarstufe I eine übergeordnete Rolle. Dies wurde einerseits an der hohen Stabilität des naturwissenschaftlichen Fachinteresses zwischen fünfter und sechster Jahrgangsstufe und zum anderen in einem kreuzversetzten Zusammenhang zum späteren Fähigkeitsselbstkonzept (bei gleichzeitiger Instabilität des Fähigkeitsselbstkonzepts im selben Zeitraum) sichtbar. Schülerinnen und Schüler, die zu Beginn der fünften Klasse Interesse am Fach Naturwissenschaften aufwiesen, trauten sich nach einem Jahr Profilunterricht folglich auch mehr in den Naturwissenschaften zu. Diese Entwicklung ging mit einer nachfolgend erhöhten Stabilität des naturwissenschaftlichen Fähigkeitsselbstkonzepts zwischen sechster und siebter Jahrgangsstufe einher, sodass in diesem Zeitraum ein hohes Fähigkeitsselbstkonzept wiederum ein höheres Interesse an den naturwissenschaftlichen Unterrichtsfächern in der siebten Jahrgangsstufe begünstigte.

Ein äquivalenter Zusammenhang fand sich auch zwischen der zweiten Wertkomponente (Sachinteresse) und dem Fähigkeitsselbstkonzept in den Naturwissenschaften. Schülerinnen und Schüler in den MINT-Klassen, die sich in der sechsten Jahrgangsstufe als kompetent erlebten, wiesen in der siebten Jahrgangsstufe auch ein entsprechend höheres Interesse an naturwissenschaftlichen Themen auf. In den Vergleichsklassen ohne Profil waren hingegen im gesamten Untersuchungszeitraum keine aussagekräftigen kreuzversetzten Zusammenhangsmuster zwischen dem naturwissenschaftlichen Sachinteresse und dem Fähigkeitsselbstkonzept zu beobachten.

10 Zusammenfassung und Diskussion der Ergebnisse

Im Folgenden werden die in Kapitel 9 berichteten Ergebnisse zusammenfassend diskutiert und unter Rückbezug auf theoretische Annahmen sowie den aktuellen Stand der Forschung eingeordnet. Dabei wird auch herausgestellt, inwiefern jeweils benannte Forschungsdesiderate in der vorliegenden Arbeit aufgegriffen wurden und welchen Beitrag die berichteten Ergebnisse zur Klärung ebendieser Forschungsdesiderate leisten können. Die Diskussion der Ergebnisse folgt der Chronologie der im siebten Kapitel aufgestellten Forschungsfragen, die sich auch in den aufeinander aufbauenden Analyseschritten in Kapitel 9 wiederfindet. Bezugnehmend auf die ersten drei Forschungsfragen werden in Kapitel 10.1 zunächst forschungsmethodische Aspekte in den Vordergrund gerückt und die Konstruktvalidität und Messinvarianz der zentralen Erhebungsinstrumente im Längsschnitt sowie im Mehrgruppenvergleich bestätigt. Anschließend werden zur Beantwortung der vierten Forschungsfrage deskriptive Befunde zum durchschnittlichen musikalischen und naturwissenschaftlichen Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept zwischen der fünften und siebten Jahrgangsstufe in den verschiedenen Subgruppen (Musik- und MINT-Klassen sowie Klassen ohne Profil) näher beleuchtet (Kapitel 10.2). Daran anknüpfend wird die längsschnittliche Beziehungsstruktur der Erwartungs- (Fähigkeitsselbstkonzept) und Wertkomponente (Fach- bzw.

Sachinteresse) zunächst für die Gesamtstichprobe sowie nachfolgend für die Profilklassen und Klassen ohne Profil getrennt beschrieben (Kapitel 10.3). In diesem Kapitel werden die Forschungsfragen fünf und sechs beantwortet. Abschließend werden aus den präsentierten Ergebnissen schulpraktische Implikationen abgeleitet (Kapitel 10.4) und auf Basis herausgestellter Forschungsdesiderate Anregungen für vertiefende Analysen sowie Folgestudien benannt (Kapitel 10.5).

10.1 Konstruktvalidität und Messinvarianz

Bezugnehmend auf die *erste Forschungsfrage* wurde zunächst die Konstruktvalidität der zentralen Messinstrumente zur Erfassung des domänenspezifischen Fach- und Sachinteresses sowie des domänenspezifischen Fähigkeitsselbstkonzepts mittels konfirmatorischer Faktorenanalysen unter Verwendung des WLSMV-Schätzers (Li, 2016a) sowie unter Berücksichtigung der Mehrebenenstruktur (*type = complex*) untersucht. Anhand der absoluten sowie inkrementellen Fit-Indizes konnte für die sechs zentralen Konstrukte (musikalisches bzw. naturwissenschaftliches Fach- und Sachinteresse sowie Fähigkeitsselbstkonzept) eine angemessene Konstruktvalidität bestätigt werden (Kapitel 9.1.1).

Um darüber hinaus die Vergleichbarkeit der zentralen Konstrukte des domänenspezifischen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts im Längsschnitt sowie im Multigruppenvergleich zu sichern, wurden die sechs zentralen Konstrukte anschließend hinsichtlich ihrer Messinvarianz geprüft (Kapitel 9.1.2 und 9.1.3). Zur Beantwortung der zweiten und dritten Forschungsfrage wurden dazu im Rahmen konfirmatorischer Faktorenanalysen verschiedene Messinvarianzmodelle mit zunehmenden Gleichheitsrestriktionen spezifiziert, um die verschiedenen Stufen der Messinvarianz im Längsschnitt sowie im Mehrgruppenvergleich zu testen. Da die Annahme eines kontinuierlichen Skalenniveaus in einer eigentlich ordinalen Datenstruktur möglicherweise falsche Schlussfolgerungen impliziert (vgl. u.a. Chen et al., 2020; Lubke & Muthén, 2004; Robitzsch, 2020; Svetina & Rutkowski, 2017; Svetina et al., 2020), wurde das ordinale Skalenniveau der vorliegenden Daten bei der Prüfung von Messinvarianz adäquat berücksichtigt. In Anlehnung an das von Liu et al. (2017) sowie Millsap und Yun-Tein (2004) empfohlene Vorgehen wurden entsprechende Messinvarianzmodelle mit ordinal-kategorialen Indikatoren spezifiziert und der für ordinale Daten besonders robuste WLSMV-Schätzer (Li, 2016a) verwendet. Für den Vergleich der unterschiedlich restriktiven Messinvarianzmodelle wurde die in *Mplus 7.11* (Muthén & Muthén, 2012) implementierte DIFFTEST-Option genutzt, um den χ^2 -Differenztest unter Verwendung des WLSMV-Schätzers angemessen zu berechnen (Asparouhov & Muthén, 2006). Die Mehrebenenstruktur der Daten wurde mittels des in *Mplus 7.11* implementierten Verfahrens für geclusterte Stichproben (*type = complex*) berücksichtigt. Zudem wurde auf die THETA-Parametrisierung zurückgegriffen, um die Residualvarianzen

der Parameter im vierten Schritt der Messinvarianzprüfung (strikte Invarianz) zwischen den Messzeitpunkten beziehungsweise zwischen den Vergleichsgruppen fixieren zu können.

Bezugnehmend auf die *zweite Forschungsfrage* konnte in den Modellen skalare Messinvarianz für die Konstrukte des musikalischen beziehungsweise naturwissenschaftlichen Fach- und Sachinteresses sowie des domänenspezifischen Fähigkeitsselbstkonzepts im Längsschnitt bestätigt werden.

Bezugnehmend auf die *dritte Forschungsfrage* konnte darüber hinaus anhand der inkrementellen sowie der absoluten Fit-Werte skalare Messinvarianz der sechs zentralen Konstrukte im Mehrgruppenvergleich bestätigt werden. Somit waren die Voraussetzungen für den Vergleich der Konstrukte zwischen verschiedenen Messzeitpunkten oder Gruppen gegeben (vgl. Brown, 2015; Christ & Schlüter, 2012). Entsprechende Entwicklungen des durchschnittlichen domänenspezifischen Fach- und Sachinteresses sowie Fähigkeitsselbstkonzepts wurden anschließend deskriptiv in Form von Mittelwertvergleichen zwischen der fünften, sechsten und siebten Jahrgangsstufe herausgestellt und dabei Unterschiede zwischen den Musik- und MINT-Klassen sowie deren jeweiligen Parallelklassen ohne Profil beleuchtet. Diese werden im Folgenden zur Beantwortung der vierten Forschungsfrage zusammenfassend diskutiert.

10.2 Entwicklung des Interesses und des Fähigkeitsselbstkonzepts in Profilklassen

Vor dem Hintergrund der bereits vielfach empirisch belegten Abnahme des durchschnittlichen Interesses der Schülerinnen und Schüler nach dem Übergang in die Sekundarstufe I (vgl. u.a. Anderhag et al., 2016; Fredricks & Eccles, 2002; Frenzel et al., 2010; Gottfried et al., 2001; Häußler & Hoffmann, 1995; Potvin & Hasni, 2014; Wigfield et al., 1997) können Lehr-Lern-Konzepte nach Höffler et al. (2014) bereits dann als erfolgreich gelten, wenn sich durch sie der Negativtrend der Interessen stoppen und sich das Interessenlevel der Schülerinnen und Schüler stabilisieren ließe. Die organisatorischen und didaktischen Gestaltungsmerkmale des musikalischen sowie naturwissenschaftlichen Profilverrichts (zusätzliche Unterrichtsstunden, praxisbezogene Lehr-Lern-Formate, der Einbezug externer Lehrkräfte beziehungsweise Lernorte sowie die Teilnahme an Auftritten oder Wettbewerben) werden mit ebendiesem Potential verbunden, die Interessenentwicklung der Schülerinnen und Schüler zu begünstigen (für eine detaillierte Beschreibung s. Kapitel 5.6 und 5.7; vgl. u.a. Brandt et al., 2008; Dolenc & Kazanis, 2020; Göllner, 2017; Heß, 2011a; Heß, 2017a; Mokhonko, 2016; Pabst-Krueger, 2020; Seidel et al., 2016; Simon, 2019; Steegh et al., 2021b). Hinsichtlich einer differenzierten Analyse zur Interessenentwicklung in Profilklassen nach dem Übergang in die Sekundarstufe I bestand jedoch ein Forschungsdesiderat (vgl. Krieg & Stubbe, 2018).

Bezugnehmend auf die *vierte Forschungsfrage* wurde daher zunächst untersucht, inwiefern sich zwischen der fünften und siebten Jahrgangsstufe Unterschiede in den durchschnittlichen Ausprägungen des domänenspezifischen Fach- und Sachinteresses zwischen Schülerinnen und Schülern in Musik- beziehungsweise MINT-Klassen und deren Parallelklassen ohne Profil beobachten lassen (Kapitel 9.2.1 und 9.2.2). Dabei wurde einerseits anhand deskriptiver Analysen geprüft, ob das jeweils domänenspezifische Fach- und Sachinteresse (d.h. in der Domäne Musik in den Klassen mit musikalischem Profil und in der Domäne Naturwissenschaften in den MINT-Profilklassen) im Untersuchungszeitraum auf einem überdurchschnittlichen Niveau gehalten werden kann. Andererseits wurde erwartet, dass die Schülerinnen und Schüler in den Profilklassen im Zeitverlauf über ein signifikant höheres domänenspezifisches Fach- und Sachinteresse verfügen als ihre Mitschülerinnen und Mitschüler in den Vergleichsgruppen (d.h. in Klassen ohne Profil und in den Profilklassen der jeweils anderen Domäne).

In Anlehnung an die von Heß (2011a) berichteten Befunde war erwartungskonform sowohl das musikalische Fach- als auch Sachinteresse in den Musik-Klassen besonders hoch ausgeprägt und verblieb auch im Zeitverlauf zwischen der fünften und siebten Jahrgangsstufe auf einem überdurchschnittlich hohen Niveau. Im Vergleich zu den übrigen Untersuchungsgruppen wiesen die Schülerinnen und Schüler in den Musik-Klassen zu allen drei Messzeitpunkten ein signifikant höheres musikalisches Fach- und Sachinteresse auf als ihre Mitschülerinnen und -schüler in den Vergleichsgruppen. Bemerkenswert war zudem, dass direkt nach dem Übergang in die Sekundarstufe I zu Beginn der fünften Jahrgangsstufe ein überdurchschnittlich hohes musikalisches Fach- und Sachinteresse in den Musik-Klassen gemessen werden konnte, wodurch die These zur Wahl von Profilangeboten durch besonders interessierte Schülerinnen und Schüler untermauert werden kann (vgl. Nonte et al., 2021; Pabst-Krueger, 2020; s. Kapitel 5.3).

In den deskriptiven Analysen zur Interessenentwicklung in den MINT-Klassen zeigte sich ebenfalls, dass die Schülerinnen und Schüler bereits mit einem überdurchschnittlich hohen naturwissenschaftlichen Fach- und Sachinteresse zu Beginn der fünften Jahrgangsstufe in die MINT-Klassen wechseln. Dieser Befund unterstreicht die Ansprache besonders naturwissenschaftlich interessierter Schülerinnen und Schüler als Zielgruppe für die MINT-Klassen, die schulintern auch als Zugangsvoraussetzung formuliert werden kann (vgl. Seidel et al., 2016; s. Kapitel 5.3). Innerhalb des ersten Schuljahres, in dem die Schülerinnen und Schüler am naturwissenschaftlichen Profilunterricht teilnahmen, konnte ihr Interesse am Naturwissenschaftsunterricht leicht gesteigert werden und lag nominell zum Teil über dem durchschnittlichen naturwissenschaftlichen Fachinteresse der Vergleichsgruppen. In der fünften und sechsten Jahrgangsstufe wiesen die Schülerinnen und Schüler in den MINT-Klassen außerdem ein signifikant höheres naturwissenschaftliches Sachinteresse auf als ihre Mitschülerinnen und Mitschüler in den Vergleichsgruppen. Diese Befunde bekräftigen die mit den profilspezifischen Gestaltungsmerkma-

len (z.B. fächerübergreifende, praxis-, projekt- oder kontextorientierte Unterrichtskonzepte) des MINT-Profilunterrichts verknüpften interessenförderlichen Potentiale (vgl. Höffler et al., 2014; Mokhonko et al., 2014; Palmer, 2009; Potvin & Hasni, 2014; Simon, 2019; Steegh et al., 2021b; Swarat et al., 2012). Auch Potvin et al. (2020) haben jüngst herausgestellt, dass das naturwissenschaftliche Interesse von Schülerinnen und Schülern der Sekundarstufe insbesondere von pädagogischen Innovationen (wie praxis-, kontext- und forschungsbasierten Lehr-Lern-Formen) profitiert, da diese als schülerzentrierte Unterrichtsmethoden die Aufmerksamkeit der Klasse erwecken und somit Interesse generieren beziehungsweise aufrechterhalten können.

Das überdurchschnittlich hohe naturwissenschaftliche Fachinteresse in den MINT-Klassen des fünften und sechsten Jahrgangs konnte jedoch nicht bis zur siebten Jahrgangsstufe aufrechterhalten werden, sodass das naturwissenschaftliche Fachinteresse zum dritten Messzeitpunkt signifikant niedriger ausfiel als zu Beginn der fünften Jahrgangsstufe. Verglichen mit den Schülerinnen und Schülern, die keine MINT-Klasse besuchen, lag das naturwissenschaftliche Fachinteresse der MINT-Klassen zu den drei Messzeitpunkten häufig zwar nominell über dem naturwissenschaftlichen Fachinteresse der Schülerinnen und Schüler in den anderen Untersuchungsgruppen. Jedoch waren die Mittelwertunterschiede nur partiell signifikant. Im siebten Jahrgang verringerte sich ebenfalls das naturwissenschaftliche Sachinteresse der Schülerinnen und Schüler der MINT-Klassen signifikant. Die negative Entwicklung des naturwissenschaftlichen Fach- und Sachinteresses insbesondere in der siebten Jahrgangsstufe steht in Einklang mit nationalen ebenso wie internationalen Forschungsbefunden (vgl. u.a. Daniels, 2008; Gottfried et al., 2001; Frenzel et al., 2010; Hoffmann & Lehrke, 1986; Hoffmann et al., 1998). Bereits Hoffmann und Lehrke (1986) stellten im Rahmen der IPN-Interessenstudie den vergleichsweise stärksten Rückgang des naturwissenschaftlichen Fach- und Sachinteresses in der siebten Jahrgangsstufe fest. Dieser Befund wurde auch in den von Daniels (2008) sowie Frenzel et al. (2010) berichteten Ergebnissen zur Entwicklung der naturwissenschaftlichen Interessen während der Sekundarstufe I untermauert. Im Vergleich zu anderen Schulfächern erwies sich zudem der Interessenrückgang in den naturwissenschaftlichen Fächern in bisherigen Studien als besonders ausgeprägt (vgl. Gottfried et al., 2001; Krapp, 1996).

Verschiedene Erklärungsansätze können zur Interpretation des Interessenrückgangs im siebten Jahrgang in den MINT-Klassen herangezogen werden: Zunächst ist die im Schulverlauf zu beobachtende zunehmende Differenzierung der individuellen Interessenstrukturen zu nennen, die mit einer gleichzeitigen Reduzierung potentieller Interessengebiete einer/eines einzelnen Schülerin/Schülers und damit einer Verringerung des durchschnittlichen Interesses (bspw. einer Klasse) einhergeht (s. Kapitel 2.5; vgl. Denissen et al., 2007; Todt & Schreiber, 1998). In der Betrachtung von Mittelwertstabilität über die Zeit sind solche intraindividuellen Interessendifferenzierungsprozesse jedoch nicht abbildbar, sodass personenzentrierte Analyseverfahren (wie *Latent Profile-Analysen*; vgl. Hagnaars & McCutcheon, 2002;

Lazarsfeld & Henry, 1968) ergänzend Aufschluss über die individuellen Interessenentwicklungen bieten könnten (vgl. Gaspard, Lauermaun et al., 2019; Krieg et al., 2021; Lazarides et al., 2020). Zudem sind entwicklungsbedingte Faktoren als Grund für den Interessenrückgang zu berücksichtigen, die insbesondere mit dem Einsetzen der Pubertät im Alter von etwa zwölf bis dreizehn Jahren (entspricht ca. der siebten Jahrgangsstufe; vgl. Lohaus & Vierhaus, 2019) mit einer verstärkten Ausbildung geschlechtsstereotypischer Interessenprofile einhergeht (vgl. Ehrtmann et al., 2019; Gaspard et al., 2015; Lazarides & Lauermaun, 2019; Marsh et al., 2005; Watt, 2004). So stellten Krieg et al. (2022) in querschnittlichen Analysen für MINT-Klassen des vorliegenden Samples fest, dass signifikante Unterschiede im naturwissenschaftlichen Sachinteresse zwischen Schülerinnen und Schülern bestehen. Entsprechend könnten Analysen zur Interessenentwicklung unter Kontrolle des Geschlechts Hinweise für die Gründe der Interessenabnahme liefern. Darüber hinaus kommen Änderungen in der Unterrichtsgestaltung und den in den jeweiligen Jahrgangsstufen angewendeten Lehr-Lern-Formen in Frage, die mit der negativen Interessenentwicklung in der siebten Jahrgangsstufe in den MINT-Klassen assoziiert sein könnten. Beispielsweise berichteten Brandt et al. (2008), Simon (2019) und Streller (2015), dass die mit dem Besuch von praxisorientierten außerschulischen Lernorten im MINT-Unterricht verknüpften positiven Effekte auf das fachspezifische Interesse der Schülerinnen und Schüler mit einem zeitlichen Abstand von wenigen Monaten verschwunden waren. Somit könnte zum Beispiel die Häufigkeit von durchgeführten Praxisphasen in den MINT-Klassen Hinweise auf die Bedingungsfaktoren der Negativentwicklung des naturwissenschaftlichen Interesses bieten. Detaillierte Auswertungen zu den in den Profilklassen des vorliegenden Samples stattfindenden Unterrichtsmethoden der einzelnen Jahrgänge befinden sich derzeit in Vorbereitung und werden im Abschlussbericht des Projektes ProBiNi erscheinen (vgl. Stubbe, Nonte et al., 2022).

Entwicklung des Fähigkeitsselbstkonzepts in Profilklassen und in Klassen ohne Profil

Neben der Interessenabnahme im Schulverlauf wurde auch die Verringerung des mittleren fachspezifischen Fähigkeitsselbstkonzepts während der Schulzeit empirisch bereits vielfach belegt (vgl. u.a. Dickhäuser et al., 2017; Jacobs et al., 2002; Marsh & Ayotte, 2003; Scherrer & Preckel, 2019; Watt, 2004). Besonders für die Phase des Übergangs in die Sekundarstufe I wird angenommen, dass sich diese auffallend negativ auf die Fähigkeitsselbsteinschätzung der Schülerinnen und Schüler auswirkt (Arens et al., 2013; Wigfield & Eccles, 1994). Diese negative Entwicklung wird mit einer Veränderung der Lernumgebung und damit auch der sozialen Vergleichsmaßstäbe als Faktor der Fähigkeitsselbstkonzeptgenese begründet (Arens, 2017; Eccles, Midgley et al., 1993; vgl. Kapitel 3.3). Aufgrund der in der Forschungsliteratur für zentrale Gestaltungsmerkmale des Profilklassenunterrichts berichteten positiven Effekte auf das Fähigkeitsselbstkonzept der Schülerinnen und Schüler (vgl. u.a. Bernecker et al., 2006; Eder, 2011; Höffler et al., 2014; Koivuhovi et al.,

2021; Kranefeld et al., 2015; Mokhonko et al., 2014; Nonte & Schwippert, 2012; Pabst-Krueger, 2020; Simon, 2019; Steegh et al., 2021b) wurde erwartet, dass das musikalische (bzw. naturwissenschaftliche) Fähigkeitsselbstkonzept in den Musik-Klassen (bzw. MINT-Klassen) im Beobachtungszeitraum auf einem überdurchschnittlichen Niveau gehalten werden kann. Zudem wurde angenommen, dass das musikalische (bzw. naturwissenschaftliche) Fähigkeitsselbstkonzept in den Musik-Klassen (bzw. MINT-Klassen) zwischen der fünften und siebten Jahrgangsstufe signifikant höher ausfällt als in den jeweiligen Vergleichsgruppen (*Forschungsfrage 4*).

Erwartungskonform wiesen die Schülerinnen und Schüler in den Musik-Klassen zu allen drei Messzeitpunkten das vergleichsweise höchste musikalische Fähigkeitsselbstkonzept auf. Die Mittelwertunterschiede zwischen den Musik-Klassen und den Vergleichsgruppen (MINT-Klassen, Klassen ohne Profil) fielen zu allen drei Messzeitpunkten signifikant aus. Zwischen der fünften und sechsten Jahrgangsstufe verloren die Schülerinnen und Schüler in den Musik-Klassen dennoch leicht (aber signifikant) an Vertrauen in die eigenen musikalischen Fähigkeiten. Der in diesem Zeitraum beobachtete Rückgang des musikalischen Fähigkeitsselbstkonzepts könnte Ausdruck des Big-Fish-Little-Pond-Effekts (BFLPE; vgl. Marsh et al., 2008) sein, der ein schlechteres Abschneiden der eigenen Fähigkeitsselbsteinschätzung gegenüber anderen leistungsstarken Kindern beschreibt und der durch einen veränderten Vergleichsmaßstab nach dem Schulübertritt begünstigt wird (Arens, 2017; Marsh et al., 2008; Marsh et al., 2017). Besonders in den Musik-Klassen ist diese Interpretation plausibel, da in diesen Klassen bereits zu Beginn der fünften Jahrgangsstufe ein überdurchschnittlich hohes Leistungsniveau zu beobachten ist (Haas & Stubbe, 2017). Zwischen der sechsten und siebten Jahrgangsstufe stabilisierte sich das musikalische Fähigkeitsselbstkonzept der Schülerinnen und Schüler in den Musik-Klassen auf einem überdurchschnittlich hohen Niveau. Dies könnte auf einen gewissen Gewöhnungseffekt an die neue Vergleichsnorm hinweisen, durch den sich nach der sechsten Jahrgangsstufe die Beurteilung des eigenen musikalischen Fähigkeitsselbstkonzepts in den Musik-Klassen normalisieren und folglich auf einem hohen Niveau einpendeln konnte. Nach einem anfänglichen verstärkten BFLPE am Schulübergang könnte nach einem Jahr Profilverricht nachfolgend der gegenteilige Reflected-Glory-Effekt (Marsh et al., 2000) überwiegen. In den Musik-Klassen spricht insbesondere das Gestaltungselement der gemeinsamen Teilnahme an öffentlichen Auftritten im Klassenensemble für die potentiell zunehmende Verstärkung des Reflected-Glory-Effekts, da die in Auftritten oder Konzerten erlebte positive Rückmeldung von außen mit positiven Effekten sowohl auf individueller Ebene (Motivation, Lernbereitschaft) als auch auf Klassenebene (Gemeinschaftsgefühl) verknüpft wird (Göllner, 2017; Jank & Stroh, 2006; Pabst-Krueger, 2020; vgl. Kapitel 5.6).

In den MINT-Klassen zeigten die deskriptiven Befunde, dass sich das naturwissenschaftliche Fähigkeitsselbstkonzept bereits zu Beginn der fünften Jahrgangsstufe auf einem überdurchschnittlich hohen Niveau befand und nach einem Jahr Profil-

unterricht in der sechsten Jahrgangsstufe sogar leicht gesteigert werden konnte. Dieser Befund stützt die mit einzelnen profilspezifischen Gestaltungsmerkmalen verknüpften Möglichkeiten zur Förderung der Fähigkeitsselbsteinschätzung (vgl. u.a. Höffler et al., 2014; Seidel et al., 2016; Simon, 2019). Da die Aufnahme in eine MINT-Klasse an bestimmte Zugangsvoraussetzungen geknüpft ist (s. Kapitel 5.3), könnten sich Schülerinnen und Schüler durch den tatsächlich realisierten Besuch der MINT-Klasse privilegiert fühlen, wovon auch ihr Fähigkeitsselbstkonzept profitiert. Für diese These spricht, dass die mit der Aufnahme in eine akademisch positiv selektive Gruppe verbundene Steigerung des Fähigkeitsselbstkonzepts (Reflected-Glory-Effekt; Marsh et al., 2000) insbesondere nach Übergangsprozessen beobachtet werden kann (Arens, 2017). Ob entsprechende Effekte in MINT-Klassen (und auch in Musik-Klassen) tatsächlich auftreten, könnte anhand von Mehrebenenanalysen unter Berücksichtigung der Beziehung der fachspezifischen Leistung und des entsprechenden fachspezifischen Fähigkeitsselbstkonzepts auf Schüler- und Klassenebene geprüft werden. Aufgrund der geringen Fallzahl des vorliegenden Samples (bzw. der Vergleichsgruppen) konnten im Rahmen dieser Arbeit die dafür notwendigen Mehrebenenanalysen jedoch nicht umgesetzt werden.

Im siebten Jahrgang sank das durchschnittliche naturwissenschaftliche Fähigkeitsselbstkonzept in den MINT-Klassen im Vergleich zum ersten Messzeitpunkt signifikant ab. Dieses Ergebnis steht in Einklang mit den von Koivuhovi et al. (2021) berichteten Befunden zur Entwicklung des mathematischen Fähigkeitsselbstkonzepts in profilierten Klassen in Finnland, die ebenfalls eine Abnahme des mathematischen Fähigkeitsselbstkonzepts in der Sekundarstufe I beobachteten. Allerdings unterschieden sich in dem von Koivuhovi et al. (2021) untersuchten Sample die Schülerinnen und Schüler in Profilklassen und ihre Mitschülerinnen und -schüler in nicht-profilieren Klassen unter Kontrolle von Hintergrundmerkmalen (z.B. Geschlecht, Mathematikleistung, Bildungsniveau der Mutter) nicht in der Entwicklung des mathematischen Fähigkeitsselbstkonzepts. Folglich könnte in vertiefenden Analysen auch für die in dieser Arbeit analysierten MINT-Klassen untersucht werden, inwieweit die Entwicklung des naturwissenschaftlichen Fähigkeitsselbstkonzepts mit weiteren Faktoren in Zusammenhang steht.

Ein enger Zusammenhang konnte bereits in der Vergangenheit zwischen dem domänenspezifischen Fähigkeitsselbstkonzept und dem domänenspezifischen Interesse empirisch beobachtet werden (vgl. u.a. Denissen et al., 2007; Dinkelmann & Buff, 2016; Fredricks & Eccles, 2002; Trautwein et al., 2012; Wigfield et al., 1997). Die beiden zunächst getrennt analysierten Entwicklungsprozesse der zwei Interessenfacetten (Fach- und Sachinteresse) und des domänenspezifischen Fähigkeitsselbstkonzepts wurden daher auf Grundlage der Expectancy-Value Theory (EVT; Eccles & Wigfield, 2002) zusammengeführt. Die Ergebnisse zur längsschnittlichen Beziehungsstruktur zwischen der Erwartungs- (Fähigkeitsselbstkonzept) und Wertkomponente (Interesse) werden im folgenden Kapitel 10.3 zusammenfassend diskutiert und zur Beantwortung die Forschungsfragen fünf und sechs herangezogen.

10.3 Zusammenhangsstruktur des Interesses und des Fähigkeitsselbstkonzepts

Wie aus der Beschreibung der empirischen Forschungsbefunde zur Expectancy-Value Theory (EVT; dt. Erwartungs-Wert Theorie; vgl. Eccles & Wigfield, 2002; vgl. Kapitel 4) deutlich wurde, stellt die EVT in der pädagogisch-psychologischen Forschung aktuell eines der am weitesten verbreiteten Modelle zur Untersuchung von Lernverhalten dar (vgl. u.a. Arens, 2021; Gaspard, Hasselhorn et al., 2019; Guo et al., 2017; Köller et al., 2019; Krieg et al., 2022; Nagy et al., 2006; Nonte et al., 2021). Die EVT unterscheidet die Erwartungs- von der Wertkomponente und nimmt an, dass diese in einem positiven Zusammenhang zueinanderstehen und durch das kulturelle Milieu, Sozialisationsinstanzen (vor allem Eltern und Lehrkräfte) sowie durch individuelle leistungsbezogene Erfahrungen geformt werden (Wigfield et al., 2016; Wigfield et al., 2017). Da eine empirische Trennung der zwei Subfacetten der Erwartungskomponente (Kompetenz- und Erfolgsüberzeugung) nicht möglich ist, hat sich in der empirischen Forschung die Operationalisierung der Erwartungskomponente über das Fähigkeitsselbstkonzept etabliert (vgl. u.a. Arens et al., 2019; Guo et al., 2017; Trautwein et al., 2012). Uneinigkeit besteht hingegen in der Operationalisierung der Wertkomponente. Es finden sich sowohl Forschungsarbeiten, die (unter Rückgriff auf die EVT) die Subfacette *intrinsic value* der Wertkomponente über Skalen zum Fachinteresse operationalisieren (vgl. u.a. Arens et al., 2019; Gaspard, Lauer mann et al., 2019; Trautwein et al., 2012) als auch Studien, die die Subfacette der Wertkomponente über Skalen zum Sachinteresse messen (vgl. u.a. Fritzsche et al., 2011, 2015; Marsh et al., 2005; Wigfield et al., 1997). Da bisherige Forschungsarbeiten jedoch entweder die eine oder die andere Interessenfacette zur Messung der Wertkomponente verwendeten, fehlte es aktuell an einer Gegenüberstellung möglicher Ergebnisunterschiede. Diesem Forschungsdesiderat wurde in der vorliegenden Arbeit (s. Kapitel 9.2 und 9.3) Rechnung getragen.

Ebenfalls bisher nur unzureichend untersucht wurde die Frage nach der längsschnittlichen Zusammenhangsstruktur zwischen der Erwartungs- und Wertkomponente. Aktuellere theoretische Abhandlungen zum EVT-Modell geben Hinweise darauf, dass die Erwartungskomponente (Fähigkeitsselbstkonzept) der Wertkomponente (Interesse) zeitlich vorgelagert ist (Wigfield et al., 2016). Auch empirisch zeigten sich häufiger Effekte zwischen dem vorherigen Fähigkeitsselbstkonzept und dem späteren Interesse als umgekehrt, wobei domänenspezifische Unterschiede zu beobachten waren (vgl. Arens et al., 2019; Lauer mann et al., 2017; Marsh et al., 2005). Eher interessenzentrierte Ansätze unterstreichen hingegen die Bedeutung der Wertkomponente, um zunächst die Hinwendung einer Person zu einem bestimmten Themenbereich zu erklären, wovon später auch die eigene Fähigkeitsselbsteinschätzung profitiert (Harackiewicz & Knogler, 2017; Wigfield et al., 2017; vgl. Eccles, 2005). Allerdings wurden die Fächer Naturwissenschaften und Musik bisher in den längsschnittlichen Analysen ausgespart, die sich vornehmlich auf die Domänen Mathematik, Englisch und Deutsch fokussierten (vgl. Arens et al., 2019;

Lauermann et al., 2017; Marsh et al., 2005). Darüber hinaus wurde die Entwicklung der Erwartungs- und Wertkomponente in der Sekundarstufe bisher vor allem anhand von zwei Messzeitpunkten sowie in höheren Jahrgängen der Sekundarstufe betrachtet (Lauermann et al., 2017; Marsh et al., 2005). Dieses Forschungsdesiderat wurde in der vorliegenden Arbeit aufgegriffen und die längsschnittliche Zusammenhangsstruktur der Wert- und Erwartungskomponente von der fünften bis siebten Jahrgangsstufe in den Domänen Musik und Naturwissenschaften untersucht.

Die längsschnittliche Beziehung zwischen dem domänenspezifischen Interesse (Fach- bzw. Sachinteresse) und dem Fähigkeitsselbstkonzept wurde anhand von latenten Cross-Lagged-Panel-Modellen (CLP) umgesetzt (vgl. Newsom, 2015). In allen berichteten latenten CLP-Modellen (vgl. Kapitel 9.3) wurde die Mehrebenenstruktur der Daten mittels des in *Mplus 7.11* implementierten Verfahrens für geclusterte Stichproben berücksichtigt (*type = complex*) und darüber hinaus der für ordinale Daten robuste WLSMV-Schätzer (Li, 2016a) verwendet. Die Ergebnisse der CLP-Modelle sind unter der Annahme skalarer Messinvarianz im Längsschnitt sowie (falls Subgruppen verglichen werden; vgl. Kapitel 9.3.3 und 9.3.4) zusätzlich unter Annahme skalarer Messinvarianz im Gruppenvergleich zu interpretieren. Ebenfalls modelliert wurden die Autokorrelationen der Indikatorresiduen eines Konstrukts über die Zeit, um die Korrelation der Messfehler zu berücksichtigen und somit Verzerrungen in den Stabilitätsannahmen zu vermeiden (vgl. Newsom, 2015).

Zur Beantwortung der *fünften Forschungsfrage* wurden zunächst vier latente CLP-Modelle für die Gesamtstichprobe analysiert, die sich hinsichtlich der untersuchten Domäne (Musik bzw. Naturwissenschaften) sowie hinsichtlich der verwendeten Wertkomponente (Fach- bzw. Sachinteresse) unterschieden. Erwartungskonform (vgl. Denissen et al., 2007; Dinkelmann & Buff, 2016; Fredricks & Eccles, 2002; Trautwein et al., 2012; Wigfield et al., 1997) zeigten sich unabhängig von der untersuchten Domäne (Musik oder Naturwissenschaften) zu allen drei Messzeitpunkten mittlere bis hohe positive korrelative Zusammenhänge zwischen dem Fähigkeitsselbstkonzept und dem Interesse. Allerdings wurde auch deutlich, dass sich in Abhängigkeit der analysierten Wertfacette (Fach- bzw. Sachinteresse) die Höhe der Korrelationskoeffizienten zwischen der Wert- und Erwartungskomponente durchaus unterschied. Die korrelativen Zusammenhänge zwischen dem domänenspezifischen Fachinteresse und dem Fähigkeitsselbstkonzept waren dabei stets höher ausgeprägt als die Zusammenhänge zwischen dem domänenspezifischen Sachinteresse und dem Fähigkeitsselbstkonzept.

Anhand der autoregressiven Pfade konnte beobachtet werden, dass sich unabhängig von der untersuchten Domäne (Musik oder Naturwissenschaften) das Fachbeziehungsweise Sachinteresse und auch das Fähigkeitsselbstkonzept nur bedingt als zeitlich stabile Konstrukte beschreiben ließen. Eine niedrige Stabilität eines Merkmals spricht für intraindividuelle Unterschiede in der Veränderung des Merkmals im Zeitverlauf (Christ & Schlüter, 2012) und ist ein Zeichen dafür, dass andere Faktoren zur Aufklärung der Varianz des Merkmals beitragen können (Reinders,

2006). Während das Fach- und Sachinteresse sowohl in Musik als auch in den Naturwissenschaften über die Zeit stabiler (und somit weniger veränderbar/vorhersagbar durch andere Faktoren) wurde, verlor hingegen sowohl das musikalische als auch naturwissenschaftliche Fähigkeitsselbstkonzept zwischen der fünften und siebten Jahrgangsstufe an Stabilität.

Hinsichtlich der Kreuzpfade zeichnete sich in den vier CLP-Modellen der Gesamtstichprobe ein einheitliches Muster ab: Unabhängig von der Domäne (Musik bzw. Naturwissenschaften) sowie der analysierten Wertfacette (Fach- bzw. Sachinteresse) ging ein höheres domänenspezifisches Fähigkeitsselbstkonzept zu Beginn der fünften Jahrgangsstufe mit einem höheren Interesse am Unterrichtsfach Musik beziehungsweise Naturwissenschaften und einem höheren Interesse an musikalischen und naturwissenschaftlichen Themen in der sechsten Klasse einher. Zwischen sechster und siebter Jahrgangsstufe war hingegen ein höheres Fach- beziehungsweise Sachinteresse mit einer späteren positiven Ausprägung des fachspezifischen Fähigkeitsselbstkonzepts in der siebten Jahrgangsstufe assoziiert. Im direkten Modellvergleich fielen die kreuzversetzten Effekte zwischen dem Fachinteresse und dem Fähigkeitsselbstkonzept (bzw. umgekehrt) im gesamten Untersuchungszeitraum sowohl in Musik als auch in den Naturwissenschaften höher aus als dies im jeweils gleichen Zeitraum für den Zusammenhang zwischen dem Sachinteresse und dem Fähigkeitsselbstkonzept der Fall war. Auch wenn die theoretische Definition der *intrinsic value*-Facette der Wertkomponente („Intrinsic value is the enjoyment the individual gets from performing the activity or the subjective interest the individual has in the subject.“; Eccles & Wigfield, 2002, S. 120) sowohl die Operationalisierung über das Interesse am Unterrichtsfach (Fachinteresse) als auch die Operationalisierung über das Interesse am Themenbereich (Sachinteresse) zulässt, waren unterschiedliche Ergebnisse in den Zusammenhangsmustern zu erwarten, da Fach- und Sachinteressen nicht gleichgesetzt werden können und nicht notwendigerweise gleichzeitig auftreten (vgl. Heß, 2017b; Häußler & Hoffmann, 1995). Demnach stellt die Operationalisierung der Wertkomponente einen entscheidenden Faktor für die Ergebnisinterpretation dar und sollte beim Vergleich von entsprechenden Forschungsbefunden, die die eine oder andere Operationalisierungsvariante nutzen, berücksichtigt werden.

Längsschnittliche Zusammenhangsstruktur des musikalischen Fach- und Sachinteresses und des Fähigkeitsselbstkonzepts in den Musik-Klassen

Um die spezifischen Rahmenbedingungen in Profilklassen (vgl. Kapitel 5.5, 5.6 und 5.7) in der Betrachtung der längsschnittlichen Zusammenhänge des Fach- beziehungsweise Sachinteresses sowie des Fähigkeitsselbstkonzepts zu berücksichtigen, wurden anschließend – bezugnehmend auf die *sechste Forschungsfrage* – für die Profilklassen und deren direkte Parallelklassen ohne Profil getrennte latente CLP-Modelle spezifiziert. Für die Musik-Klassen und deren Parallelklassen ohne Profil wurde die

längsschnittliche Beziehungsstruktur zwischen dem musikalischen Fach- beziehungsweise Sachinteresse sowie dem musikalischen Fähigkeitsselbstkonzept analysiert (s. Kapitel 9.3.3). Für die MINT-Klassen und deren Parallelklassen ohne Profil wurde die längsschnittliche Beziehungsstruktur zwischen dem naturwissenschaftlichen Fach- beziehungsweise Sachinteresse sowie dem naturwissenschaftlichen Fähigkeitsselbstkonzept im Multigruppenansatz untersucht (s. Kapitel 9.3.4). Die latenten CLP-Modelle wurden unter Annahme skalarer Messinvarianz im Längsschnitt sowie zusätzlich unter Annahme skalarer Messinvarianz im Gruppenvergleich berechnet. Analog zu den vorherigen CLP-Modellen der Gesamtstichprobe wurden ebenfalls die Autokorrelationen der Indikatorresiduen eines Konstrukts über die Zeit modelliert, um die Korrelation der Messfehler zu berücksichtigen. Die Mehrebenenstruktur der Daten wurde erneut mittels des in *Mplus 7.11* implementierten Verfahrens für geclusterte Stichproben einbezogen (*type = complex*) und darüber hinaus der für ordinale Daten robuste WLSMV-Schätzer (Li, 2016a) verwendet.

Unter Berücksichtigung der korrelativen und autoregressiven Effekte konnten im Modell der Musik-Klassen zwischen allen drei Messzeitpunkten signifikant positive kreuzversetzte Pfade zwischen dem vorherigen musikalischen Fähigkeitsselbstkonzept und dem späteren musikalischen Fachinteresse beobachtet werden (s. Abbildung 9.08). Schülerinnen und Schüler, die sich zwischen der fünften und siebten Jahrgangsstufe viel in Musik zutrauen, weisen demnach jeweils ein Schuljahr später auch ein höheres musikalisches Fachinteresse auf. Dieser Befund steht sowohl in Einklang zu der von Wigfield et al. (2016) formulierten theoretischen Annahme, die Erwartungskomponente sei der Wertkomponente zeitlich vorgelagert, als auch in Einklang mit den empirischen Befunden von Arens et al. (2019), die für die Domänen Mathematik und Englisch für Schülerinnen und Schüler des fünften bis neunten Jahrgangs Kreuzfade des gleichen Musters zwischen dem vorherigen Fähigkeitsselbstkonzept und dem späteren Interesse feststellen konnten (vgl. außerdem Lauer et al., 2017 für Schülerinnen und Schüler des neunten bis zwölften Jahrgangs in der Domäne Mathematik).

Die berichteten signifikanten Kreuzpfade zwischen dem vorherigen musikalischen Fähigkeitsselbstkonzept und dem späteren musikalischen Fachinteresse sind vor dem Hintergrund einer relativ hohen zeitlichen Stabilität des musikalischen Fähigkeitsselbstkonzepts bei gleichzeitiger Instabilität (d.h. Veränderbarkeit) des musikalischen Fachinteresses zu interpretieren. Somit kommt dem musikalischen Fähigkeitsselbstkonzept in den Musik-Klassen eine besondere Rolle zu – auch bezüglich der Erklärung der Entwicklung des musikalischen Fachinteresses.

In einem weiteren Schritt wurden die Zusammenhänge zwischen der alternativen Wertfacette (Sachinteresse) und dem Fähigkeitsselbstkonzept in der Domäne Musik in einem latenten CLP-Modell getrennt für die Musik-Klassen und deren direkte Parallelklassen ohne Profil spezifiziert (s. Abbildung 9.09). Unter Berücksichtigung der korrelativen und autoregressiven Beziehungen stand sowohl in den

Musik-Klassen als auch in den Parallelklassen ohne Profil das vorherige musikalische Sachinteresse mit einem späteren Fähigkeitsselbstkonzept in Musik in Zusammenhang. Ähnliche Ergebnisse fanden sich bereits in den von Arens et al. (2019) analysierten CLP-Modelle für Schülerinnen und Schüler der fünften bis neunten Jahrgangsstufe – allerdings für die Domäne Deutsch. Somit wird deutlich, dass die Zusammenhangsstruktur zwischen der Wert- und Erwartungskomponente nicht nur domänenspezifisch ist (vgl. Ergebnisse von Arens et al., 2019 in den Domänen Mathematik, Englisch und Deutsch; vgl. Gaspard, Hasselhorn et al., 2019; Gaspard, Nagengast et al., 2019), sondern sich zusätzlich je nach Wertfacette (Fach- bzw. Sachinteresse) sogar konträre Zusammenhangsmuster ergeben.

Auch hinsichtlich der Stabilität im Längsschnitt zeigten sich in den Musik-Klassen differente Befunde zwischen dem musikalischen Fachinteresse und dem Sachinteresse. Das musikalische Sachinteresse in den Musik-Klassen erwies sich gegenüber dem musikalischen Fachinteresse in den Musik-Klassen als relativ stabiles und somit nur bedingt veränderbares Konstrukt. Die Instabilität des musikalischen Fachinteresses spricht hingegen für eine hohe Veränderbarkeit des Konstrukts über die Zeit. Ein kombiniertes Modell, das das Fach- und Sachinteresse sowohl in Bezug zueinander als auch in Beziehung zum Fähigkeitsselbstkonzept setzt, könnte Aufschluss darüber geben, ob sich über das Interesse an musikalischen Themen (Sachinteresse) möglicherweise das musikalische Fähigkeitsselbstkonzept stärken ließ, das wiederum positive Effekte auf das Interesse am Unterrichtsfach Musik entfalten könnte. Die in dieser Arbeit präsentierten CLP-Modelle geben einen ersten Anhaltspunkt für diese These.

Längsschnittliche Zusammenhangsstruktur des naturwissenschaftlichen Fach- und Sachinteresses und des Fähigkeitsselbstkonzepts in den MINT-Klassen

Analog zu den CLP-Modellen in der Domäne Musik wurden auch für die Domäne Naturwissenschaften getrennte Modelle für die MINT-Klassen sowie deren direkte Parallelklassen ohne Profil zum Zusammenhang zwischen dem Fach- beziehungsweise Sachinteresse und dem Fähigkeitsselbstkonzept berechnet. Anhand der CLP-Modelle der Domäne Naturwissenschaften im Multigruppenvergleich wurde erneut deutlich, dass sich sowohl in Abhängigkeit der jeweils betrachteten Wertkomponente (Fach- oder Sachinteresse) als auch zwischen den MINT-Klassen und den Vergleichsklassen ohne Profil unterschiedliche Effekte im Zusammenhangsmuster der Wert- und Erwartungskomponente ergeben.

In den MINT-Klassen ist insbesondere die Rolle des naturwissenschaftlichen Fachinteresses zu Beginn der Sekundarstufe I hervorzuheben. Dies wurde im ersten CLP-Modell (s. Abbildung 9.10) einerseits an der hohen Stabilität des naturwissenschaftlichen Fachinteresses zwischen fünfter und sechster Jahrgangsstufe deutlich. Andererseits verwies das CLP-Modell der MINT-Klassen auf einen kreuzversetzten Zusammenhang zwischen dem naturwissenschaftlichen Fachinteresse der fünften

Jahrgangsstufe und dem späteren Fähigkeitsselbstkonzept (bei gleichzeitiger Instabilität des Fähigkeitsselbstkonzepts im selben Zeitraum). Schülerinnen und Schüler, die zu Beginn der fünften Klasse Interesse am Fach Naturwissenschaften aufwiesen, trauten sich nach einem Jahr Profilunterricht folglich auch mehr in den Naturwissenschaften zu. Diese Entwicklung ging mit einer nachfolgend erhöhten Stabilität des naturwissenschaftlichen Fähigkeitsselbstkonzepts zwischen dem zweiten und dritten Messzeitpunkt einher. Zwischen sechster und siebter Jahrgangsstufe begünstigte anschließend ein hohes Fähigkeitsselbstkonzept ein späteres höheres Interesse an den naturwissenschaftlichen Unterrichtsfächern. Dieser Effekt ließ sich auch zwischen der alternativen Wertfacette (Sachinteresse) und dem Fähigkeitsselbstkonzept in den Naturwissenschaften zwischen dem sechsten und siebten Jahrgang für Schülerinnen und Schüler der MINT-Klassen beobachten. Die übrigen Kreuzpfade zwischen dem naturwissenschaftlichen Sachinteresse und dem Fähigkeitsselbstkonzept waren jedoch sowohl in den MINT-Klassen als auch in den Vergleichsklassen ohne Profil inhaltlich nicht zu interpretieren, wodurch Rückschlüsse auf die Beziehungsstruktur der Konstrukte erschwert wurden.

Im direkten Vergleich der für die MINT-Klassen spezifizierten CLP-Modelle (Fachinteresse und Fähigkeitsselbstkonzept vs. Sachinteresse und Fähigkeitsselbstkonzept) fiel erneut auf, dass nicht nur die korrelativen Zusammenhänge zwischen dem naturwissenschaftlichen Fachinteresse und dem Fähigkeitsselbstkonzept stärker ausgeprägt waren als die entsprechenden Zusammenhänge zwischen dem Sachinteresse und dem Fähigkeitsselbstkonzept, sondern auch die kreuzversetzten Effekte zwischen dem Fachinteresse und dem Fähigkeitsselbstkonzept höher ausfielen als die entsprechenden Effekte zum Sachinteresse. Diese Beobachtung, die sich bereits in den CLP-Modellen der Gesamtstichprobe sowie in den Modellen der Musik-Klassen und deren Vergleichsklassen zeigte, ist durchaus plausibel, da die Konstrukte und folglich auch die Items der Skala zum Fachinteresse und zum Fähigkeitsselbstkonzept einen starken Unterrichtsbezug und somit inhaltlich eine größere Schnittmenge aufweisen (vgl. Kapitel 8.1.2). In der Gesamtschau der in dieser Arbeit berichteten Ergebnisse liegt daher der Schluss nahe, dass für die Betrachtung der Zusammenhänge der Wert- und Erwartungskomponente im Schulkontext eine Operationalisierung der Wertkomponente über das Fachinteresse zu bevorzugen ist.

10.4 Implikationen für die Praxis

Die Frage, ob die spezifischen Lehr-Lern-Bedingungen in Profilklassen das Interesse und das Fähigkeitsselbstkonzept von Schülerinnen und Schülern fördern können, ist von besonderer praktischer Relevanz. Denn die beiden Konstrukte gelten nicht nur als zentrale Indikatoren für schulisches Lern- und Leistungsverhalten, sie stehen auch in einem engen Beziehungsgefüge zueinander und müssen folglich zusammen gedacht und bestenfalls (gemeinsam) gefördert werden. Eine wesentliche

Erkenntnis dieser Arbeit ist, dass das Interesse und Fähigkeitsselbstkonzept in den ersten Jahren der Sekundarstufe I nur bedingt stabil und somit durch äußere Faktoren (positiv wie negativ) beeinflussbar ist. Darin stecken Chancen und Risiken gleichermaßen. Das Risiko eines ‚weiter so wie bisher‘ sollte durch die in der Vergangenheit vielfach empirisch belegte Verringerung des Interesses und des Fähigkeitsselbstkonzepts im Schulverlauf (vgl. Kapitel 2.6 und Kapitel 3.4) deutlich geworden sein. Auf der Seite der Chancen ist hingegen zu verorten, dass in dieser Arbeit Möglichkeiten aufgezeigt wurden, um das Interesse und das Fähigkeitsselbstkonzept von Schülerinnen und Schülern durch das Potential von Profilklassen oder einzelner profilspezifischer Gestaltungsmerkmale zu stärken. Diese Potentiale ließen sich auch auf andere schulische Kontexte und auf andere Domänen übertragen.

Wie sich in den Analysen herausstellte, herrscht in den Musik-Profilklassen offenbar eine Lehr-Lern-Atmosphäre, von der insbesondere das Fähigkeitsselbstkonzept der Schülerinnen und Schüler profitiert. Daher sollte ein besonderes Augenmerk auf das Potential der Musik-Klassen bezüglich der Stärkung des Fähigkeitsselbstkonzepts gelegt werden. Eine entscheidende Rolle für den Profit des Fähigkeitsselbstkonzepts dürfte hier auch die Art des Feedbacks spielen, die die Schülerinnen und Schüler durch ihre Teilnahme an der Musik-Klasse und die im Rahmen der Profilklassen stattfindenden Aktivitäten erhalten (vgl. O'Mara, Green & Marsh, 2006). Denn die für Musik-Klassen typischen sozial verbindenden und praxisorientierten Gestaltungselemente (z.B. Bildung eines Klassenensembles) strahlen nicht nur positiv auf die Gruppendynamik und die Klassengemeinschaft aus, sondern die bei Auftritten oder Konzerten erlebte positive Rückmeldung stärkt auch die eigene Fähigkeitsselbsteinschätzung. Von einem positiven musikalischen Fähigkeitsselbstkonzept profitiert schließlich auch das musikalische Fachinteresse, wie die Analysen der vorliegenden Arbeit zeigen. Übertragen auf weitere Kontexte und Domänen ließe sich dieser Effekt auch durch andere praxis- und projektorientierte Unterrichtskonzepte erzeugen, in denen die Schülerinnen und Schüler die Chance erhalten, sich als selbstwirksam wahrzunehmen und ihre Stärken zum gemeinsamen Erreichen eines (Projekt-)Ziels einzubringen. Wichtig ist dabei zu bedenken, dass solche Interventionen zur Stärkung motivational-affektiver Merkmale (wie des Fähigkeitsselbstkonzepts oder des Interesses) der Schülerinnen und Schüler immer domänenspezifisch ihre größte Wirkung entfalten (vgl. Gaspard et al., 2016).

In den MINT-Klassen profitierte hingegen zu Beginn der Sekundarstufe I vor allem das naturwissenschaftliche Interesse der Schülerinnen und Schüler. Das naturwissenschaftliche Fachinteresse konnte in den MINT-Klassen zwischen der fünften und sechsten Jahrgangsstufe nicht nur gesteigert werden, sondern war zudem mit einem höheren naturwissenschaftlichen Fähigkeitsselbstkonzept nach der sechsten Jahrgangsstufe verknüpft. Den Schülerinnen und Schülern im Rahmen des MINT-Unterrichts das selbstständige Ausprobieren und Experimentieren sowie projektorientiertes Arbeiten zu ermöglichen, kann demnach sowohl das naturwissenschaftliche Interesse aufrechterhalten und (indirekt) auch das Selbstvertrauen in

die eigenen Fähigkeiten stärken. Ein weiterer Faktor für die positive Interessengene-
nese könnte im profilspezifischen Gestaltungsmerkmal des Einbezugs außerschuli-
scher Lernorte (z.B. Schülerlabore) liegen. Hier werden Unterrichtsthemen prak-
tisch erfahrbar und außerhalb des schulischen Kontexts vermittelbar. Schülerlabore
finden sich mittlerweile nicht nur im naturwissenschaftlichen Bereich, sondern grei-
fen auch gesellschaftswissenschaftliche, politische, historische, künstlerische oder
sprachliche Themen auf und ermöglichen so einen alltagspraktischen Zugang zu
diesen Bereichen²⁰. Um den Einbezug außerschulischer Lernorte möglichst lang-
fristig nachhaltig zu gestalten, fügt sich der Besuch idealerweise thematisch-passend
in die Unterrichtskonzeption des Regelunterrichts ein, wird durch eine angemessene
Vor- und Nachbereitung im schulischen Kontext begleitet und in regelmäßigen Ab-
ständen wiederholt (vgl. Brandt et al., 2008; Mokhonko et al., 2014; Reimann et al.,
2020).

10.5 Forschungsdesiderate

Die Ergebnisse dieser Arbeit zeigen, dass es sich lohnt, innovative Lehr-Lern-For-
men, wie sie in Profilklassen zu finden sind, zu entwickeln und anzuwenden, diese
wissenschaftlich zu evaluieren, aus den Studienergebnissen zu lernen und Implika-
tionen für die Schulentwicklung abzuleiten. Im Anschluss an die in dieser Arbeit
gewonnen Ergebnisse zur Entwicklung des Interesses und des Fähigkeitsselbstkon-
zepts in Profilklassen ergeben sich weiterführend sowohl methodisch als auch in-
haltlich weitere Forschungsfragen, um das Phänomen der Profilklassen noch besser
verstehen zu können. So könnten ergänzend zu dem in dieser Arbeit verfolgten
variablenzentrierten Ansatz personenzentrierte Analyseverfahren (wie *Latent Profile*-
Analyse; Hagenaars & McCutcheon, 2002; Lazarsfeld & Henry, 1968) zum einen
Aufschluss über intraindividuelle Differenzierungsprozesse des Interesses sowie
des Fähigkeitsselbstkonzepts während der Sekundarstufe I bieten (vgl. Gaspard,
Lauermaun et al., 2019; Krieg et al., 2021; Lazarides et al., 2020; Rosenzweig &
Wigfield, 2017). Zum anderen könnte im Rahmen personenzentrierter Analysever-
fahren der Zusammenhang zwischen schülerspezifischen Erwartungs-Wert-Profi-
len und weiteren Merkmalen (z.B. Unterrichtsgestaltungsmerkmale oder Eltern-
bzw. Lehrkräfteeinschätzungen) sowie Veränderungen der Profile im Zeitverlauf
betrachtet werden (*Latent Transition*-Analyse; Collins & Lanza, 2013).

Anknüpfend an bestehende Forschungsbefunde, die insbesondere (Musik-)Profi-
filklassen als sozial positiv selektierte und leistungsstarke Gruppe (vgl. Koivuhovi
et al., 2020; Nonte, 2022) sowie den musikalischen und naturwissenschaftlichen Be-
reich als besonders geschlechtsstereotypen Bereich ausweisen (vgl. Ehrtmann et al.,
2019; Lazarides & Lauermaun, 2019; McPherson, Osborne, Barrett, Davidson &

²⁰ vgl. z.B. <https://www.forschungs-werkstatt.de/>

Faulkner, 2015), könnten die in dieser Arbeit berichteten Ergebnisse unter Berücksichtigung weiterer Aspekte, wie des sozialen Status, des Geschlechts oder der fachspezifischen Kompetenzen der Schülerinnen und Schüler betrachtet werden. Erste vertiefende Analysen zu den in dieser Arbeit dargestellten latenten Cross-Lagged-Panel-Modellen weisen jedoch darauf hin, dass die berichteten Zusammenhangsmuster zwischen dem domänenspezifischen Interesse und dem Fähigkeitsselbstkonzept auch unter Kontrolle des sozioökonomischen Hintergrunds konstant bleiben.

Zudem könnten Mehrebenenanalysen Aufschluss über soziale Vergleichsprozesse (Big-Fish-Little-Pond-Effekt bzw. Reflected-Glory-Effekt) in Profilklassen geben. Da diese anhand des vorliegenden Samples jedoch nicht umsetzbar waren, bedarf es dazu eines größeren Samples, das im Rahmen einer Folgestudie erhoben werden könnte. Um die Herausforderungen der Datenerhebung in Schulen besser bewältigen und das Potential wissenschaftlich begleiteter Schulentwicklung besser nutzen können, bleibt daher zu diskutieren, welche schulpolitischen Regelungen aber auch Anreize geschaffen werden müssen, um Schulen die Teilnahme an wissenschaftlichen Untersuchungen zu erleichtern. Über einen verstärkten Erkenntnis-transfer zwischen Wissenschaft und Schulpraxis könnte darüber hinaus ein echter gegenseitiger Austausch zwischen Wissenschaft und Schulpraxis ermöglicht werden, um über Maßnahmen und Effekte von Schul- und Klassenprofilierungsmaßnahmen ins Gespräch zu kommen.

Not everything that can be counted counts and not everything that counts can be counted.

Albert Einstein

11 Verzeichnisse

11.1 Literaturverzeichnis

- Abeles, H. (2009). Are musical instrument gender associations changing? *Journal of Research in Music Education*, 57(2), 127–139. <https://doi.org/10.1177/0022429409335878>
- Abramo, J. M. (2011). Gender differences of popular music production in secondary schools. *Journal of Research in Music Education*, 59(1), 21–43. <https://doi.org/10.1177/0022429410396095>
- Altrichter, H. & Heinrich, M. (2007). Kategorien der Governance-Analyse und Transformationen der Systemsteuerung in Österreich. In H. Altrichter, T. Brüsemeister & J. Wissinger (Hrsg.), *Educational Governance. Handlungskoordination und Steuerung im Bildungssystem* (S. 55–103). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90498-6_3
- Altrichter, H., Heinrich, M. & Soukup-Altrichter, K. (2011). Schulprofilierung – Annäherung an ein Phänomen. In H. Altrichter, M. Heinrich & K. Soukup-Altrichter (Hrsg.), *Schulentwicklung durch Schulprofilierung? Zur Veränderung von Koordinationsmechanismen im Schulsystem* (S. 11–45). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92825-8_1

- Altrichter, H. & Rürup, M. (2010). Schulautonomie und die Folgen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 111–144). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92245-4_5
- Altrichter, H., Rürup, M. & Schuchart, C. (2016). Schulautonomie und die Folgen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (2. überarb. und aktual. Aufl., S. 107–149). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-18942-0_5
- Anderhag, P., Wickman, P.-O., Bergqvist, K., Jakobson, B., Hamza, K. M. & Säljö, R. (2016). Why do secondary school students lose their interest in science? Or does it never emerge? A possible and overlooked explanation. *Science Education*, 100(5), 791–813. <https://doi.org/10.1002/sce.21231>
- Arbeitsgruppe MINT-EC-Zertifikat. (2020). *Das MINT-EC-Zertifikat. Die Würdigung besonderer Leistungen im MINT-Bereich*. Zugriff am 24.09.2021 unter https://www.mint-ec.de/fileadmin/mediathek/MINT-EC-Zertifikat_Handbuch_2020_fotolia_.pdf
- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Hrsg.), *Advanced structural equation modeling* (S. 243–277). Mahwah, NJ: Lawrence Erlbaum Publishers.
- Arens, A. K. (2017). Schulisches Selbstkonzept: Struktur, Entwicklung und Einfluss auf Lernen und Leistung. In U. Hartmann, M. Hasselhorn & A. Gold (Hrsg.), *Entwicklungsverläufe verstehen – Kinder mit Bildungsrisiken wirksam fördern. Forschungsergebnisse des Frankfurter IDeA-Zentrums* (S. 240–253). Stuttgart: Kohlhammer.
- Arens, A. K. (2021). Wertfacetten im Grundschulalter in drei Fächern: Differenzierung, Entwicklung, Geschlechtseffekte und Zusammenhänge zu Noten. *Zeitschrift für Pädagogische Psychologie*, 35(1), 32–52. <https://doi.org/10.1024/1010-0652/a000257>
- Arens, A. K., Schmidt, I. & Preckel, F. (2019). Longitudinal relations among self-concept, intrinsic value and attainment value across secondary school years in three academic domains. *Journal of Educational Psychology*, 111(4), 663–684. <https://doi.org/10.1037/edu0000313>
- Arens, A. K. & Watermann, R. (2015). How an early transition to high-ability secondary schools affects students' academic self-concept. Contrast effects, assimilation effects and differential stability. *Learning and Individual Differences*, 37, 64–71. <https://doi.org/10.1016/j.lindif.2014.11.007>
- Arens, A. K., Yeung, A. S., Craven, R. G., Watermann, R. & Hasselhorn, M. (2013). Does the timing of transition matter? Comparison of German students' self-perceptions before and after transition to secondary school. *International Journal of Educational Research*, 57, 1–11. <https://doi.org/10.1016/j.ijer.2012.11.001>

- Artelt, C., Blossfeld, H.-P., Faust, G., Roßbach, H.-G. & Weinert, S. (2013). *Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS-8-14)*. Bamberg: Institut zur Qualitätsentwicklung im Bildungswesen. https://doi.org/10.5159/IOB_BIKS_8_14_v2
- Asparouhov, T. & Muthén, B. O. (2006). *Robust chi square difference testing with mean and variance adjusted test statistics (Mplus web notes: No. 10)*. Zugriff am 24.01.2021 unter <https://www.statmodel.com/download/webnotes/webnote10.pdf>
- Asparouhov, T. & Muthén, B. O. (2010). *Weighted least squares estimation with missing data*. Zugriff am 13.02.2021 unter <https://www.statmodel.com/download/GstrucMissingRevision.pdf>
- Atkinson, J. W. (1957). Motivational determinants of risk taking behavior. *Psychology Review*, 64(6), 359–372. <https://doi.org/10.1037/h0043445>
- Atkinson, J. W. & Lens, W. (1980). Fähigkeit und Motivation als Determinanten momentaner und kumulativer Leistung. In H. Heckhausen (Hrsg.), *Fähigkeit und Motivation in erwartungswidriger Schulleistung* (S. 129–192). Göttingen: Hogrefe.
- Avenarius, H. (1994). Schulische Selbstverwaltung – Grenzen und Möglichkeiten. *Recht der Jugend und des Bildungswesens*, 42(2), 256–269.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2018). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (15. Aufl.). Heidelberg: Springer-Verlag. <https://doi.org/10.1007/978-3-662-56655-8>
- Bähr, J. (2005). Klassenmusizieren. In W. Jank (Hrsg.), *Musik-Didaktik. Praxishandbuch für die Sekundarstufe I und II* (S. 159–167). Berlin: Cornelsen.
- Bähr, J., Jank, W. & Schwab, C. (2018). Musikunterricht und Ensemblespiel im Rahmen der Kooperation von allgemein bildender Schule und Musikschule. In R.-D. Kraemer & W. Rüdiger (Hrsg.), *Ensemblespiel und Klassenmusizieren. Ein Handbuch für die Praxis in Schule und Musikschule* (S. 131–152). Augsburg: Wißner-Verlag.
- Bähr, J. & Schwab, C. (2002). Vom Modellversuch zum Projekt: Kooperation von Schule und Musikschule in Hessen. In S. Helms (Hrsg.), *Musik im Diskurs. Allgemein bildende Schule und Musikschule in europäischen Ländern* (S. 67–78). Kassel: Gustav Bosse Verlag.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bardone, L. & Guio, A.-C. (2005). In-work poverty. New commonly agreed indicators at EU level. *Statistics in focus*, 5, 1–11.
- Bastian, H. G. (2000). *Musik(erziehung) und ihre Wirkung. Eine Langzeitstudie an Berliner Grundschulen*. Mainz: Schott Musik International.
- Baumert, J., Gruehn, S., Heyn, S., Köller, O. & Schnabel, K. (2003). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). Dokumentation: Bd. 1. Skalen Längsschnitt 1. Wellen 1–4*. Berlin: Max-Planck-Institut für Bildungsforschung.

- Baumert, J. & Köller, O. (1998). Interest research concerning secondary level I: An overview. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon conference on interest and gender* (S. 241–256). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Baumert, J., Roeder, P. M., Sang, F. & Schmitz, B. (1986). Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen. *Zeitschrift für Pädagogik*, 32(5), 639–660. <https://doi.org/10.25656/01:14408>
- Beauducel, A. & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2
- Becker, D., Willems, A. S. & Bos, W. (2014). Kalter Fischteich oder warmes Sonnenbad? Kontrolle von Mess- und Aggregationsfehlern bei der Schätzung von Klassenkontexteffekten auf das akademische Selbstkonzept in den Naturwissenschaften. In H. G. Holtappels, A. S. Willems, M. Pfeiffer, W. Bos & N. McElvany (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (S. 321–355). Weinheim: Beltz Juventa.
- Belfield, C. R., Levin, H. M. (2002). The effects of competition between schools on educational outcomes: A review for the United States. *Review of Educational Research*, 72(2), 279–341. <https://doi.org/10.3102/00346543072002279>
- Bellmann, J. & Weiß, M. (2009). Risiken und Nebenwirkungen Neuer Steuerung im Schulsystem. *Zeitschrift für Pädagogik*, 55(2), 286–308. <https://doi.org/10.25656/01:4251>
- Bennett, J., Lubben, F. & Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education*, 91(3), 347–370. <https://doi.org/10.1002/sce.20186>
- Bennetts, K. S. (2013). Boys' music? School context and middle-school boys' musical choices. *Music Education Research*, 15(2), 214–230. <https://doi.org/10.1080/14613808.2012.759550>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Bernecker, C., Haag, L. & Pfeiffer, W. (2006). Musikalisches Selbstkonzept – eine empirische Untersuchung. *Diskussion Musikpädagogik*, 29, 53–56.
- Biedermann, H., Weber, C., Herzog-Punzenberger, B. & Nagel, A. (2016). Auf die Mitschüler/innen kommt es an? Schulische Segregation – Effekte der Schul- und Klassenzusammensetzung in der Primarstufe und der Sekundarstufe I. In M. Bruneforth, F. Eder, K. Krainer, C. Schreiner, A. Seel & C. Spiel (Hrsg.), *Nationaler Bildungsbericht Österreich 2015* (S. 133–174). Graz: Leykam. <https://doi.org/10.17888/nbb2015-2-4>

- Blankenburg, J. S., Höffler, T. N. & Parchmann, I. (2015). Design und Evaluation eines naturwissenschaftlichen Wettbewerbstages. In S. Bernholt (Hrsg.), *Heterogenität und Diversität – Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht* (S. 76–78). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Blossfeld, H.-P., Bos, W., Daniel, H. D., Hannover, B., Lenzen, D., Prenzel, M. & Wöbmann, L. (2010). Aktionsrat Bildung. Bildungsautonomie: Zwischen Regulierung und Eigenverantwortung – die Bundesländer im Vergleich. Expertenrating der Schul- und Hochschulgesetzte der Länder zum Jahresgutachten 2010. In vbw – Vereinigung der Bayerischen Wirtschaft e.V. (Hrsg.), *Bildungsautonomie: Zwischen Regulierung und Eigenverantwortung* (S. 1–171). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task-value and achievement goals. *Journal of Educational Psychology*, 93(1), 23–34. <https://doi.org/10.1037/0022-0663.93.1.23>
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. überarb. Aufl.). Heidelberg: Springer Medizin.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (S. 183–198). Göttingen: Schwartz.
- Brandt, A., Möller, J. & Kohse-Höinghaus, K. (2008). Was bewirken außerschulische Experimentierlabors? Ein Kontrollgruppenexperiment mit Follow up-Erhebung zu Effekten auf Selbstkonzept und Interesse. *Zeitschrift für Pädagogische Psychologie*, 22(1), 5–12. <https://doi.org/10.1024/1010-0652.22.1.5>
- Brandstätter, V., Schüler, J., Puca, R. M. & Lozo, L. (2013). *Motivation und Emotion. Allgemeine Psychologie für Bachelor*. Heidelberg: Springer Medizin. <https://doi.org/10.1007/978-3-642-30150-6>
- Braund, M. & Reiss, M. J. (2006). Towards a more authentic science curriculum: The contribution of out-of-school learning. *International Journal of Science Education*, 28(12), 1373–1388. <https://doi.org/10.1080/09500690500498419>
- Brown, T. A. (2015). *Confirmatory factor analysis* (2. Aufl.). New York: Guilford.
- Buchborn, T. (2011). *Neue Musik im Musikunterricht mit Blasinstrumenten* (Detmolder Hochschulschriften, Bd. 6). Essen: Die Blaue Eule.
- Budde, J., Kansteiner, K. & Bossen, A. (2016). *Zwischen Differenz und Differenzierung*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-02698-1>
- Buddeberg, M. & Hußmann, A. (2021). Verlängerte Unterrichtsstunden aus Sicht von Lehrkräften sowie Schülerinnen und Schülern. In M. Haas & M. Krieg (Hrsg.), *Lernen gestalten. Beiträge der empirischen Forschung zu Innovationen in Schule und Unterricht* (S. 15–28). Göttingen: Universitätsverlag. <https://doi.org/10.17875/gup2021-1578>

- Busch, T. (2013). „Was glaubst Du, kannst Du in Musik?“ *Musikalische Selbstwirksamkeitserwartungen in ihre Entwicklung zu Beginn der Sekundarstufe I*. Münster: LIT Verlag.
- Busch, T. & Kranefeld, U. (2013). Selbstüberzeugungen und Programmteilnahme: Wie entwickelt sich das musikalische Selbstkonzept im Programm „Jedem Kind ein Instrument“?. *Beiträge empirischer Musikpädagogik*, 4(2), 1–27.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Chen, P.-Y., Wu, W., Garnier-Villareal, M., Kite, B. A. & Jia, F. (2020). Testing measurement invariance with ordinal missing data: A comparison of estimators and missing data techniques. *Multivariate Behavioral Research*, 55(1), 87–101. <https://doi.org/10.1080/00273171.2019.1608799>
- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5
- Christ, O. & Schlüter, E. (2012). *Strukturgleichungsmodelle mit Mplus. Eine praktische Einführung*. München: Oldenbourg Verlag. <https://doi.org/10.1524/9783486714807>
- Clausen, M. (2006). Warum wählen Sie genau diese Schule? Eine inhaltsanalytische Untersuchung elterlicher Begründungen der Wahl der Einzelschule innerhalb eines Bildungsgangs. *Zeitschrift für Pädagogik*, 52(1), 69–89. <https://doi.org/10.25656/01:4446>
- Clausen, M. (2007). *Einzelschulwahl. Zur Wahl der konkreten weiterführenden Einzelschule aus der Sicht von Bildungsnachfragenden und Bildungsanbietenden* (Unveröffentlichte Habilitationsschrift). Universität Mannheim, Mannheim.
- Collins, L. M. & Lanza, S. T. (2013). *Latent class and latent transition analysis. With applications in the social, behavioral and health sciences*. Hoboken: Wiley.
- Csikszentmihalyi, M. (2010). *Das Flow-Erlebnis. Jenseits von Angst und Langeweile: Im Tun aufgehen*. Stuttgart: Klett-Cotta.
- Daniels, Z. (2008). *Entwicklung schulischer Interessen im Jugendalter*. Münster: Waxmann.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238. <https://doi.org/10.25656/01:11173>
- Denissen, J. J. A., Zarett, N. R. & Eccles, J. S. (2007). I like to do it, I'm able and I know I am: Longitudinal couplings between domain-specific achievement, self-concept and interest. *Child Development*, 78(2), 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>

- Dewey, J. (1913). *Interest and effort in education*. Boston: Riverside Press.
- Dickhäuser, O., Janke, S., Praetorius, A.-K. & Dresel, M. (2017). The effects of teachers' reference norm orientations on students' implicit theories and academic self-concepts. *Zeitschrift für Pädagogische Psychologie*, 31(3–4), 205–219. <https://doi.org/10.1024/1010-0652/a000208>
- Dinkelmann, I. & Buff, A. (2016). Children's and parents' perceptions of parental support and their effects on children's achievement motivation and achievement in mathematics. A longitudinal predictive mediation model. *Learning and Individual Differences*, 50, 122–132. <https://doi.org/10.1016/j.lindif.2016.06.029>
- Dolenc, N. R. & Kazanis, W. H. (2020). A potential for interest driven learning to enhance the inquiry based learning process. *Science Educator*, 27(2), 121–128.
- Eccles, J. S. (2005). Subjective task value and the Eccles et al. model of achievement related choices. In A. J. Elliot & C. S. Dweck (Hrsg.), *Handbook of competence and motivation* (S. 105–121). New York: Guilford Press.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives* (S. 74–146). San Francisco: W. H. Freeman.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C. & Mac Iver, D. (1993). Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and families. *American Psychologist*, 48(2), 90–101. <https://doi.org/10.1037/0003-066X.48.2.90>
- Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values and goals. *Annual Review of Psychology*, 53(1), 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Eccles, J. S., Wigfield, A., Harold, R. D. & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, 64(3), 830–847. <https://doi.org/10.2307/1131221>
- Eder, F. (2011). Wie gut sind Musikhauptschulen? In H. Altrichter, M. Heinrich & K. Soukup-Altrichter (Hrsg.), *Schulentwicklung durch Schulprofilierung? Zur Veränderung von Koordinationsmechanismen im Schulsystem* (S. 165–193). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92825-8_5
- Ehrtmann, L., Wolter, I. & Hannover, B. (2019). The interrelatedness of gender-stereotypical interest profiles and students' gender-role orientation, gender and reasoning abilities. *Frontiers in Psychology*, 10, 1–15. <https://doi.org/10.3389/fpsyg.2019.01402>
- Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.
- Engeln, K. & Euler, M. (2004). Forschen statt Pauken. Aktives Lernen im Schülerlabor. *Physik Journal*, 3(11), 45–48.

- Erikson, E. H. (1959). *Identity and the life circle*. New York: Norton.
- Euler, M., Schüttler, T. & Hausamann, D. (2015). Schülerlabore: Lernen durch Forschen und Entwickeln. In E. Kircher, R. Girwidz & P. Häußler (Hrsg.), *Physikdidaktik. Theorie und Praxis* (S. 759–782). Berlin: Springer Spektrum.
- Eurydice. (2007). *Schulautonomie in Europa. Strategien und Maßnahmen*. Zugriff am 08.06.2021 unter <https://op.europa.eu/de/publication-detail/-/publication/102bb131-8105-4599-9367-377946471af3/language-de>
- Ferdinand, H. (2014). *Entwicklung von Fachinteresse. Längsschnittstudie zu Interessenverläufen und Determinanten positiver Entwicklung in der Schule*. Münster: Waxmann.
- Feuerborn, W., Hollenstein, M., Janssen, W., Kürner, G. & Staub, P. (2002). *Essential Elements – Lehrerhandbuch*. Hagedorn: DeHaske.
- Fiedler, D. & Hasselhorn, J. (2020). Zum Zusammenhang von musikalischem Selbstkonzept und Motivation im Musikunterricht. Eine empirische Querschnittsuntersuchung von Schülerinnen und Schülern der 7. bis 11. Klassenstufe an allgemeinbildenden Schulen in Baden-Württemberg und Nordrhein-Westfalen. *Beiträge empirischer Musikpädagogik*, 11, 1–34.
- Fredricks, J. A. & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, 38(4), 519–533. <https://doi.org/10.1037//0012-1649.38.4.519>
- Frenzel, A. C., Goetz, T., Pekrun, R. & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family and school context. *Journal of Research on Adolescence*, 20(2), 507–537. <https://doi.org/10.1111/j.1532-7795.2010.00645.x>
- Fritzsche, E. S., Kröner, S. & Pfeiffer, W. (2011). Chorknaben und andere Gymnasiasten – Determinanten musikalischer Aktivitäten an Gymnasien mit unterschiedlichen Schulprofilen. *Journal for Educational Research Online*, 3(2), 94–118.
- Fritzsche, E. S., Pfeiffer, W. & Kröner, S. (2015). Musical elite gymnasias as learning environments and settings for personality development in secondary students? The case of musical self-concept. *Psychological Test and Assessment Modeling*, 57(2), 236–261.
- Gardner, P. L. (1985). Students' attitudes to science and technology: An international overview. In M. Lehrke, L. Hoffmann & P. L. Gardner (Hrsg.), *Interests in science and technology education* (S. 15–34). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B. M., Trautwein, U. & Nagengast, B. (2016). Side effects of motivational interventions? Effects of an intervention in math classrooms on motivation in verbal domains. *AERA Open*, 2(2), 1–14. <https://doi.org/10.1177/2332858416649168>

- Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U. & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663–677. <https://doi.org/10.1037/edu0000003>
- Gaspard, H., Hasselhorn, M., Nagengast, B. & Trautwein, U. (2019). Motivation und Volition im Schulalter: Einführung und Überblick. In H. Gaspard, U. Trautwein & M. Hasselhorn (Hrsg.), *Diagnostik und Förderung von Motivation und Volition* (S. 1–18). Göttingen: Hogrefe. <https://doi.org/10.1026/03001-000>
- Gaspard, H., Lauermaun, F., Rose, N., Wigfield, A. & Eccles, J. S. (2019). Cross-domain trajectories of students' ability self-concepts and intrinsic values in math and language arts. *Child Development, 91*(5), 1800–1818. <https://doi.org/10.1111/cdev.13343>
- Gaspard, H., Nagengast, B. & Trautwein, U. (2019). Erfassung von Wertüberzeugungen. In H. Gaspard, U. Trautwein & M. Hasselhorn (Hrsg.), *Diagnostik und Förderung von Motivation und Volition* (S. 87–100). Göttingen: Hogrefe. <https://doi.org/10.1026/03001-000>
- Geiser, C. (2010). *Datenanalyse mit MPlus. Eine anwendungsorientierte Einführung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gess, C., Rueß, J. & Deicke, W. (2014). Design-based Research als Ansatz zur Verbesserung der Lehre an Hochschulen – Einführung und Praxisbeispiel. *Qualität in der Wissenschaft, 8*(1), 10–16.
- Gewirtz, S., Ball, S & Bowe, R. (1995). *Markets, choice and equity in education*. Buckingham: Open University Press.
- Gibson, H. L. & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Science Education, 86*(5), 693–705.
- Göllner, M. (2017). *Perspektiven von Lehrenden und SchülerInnen auf Bläserklassenunterricht. Eine qualitative Interviewstudie* (Perspektiven musikpädagogischer Forschung, Bd. 6). Münster: Waxmann.
- Göllner, M. & Niessen, A. (2015). Ansätze von Öffnung im Musikklassenunterricht in der Wahrnehmung von Lehrenden und Schülern. Eine Fallstudie auf Basis qualitativer Interviews. *Beiträge empirischer Musikpädagogik, 6*(2), 2–30.
- Gottfried, A. E., Fleming, J. S. & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*(1), 3–13. <https://doi.org/10.1037/0022-0663.93.1.3>
- Green, L. (2002). Exposing the gendered discourse of music education. *Feminism and Psychology, 12*(2), 137–144. <https://doi.org/10.1177/0959353502012002003>

- Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S. & Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: Dimensional comparison processes and expectancy-by-value interactions. *Learning and Instruction*, 49, 81–91. <https://doi.org/10.1016/j.learninstruc.2016.12.007>
- Gymnasium Alfeld. (2021). *Profilklasse Junge Wissenschaftler*. Zugriff am 19.08.2021 unter <https://www.gymalfeld.de/index.php/component/jdownloads/?task=download.send&id=69&catid=32&m=0&Itemid=435>
- Gymnasium Martino Katharineum. (2017). *Schulprogramm*. Zugriff am 19.08.2021 unter https://mk-braunschweig.de/fileadmin/Unsere_Schule/Wir_ueber_uns/Schulprogramm_2017_2.pdf
- Gymnasium Stolzenau. (2020a). *Fachbereich Musik. Bläserklasse und Musicklasse in den Jahrgängen 5 und 6 am Gymnasium Stolzenau*. Zugriff am 19.08.2021 unter <https://www.gymnasium-stolzenau.de/wp-content/uploads/2020/05/Informationen-zur-BI%C3%A4ser-und-zur-Musicklasse.pdf>
- Gymnasium Stolzenau. (2020b). *Informationen zum besonderen Unterrichtsangebot „Entdeckerklasse“*. Zugriff am 19.08.2021 unter <https://www.gymnasium-stolzenau.de/wp-content/uploads/2020/05/Informationen-zur-Entdeckerklasse.pdf>
- Haas, M., Nonte, S., Krieg, M. & Stubbe, T. C. (2019). Unterrichtsqualität in Musikklassen. Befunde aus der quasi-experimentellen Studie ProBiNi. In V. Weidner & C. Rolle (Hrsg.), *Praxen und Diskurse aus Sicht musikpädagogischer Forschung* (Musikpädagogische Forschung, Bd. 40, 1. Aufl., S. 137–154). Münster: Waxmann.
- Haas, M. & Stubbe, T. C. (2017, September). Soziale Selektion durch Klassenprofilierung. In S. Nonte & T. C. Stubbe (Chair), *Wer entscheidet sich für eine Profilklasse? Vertiefende Analysen zu ProBiNi*. Symposium durchgeführt auf der 82. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF), Tübingen.
- Habig, S., Blankenburg, J., van Vorst, H., Fechner, S., Parchmann, I. & Sumfleth, E. (2018). Context characteristics and their effects on students' situational interest in chemistry. *International Journal of Science Education*, 40(10), 1154–1175. <https://doi.org/10.1080/09500693.2018.1470349>
- Hagenaars, J. A. & McCutcheon, A. L. (Hrsg.). (2002). *Applied latent class analysis*. Cambridge: Cambridge University Press.
- Hall, C. (2005). Gender and boys' singing in early childhood. *British Journal of Music Education*, 22(1), 5–20. <https://doi.org/10.1017/S0265051704005960>
- Hannover, B. & Kessels, U. (2002). Challenge the science-stereotype! In M. Prenzel & J. Doll (Hrsg.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (S. 341–358). Weinheim: Beltz.

- Harackiewicz, J. M. & Knogler, M. (2017). Interest. Theory and application. In A. J. Elliot, C. S. Dweck & D. S. Yeager (Hrsg.), *Handbook of competence and motivation* (S. 334–352). New York: The Guilford Press.
- Harrison, S. D. (2005). Gender in music education. Does it really matter? *Music in Action for Australian Educators*, 3(1), 8–10.
- Harrison, S. D. (2007). A perennial problem in gendered participation in music: What's happening to the boys? *British Journal of Music Education*, 24(3), 267–280. <https://doi.org/10.1017/S0265051707007577>
- Harter, S. (1998). The development of self-representations. In W. Damon & N. Eisenberg (Hrsg.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5. Aufl., S. 553–617). New York: Wiley.
- Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.
- Hasse, R. (2015). Organisationssoziologische Bildungsanalysen. Eine Situations-einschätzung. In R. Leemann, C. Imdorf, J. Powell & M. Sertl (Hrsg.), *Wie Bildung organisiert wird. Bildungssoziologische Beiträge* (S. 47–67). Weinheim: Beltz Juventa.
- Haupt, O. J. & Hempelmann, R. (2013). Schülerlabor Kategorien: Das Klassische Schülerlabor. *LeLa magazin*, 5, 2–4.
- Häußler, P. (1987). Measuring students' interest in physics – design and results of a cross-sectional study in the Federal Republic of Germany. *International Journal of Science Education*, 9(1), 79–92. <https://doi.org/10.1080/0950069870090109>
- Häußler, P., Frey, K., Hoffmann, L., Rost, J. & Spada, H. (1980). *Physikalische Bildung: Eine curriculare Delphi-Studie, Teil 1. IPN-Arbeitsberichte 41*. Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Häußler, P., Frey, K., Hoffmann, L., Rost, J. & Spada, H. (1988). *Physikalische Bildung für heute und morgen. IPN-Arbeitsberichte 116*. Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Häußler, P. & Hoffmann, L. (1995). Physikunterricht – an den Interessen von Mädchen und Jungen orientiert. *Unterrichtswissenschaft*, 23(2), 107–126. <https://doi.org/10.25656/01:8124>
- Häußler, P. & Hoffmann, L. (1998). Qualitative differences in student's interest in physics and dependence on age and gender. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon conference on interest and gender* (S. 280–189). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- He, J., Buchholz, J. & Klieme, E. (2017). Effects of anchoring vignettes on comparability and predictive validity of student self-reports in 64 cultures. *Journal of Cross-Cultural Psychology*, 48(3), 319–334. <https://doi.org/10.1177/0022022116687395>

- Heckhausen, H. (1989). *Motivation und Handeln* (2. Aufl.). Berlin: Springer-Verlag.
- Heinrich, M., Altrichter, H. & Soukup-Altrichter, K. (2011). Neue Ungleichheiten durch Schulprofilierung? Autonomie, Wettbewerb und Selektion in profilorientierten Schulentwicklungsprozessen. In F. Dietrich, M. Heinrich & N. Thieme (Hrsg.), *Neue Steuerung – alte Ungleichheit? Steuerung und Entwicklung im Bildungssystem* (S. 271–290). Münster: Waxmann.
- Helmke, A. (1993). Die Entwicklung der Lernfreude vom Kindergarten bis zur 5. Klassenstufe. *Zeitschrift für Pädagogische Psychologie*, 7, 77–86.
- Helsper, W., Dreier, L., Gibson, A., Kotzyba, K. & Niemann, M. (2015). „Exklusive“ und private Gymnasien in städtischen Bildungsregionen. Wettbewerb und Schülersauswahl am städtischen höheren „Bildungsmarkt“. In M. Kraul (Hrsg.), *Private Schulen* (Schule und Gesellschaft, Bd. 58, S. 45–61). Berlin: Springer VS. https://doi.org/10.1007/978-3-658-07735-8_3
- Helsper, W., Dreier, L., Gibson, A. & Niemann, M. (2015). Auswahlverfahren an „exklusiven“ Gymnasien. Auswahlpraktiken und ihre Legitimation. In W. Helsper & H.-H. Krüger (Hrsg.), *Auswahl der Bildungsklientel. Zur Herstellung von Selektivität in „exklusiven“ Bildungsinstitutionen* (S. 135–164). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-09375-4_6
- Helsper, W., Kramer, R.-T. & Thiersch, S. (2014). Habitus – Schule – Schüler: Eine Einleitung. In W. Helsper, R.-T. Kramer & S. Thiersch (Hrsg.), *Schülerhabitus. Theoretische und empirische Analysen zum Bourdieuschen Theorem der kulturellen Passung* (S. 7–29). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-00495-8_1
- Helsper, W. & Krüger, H.-H. (2015). Auswahlverfahren in Bildungsinstitutionen – eine Einleitung. In W. Helsper & H.-H. Krüger (Hrsg.), *Auswahl der Bildungsklientel. Zur Herstellung von Selektivität in „exklusiven“ Bildungsinstitutionen* (S. 9–27). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-09375-4_1
- Herbart, J. F. (1806). Allgemeine Pädagogik, aus dem Zweck der Erziehung abgeleitet. In J. F. Herbart (Hrsg.), *Pädagogische Schriften* (Bd. 2). Düsseldorf: Küpper.
- Heß, C. (2017a). *Konzeptionelle Spannungsfelder des Klassenmusizierens mit Blasinstrumenten. Eine Analyse divergenter Prämissen und Zielvorstellungen*. Münster: Waxmann.
- Heß, F. (2005). Pauken statt „pauken“. Überlegungen zur Nachhaltigkeit musikbezogenen Lernens durch Klassenmusizieren. In F. Heß (Hrsg.), *Zugänge zur Musik. Formen des Musik-Lernens von der Kindheit bis ins Alter* (S. 81–98). Kassel: Gustav Bosse.
- Heß, F. (2011a). Musikunterricht zwischen Sach- und Fachinteresse. Ergebnisse aus der Pilotstudie Musikunterricht aus Schülersicht. *Beiträge empirischer Musikpädagogik*, 2(1), 1–26.

- Heß, F. (2011b). *Skalenhandbuch zur Studie „Musikunterricht aus Schülersicht“ (MASS 2011). Dokumentation der Erhebungsinstrumente*. Zugriff am 20.03.2021 unter www.uni-kassel.de/fb01/fileadmin/datas/fb01/Institut_fuer_Musik/Daten/Frauke_He%C3%9F__Mass_Forschungsprojekt/Skalenhandbuch_Mass_2011.pdf
- Heß, F. (2017b). *Gendersensibler Musikunterricht. Empirische Studien und didaktische Konsequenzen*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-19166-5>
- Hidi, S. & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, 70(2), 151–179. <https://doi.org/10.2307/1170660>
- Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
- Hidi, S., Renninger, K. A. & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Hrsg.), *Motivation, emotion and cognition: Integrative perspective on intellectual functioning and development* (S. 89–115). Hillsdale: Erlbaum.
- Ho, D., Imai, K., King, G. & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1–28. <https://doi.org/10.18637/jss.v042.i08>
- Höffler, T. N., Lüthjohann, F. & Parchmann, I. (2014). Welche Wirkung erzielt ein naturwissenschaftlicher Anfangsunterricht? Befunde einer Untersuchung zum Projekt NaWi-aktiv im Zuge der Einführung des Nawi-Unterrichts an nicht-gymnasialen Schulen in Schleswig-Holstein. *Zeitschrift für die Didaktik der Naturwissenschaften*, 20(1), 87–99. <https://doi.org/10.1007/s40573-014-0009-1>
- Hoffmann, L. (2002). Promoting girls' interest and achievement in physics classes for beginners. *Learning and Instruction*, 12(4), 447–465. [https://doi.org/10.1016/S0959-4752\(01\)00010-X](https://doi.org/10.1016/S0959-4752(01)00010-X)
- Hoffmann, L., Häußler, P. & Lehrke, M. (1998). *Die IPN-Interessenstudie Physik*. Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Hoffmann, L. & Lehrke, M. (1986). Eine Untersuchung über Schülerinteressen an Physik und Technik. *Zeitschrift für Pädagogik*, 32(2), 189–204.
- Holland, J. L. (1973). *Making vocational choices. A theory of careers*. Englewood Cliffs, NJ: Prentice-Hall.
- Holmeier, M. (2017). Evaluation MINT-Klasse – Thesen zur MINT-Nachwuchsförderung. In C. Maurer (Hrsg.), *Qualitätvoller Chemie- und Physikunterricht – normative und empirische Dimensionen* (S. 396–399). Regensburg: Universität Regensburg.

- Höttecke, D. (2012). Forschend-entdeckenden Unterricht authentisch gestalten. Ein Problemaufriss. In S. Bernholt (Hrsg.), *Inquiry-based learning – Forschendes Lernen* (S. 32–45). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling. A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Huber, L. (2009). Warum Forschendes Lernen nötig und möglich ist. In L. Huber, J. Hellmer & F. Schneider (Hrsg.), *Forschendes Lernen im Studium: Aktuelle Konzepte und Erfahrungen* (S. 9–35). Bielefeld: Universitätsverlag Webler.
- Huber, L. (2014). *Forschungsbasiertes, Forschungsorientiertes, Forschendes Lernen: Alles das selbe? Ein Plädoyer für eine Verständigung über Begriffe und Unterscheidungen im Feld forschungsnahen Lehrens und Lernens*. Zugriff am 22.08.2021 unter https://www.fh-potsdam.de/fileadmin/user_upload/forschen/material-publication/HSW1_2_2014_Huber.pdf
- IBM Corp. (2021). *IBM SPSS Statistics for Windows, Version 27* [Computer Software]. Armonk, NY: IBM Corp.
- IGS Roderbruch. (2020). *Bläserklasse auf einen Blick*. Zugriff am 19.08.2021 unter https://www.igs-roderbruch.de/images/fachbereiche/musik/Flyer_BI%C3%A4serklasse_IGS_Roderbruch_2020.pdf
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. E. & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73(2), 509–527. <https://doi.org/10.1111/1467-8624.00421>
- Jank, W. & Stroh, W. M. (2006). Aufbauender Musikunterricht – Königsweg oder Sackgasse? In W. Pfeiffer & J. Terhag (Hrsg.), *Schülerorientierter Musikunterricht – Wunsch und Wirklichkeit* (S. 52–64). Oldershausen: Lugert Verlag.
- Jia, F. & Wu, W. (2019). Evaluating methods for handling missing ordinal data in structural equation modeling. *Behavior Research Methods*, 51, 2337–2355. <https://doi.org/10.3758/s13428-018-1187-4>
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), 409–426.
- Jöreskog, K. G. (1993). Latent variable modeling with ordinal variables. In K. Haagen, D. J. Bartholomew & M. Deistler (Hrsg.), *Statistical modeling and latent variables* (S. 163–171). Amsterdam: Elsevier.
- Jünger, H. (2013). *Auf dem Wege zu musikalischer Tätigkeit*. Zugriff am 26.09.2021 unter http://www.epb.uni-hamburg.de/erzwiss/juenger/download/juenger_2013_taetigkeit.pdf

- Juraneck, M. (2019). Eigenverantwortlich – selbstständig – autonom. Was wir aus einem Rechtsvergleich hinsichtlich schulischer Entscheidungsfreiräume lernen können. In E. Rauscher, C. Wiesner, D. Paasch & P. Heißenberger (Hrsg.), *Schulautonomie – Perspektiven in Europa. Befunde aus dem EU-Projekt INNOVITAS* (S. 21–67). Münster: Waxmann.
- Jurczok, A. (2019). *Schulwahl unter „gleichwertigen“ Einzelschulen. Elterliche Übergangsentscheidungen im zweigliedrigen Sekundarschulsystem*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-26199-3>
- Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to linking for school subjects. *European Journal of Psychology of Education*, 20(3), 309–323. <https://doi.org/10.1007/BF03173559>
- Kessels, U. & Hannover, B. (2004). Empfundene „Selbstnähe“ als Mediator zwischen Fähigkeitsselbstkonzept und Leistungskurswahlintention. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 36(3), 130–138. <https://doi.org/10.1026/0049-8637.36.3.130>
- Keßler, C. I., Nonte, S., Krieg, M. & Brauns, J. (2022). Schul- und Klassenprofilierung aus schulentwicklungstheoretischer Betrachtung. In T. C. Stubbe, S. Nonte, M. Haas & M. Krieg (Hrsg.), *Musik- und MINT-Profile an Niedersächsischen Gymnasien und integrierten Gesamtschulen. Ergebnisse der Studie ProBiNi*. Münster: Waxmann. Manuskript in Vorbereitung.
- Kim, E. S. & Yoon, M. (2011). Testing measurement invariance: A comparison of multiple-group categorical CFA and IRT. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(2), 212–228. <https://doi.org/10.1080/10705511.2011.557337>
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237. <https://doi.org/10.25656/01:4348>
- Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3. Aufl.). New York: Guilford.
- Klose, P. (2018). Musik als Praxis. Ein anderer Weg, Musikunterricht und Bläserklassen zu denken. In I. Siedenburg & G. Harbig (Hrsg.), *Kreatives Musizieren in der Bläserklasse. Wege zum Improvisieren und Komponieren zwischen Jazz und experimentellem Klang* (S. 9–29). Münster: Waxmann.
- Köhler, C., Bernholt, S., Höffler, T. & Parchmann, I. (2015). Charakteristika von Wettbewerbsteilnehmenden. In S. Bernholt (Hrsg.), *Heterogenität und Diversität – Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht* (S. 67–69). Kiel: Institut für Pädagogik der Naturwissenschaften und Mathematik (IPN).

- Köller, O., Baumert, J. & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, 32(5), 448–470. <https://doi.org/10.2307/749801>
- Köller, O., Meyer, J., Saß, S. & Baumert, J. (2019). New analyses of an old topic: Effects of intelligence and motivation on academic achievement. *Journal for Educational Research Online*, 11(1), 166–189. <https://doi.org/10.25656/01:16792>
- Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie*, 20(1), 27–39. <https://doi.org/10.1024/1010-0652.20.12.27>
- Koivuhovi, S. (2021). *Studying in a class with special emphasis. Changes in children's competence beliefs and mathematical thinking skills and the role of social comparison.* (Dissertation, Universität Helsinki). Zugriff am 23.09.2021 unter https://helda.helsinki.fi/bitstream/handle/10138/330292/koivuhovi_satu_dissertation_2021.pdf?sequence=3&isAllowed=y
- Koivuhovi, S., Vainikainen, M.-P. & Kalalahti, M. (2020). The effect of studying in selective classes on the change of pupils' action-control beliefs during lower secondary school in Finland. *Scandinavian Journal of Educational Research*, 66(1), 1–14. <https://doi.org/10.1080/00313831.2020.1833246>
- Koivuhovi, S., Vainikainen, M.-P. & Kalalahti, M. (2021). Oppilaiden matemaattisten ajattelutaitojen ja matematiikkaminäkäsityksen kehitys painotetun opetuksen ja yleisopetuksen luokilla neljänneltä luokalta kuudennelle [Development of pupils' mathematical thinking skills and mathematical self-concept in classes with and without a special emphasis from fourth to sixth grade]. *Kasvatus*, 52(1), 22–36.
- Kosunen, S., Bernelius, V., Seppänen, P. & Porkka, M. (2020). School choice to lower secondary schools and mechanisms of segregation in urban Finland. *Urban Education*, 55(10), 1416–1488. <https://doi.org/10.1177/0042085916666933>
- Kranefeld, U. (Hrsg.). (2015). *Instrumentalunterricht in der Grundschule. Prozess- und Wirkungsanalysen zum Programm Jedem Kind ein Instrument.* Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Kranefeld, U., Busch, T. & Dücker, J. (2015). BEGIn – Instrumentaler Gruppenunterricht in der Grundschule: Teilnahme, Selbstkonzepte, Gestaltungsmuster. In U. Kranefeld (Hrsg.), *Instrumentalunterricht in der Grundschule. Prozess- und Wirkungsanalysen zum Programm Jedem Kind ein Instrument.* (S. 49–89). Berlin: Bundesministerium für Bildung und Forschung (BMBF).

- Krapp, A. (1992a). Das Interessenskonstrukt. Bestimmungsmerkmale der Interessenhandlung und des individuellen Interesses aus der Sicht einer Person-Gegenstands-Konzeption. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze der pädagogisch-psychologischen Interessenforschung* (S. 297–329). Münster: Aschendorff.
- Krapp, A. (1992b). Konzepte und Forschungsansätze zur Analyse des Zusammenhangs von Interesse, Lernen und Leistung. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze der pädagogisch-psychologischen Interessenforschung* (S. 9–52). Münster: Aschendorff.
- Krapp, A. (1996). Psychologische Bedingungen naturwissenschaftlichen Lernens: Untersuchungsansätze und Befunde zu Motivation und Interesse. In R. Duit & C. von Rhöneck (Hrsg.), *Lernen in den Naturwissenschaften* (S. 37–68). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Krapp, A. (1998). Entwicklung und Förderung von Interessen und Unterricht. *Psychologie in Erziehung und Unterricht*, 45(3), 185–201.
- Krapp, A. (1999). Intrinsische Lernmotivation und Interesse. Forschungsansätze und konzeptuelle Überlegungen. *Zeitschrift für Pädagogik*, 45(3), 387–406. <https://doi.org/10.25656/01:5958>
- Krapp, A. (2001). Interesse. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 286–294). Weinheim: Beltz.
- Krapp, A. (2002). An educational-psychological theory of interest and it's relation to self-determination theory. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 405–427). Rochester: University of Rochester Press.
- Krapp, A. (2007). An educational-psychological conceptualisation of interest. *Journal for Educational and Vocational Guidance*, 7(1), 5–21. <https://doi.org/10.1007/s10775-007-9113-9>
- Krapp, A. (2018). Interesse. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 286–297). Weinheim: Beltz.
- Krapp, A., Hidi, S. & Renninger, K. A. (1992). Interest, learning and development. In K. A. Renninger & S. Hidi (Hrsg.), *The role of interest in learning and development* (S. 3–25). Hillsdale, NJ: Erlbaum.
- Krieg, M., Nonte, S., Haas, M. & Stubbe, T. C. (2022). Welche Rolle spielt das Geschlechtsrollen-Selbstkonzept bei der Wahl einer naturwissenschaftlichen Profilklassen? In S. Hornberg & M. Buddeberg (Hrsg.), *Schule als Ort gesellschaftlicher Teilhabe. Bildungswissenschaftliche Perspektiven, Schulkonzepte und Schulprofile im Fokus* (S. 129–148). Münster: Waxmann.
- Krieg, M. & Stubbe, T. C. (2018, September). *Die Bedeutung der Klassenzusammensetzung von Profilklassen für die Entwicklung von Interessen*. Einzelbeitrag präsentiert auf der 83. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF), Lüneburg.

- Krieg, M., Stubbe, T. C., Tarelli, I. & Willems, A. S. (2021). Zur Interessenstruktur von Grundschulkindern: Eine sekundäranalytische Auswertung der nationalen IGLU 2006 Daten. In M. Haas & M. Krieg (Hrsg.), *Lernen gestalten. Beiträge der empirischen Forschung zu Innovationen in Schule und Universität* (S. 47–74). Göttingen: Universitätsverlag.
- Krüger, J. O. (2015). Auswahl an privaten Grundschulen. Annäherungen an Legitimationspraktiken einer schulischen Elternwahl. In W. Helsper & H.-H. Krüger (Hrsg.), *Auswahl der Bildungsklientel. Zur Herstellung von Selektivität in „exklusiven“ Bildungsinstitutionen* (S. 119–133). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-09375-4_5
- Krüger, H.-H., Keßler, C. I. & Winter, D. (2015). Schulkultur und soziale Ungleichheit. Perspektiven von Schulleitungen an exklusiven Gymnasien auf den Elite- und Exzellenzdiskurs. In J. Böhme, M. Hummrich & R.-T. Kramer (Hrsg.), *Schulkultur. Theoriebildung im Diskurs* (S. 183–210). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-03537-2_9
- Kühnel, S. & Mays, A. (2019). Probleme von Cross-Lagged-Panelmodellen zur Analyse gegenseitiger Beeinflussung von Einstellung und Verhalten. Das Beispiel des Zusammenhangs von politischem Interesse und politischer Beteiligung mit den Daten des SOEP. In J. Mayerl, T. Krause, A. Wahl & M. Wuketich (Hrsg.), *Einstellungen und Verhalten in der empirischen Sozialforschung. Analytische Konzepte, Anwendungen und Analyseverfahren* (S. 359–386). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-16348-8_15
- Kulin, S., Schwippert, K. & Rieckmann, T. (2015). Teilprojekt „Kooperation“ – Rahmen- und Gelingensbedingungen der Zusammenarbeit zwischen Lehrkräften aus Grund- und Musikschule. In U. Kranefeldt (Hrsg.), *Instrumentalunterricht in der Grundschule. Prozess- und Wirkungsanalysen zum Programm Jedem Kind ein Instrument* (S. 200–220). Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Lauermann, F., Meißner, A. & Steinmayr, R. (2020). Relative importance of intelligence and ability self-concept in predicting test performance and school grades in the math and language arts domains. *Journal of Educational Psychology*, 112(2), 364–383. <https://doi.org/10.1037/edu0000377>
- Lauermann, F., Tsai, Y.-M. & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy-value theory of achievement-related behaviors. *Developmental Psychology*, 53(8), 1540–1559. <https://doi.org/10.1037/dev0000367>
- Lazarides, R., Dicke, A.-L., Rubach, C. & Eccles, J. S. (2020). Profiles of motivational beliefs in math: Exploring their development, relations to student-perceived classroom characteristics and impact on future career aspirations and choices. *Journal of Educational Psychology*, 112(1), 70–92. <https://doi.org/10.1037/edu0000368>

- Lazarides, R. & Lauermaun, F. (2019). Gendered path into STEM-related and language-related careers: Girls' and boys' motivational beliefs and career plans in math and language arts. *Frontiers in Psychology, 10*, 1–17. <https://doi.org/10.3389/fpsyg.2019.01243>
- Lazarsfeld, P. F. (1940). Panel studies. *Public Opinion Quarterly, 4*(1), 122–128. <https://doi.org/10.1086/265373>
- Lazarsfeld, P. F. & Henry, N. W. (Hrsg.). (1968). *Latent structure analysis*. New York: Houghton Mifflin.
- Lehmann-Wermser, A., Naacke, S., Nonte, S. & Ritter, B. (Hrsg.). (2010). *Musisch-kulturelle Bildung an Ganztagschulen. Empirische Befunde, Chancen und Perspektiven*. Weinheim: Juventa.
- Leist, S. & Pietsch, M. (2017). Bordering the area of spatial relevance for schools: A stochastic network approach using the example of Hamburg, Germany. *Belgeo, 2–3*, 1–19. <https://doi.org/10.4000/belgeo.20332>
- Lewin, K. (1938). *The conceptual representation and the measurement of psychology forces*. Durham, NC: Duke University Press.
- Li, C.-H. (2016a). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Li, C.-H. (2016b). The performance of ML, DWLS and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods, 21*(3), 369–387. <https://doi.org/10.1037/met0000093>
- Liang, W. & Yang, Y. (2014). An evaluation of WLSMV and bayesian methods for confirmatory factor analysis with categorical indicators. *International Journal of Quantitative Research in Education, 2*(1), 17–38. <https://doi.org/10.1504/IJQRE.2014.060972>
- Little, T. D. (2013). *Longitudinal structure equation modelling*. New York: The Guilford Press.
- Liu, Y., Millsap, R. E., West, S. G., Tein, J., Tanaka, R. & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods, 22*(3), 486–506. <https://doi.org/10.1037/met0000075>
- Lohaus, A. & Vierhaus, M. (2019). *Entwicklungspsychologie des Kindes- und Jugendalters für Bachelor* (4. Aufl.). Berlin: Springer-Verlag. <https://doi.org/10.1007/978-3-662-59192-5>
- Lorke, J. (2018). Museen und Science Center. In K. Sommer, J. Wambach-Laicher & P. Pfeifer (Hrsg.), *Konkrete Fachdidaktik Chemie. Grundlagen für das Lernen und Lehren im Chemieunterricht* (S. 678–683). Seelze: Aulis Verlag.

- Lowe, G. & Coy, N. (2016). Immersion, relevance and transferability: The motivational preferences of lower secondary students towards a newly created praxis-based class music program. *Australian Journal of Music Education*, 50(2), 39–47.
- Lubke, G. H. & Muthén, B. O. (2004). Applying multigroup confirmatory factor models for continuous outcomes to likert scale data complicates meaningful group comparisons. *Structural Equation Modeling: A Multidisciplinary Journal*, 11(4), 514–534. https://doi.org/10.1207/s15328007sem1104_2
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58(2), 103–117. <https://doi.org/10.1026/0033-3042.58.2.103>
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23(1), 129–149. <https://doi.org/10.3102/00028312023001129>
- Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, 82(4), 623–636. <https://doi.org/10.1037/0022-0663.82.4.623>
- Marsh, H. W. & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology*, 95(4), 687–706. <https://doi.org/10.1037/0022-0663.95.4.687>
- Marsh, H. W., Barnes, J., Cairns, L. & Tidman, M. (1984). Self-description questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology*, 76(5), 940–956. <https://doi.org/10.1037/0022-0663.76.5.940>
- Marsh, H. W., Bryne, B. M. & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80(3), 366–380. <https://doi.org/10.1037/0022-0663.80.3.366>
- Marsh, H. W., Kong, C.-K. & Hau, K.-T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glorry effects in Hong Kong Schools. *Journal of Personality and Social Psychology*, 78(2), 337–349. <https://doi.org/10.1037//0022-3514.78.2.337>
- Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J. S., Möller, J. & Abduljabbar, A. S. (2014). Dimensional comparison theory: An extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology*, 39(4), 326–341. <https://doi.org/10.1016/j.cedpsych.2014.08.003>

- Marsh, H. W., Lüdtke, O., Trautwein, U. & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(2), 191–225. <https://doi.org/10.1080/10705510902751010>
- Marsh, H. W., Martin, A. J., Yeung, A. S. & Craven, R. G. (2017). Competence self-perceptions. In A. J. Elliot, C. S. Dweck & D. S. Yeager (Hrsg.), *Handbook of competence and motivation. Theory and application* (S. 85–115). New York: The Guilford Press.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J. & Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, 44(3), 631–669. <https://doi.org/10.3102/0002831207306728>
- Marsh, H. W., Trautwein, U., Lüdtke, O. & Köller, O. (2008). Social comparison and big-fish–little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology*, 100(3), 510–524. <https://doi.org/10.1037/0022-0663.100.3.510>
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>
- McCormick, J. & McPherson, G. E. (2007). Expectancy-value motivation in the context of a music performance examination. *Musicae Scientiae*, 37–52.
- McGregor, G. & Mills, M. (2006). Boys and music education: RMXing the curriculum. *Pedagogy, Culture & Society*, 14(2), 221–233. <https://doi.org/10.1080/14681360600738350>
- McPherson, G. E. & McCormick, J. (2006). Self-efficacy and music performance. *Psychology of Music*, 34(3), 325–339. <https://doi.org/10.1177/0305735606064841>
- McPherson, G. E., Osborne, M. S., Barrett, M. S., Davidson, J. W. & Faulkner, R. (2015). Motivation to study music in Australian schools: The impact of music learning, gender and socio-economic status. *Research Studies in Music Education*, 37(2), 141–160. <https://doi.org/10.1177/1321103X15600914>
- Mead, G. H. (1934). *Mind, self and society from the standpoint of a social behaviorist*. Chicago: University of Chicago Press.
- Meade, A. W. & Lautenschlager, G. Y. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7(4), 361–388. <https://doi.org/10.1177/1094428104268027>
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <https://doi.org/10.1007/BF02294825>

- Meredith, W. & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins, A. G. Sayer (Hrsg.), *New methods for the analysis of change* (S. 203–240). Washington, D.C: American Psychological Association.
- Messner, R. (2012). Forschendes Lernen als Element praktischer Lehr-Lernkultur. In W. Blum, R. Borromeo Ferri & K. Maaß (Hrsg.), *Mathematikunterricht im Kontext von Realität, Kultur und Lehrerprofessionalität* (S. 334–346). Wiesbaden: Springer Spektrum. https://doi.org/10.1007/978-3-8348-2389-2_34
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.
- Millsap, R. E. & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, 39(3), 479–515. https://doi.org/10.1207/S15327906MBR3903_4
- MINT-EC. (2021). MINT-EC – Das nationale Excellence-Schulnetzwerk. Zugriff am 24.09.2021 unter <https://www.mint-ec.de/mint-ec/ueber-uns/>
- Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, 85(3), 424–436. <https://doi.org/10.1037/0022-0663.85.3.424>
- Mokhonko, S. (2016). *Nachwuchsförderung im MINT-Bereich. Aktuelle Entwicklungen, Fördermaßnahmen und ihre Effekte*. Stuttgart: Franz Steiner Verlag.
- Mokhonko, S., Nickolaus, R. & Windaus, A. (2014). Förderung von Mädchen in Naturwissenschaften: Schülerlabore und ihre Effekte. *Zeitschrift für Didaktik der Naturwissenschaften*, 20, 143–159. <https://doi.org/10.1007/s40573-014-0016-2>
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55, 19–27. <https://doi.org/10.1026/0033-3042.55.1.19>
- Möller, J., Pohlmann, B., Köller, O. & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, 79(3), 1129–1167. <https://doi.org/10.3102/0034654309337522>
- Möller, J. & Trautwein, U. (2020). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (3. Aufl., S. 187–209). Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-662-61403-7_8
- Morris-Lange, S., Wendt, H. & Wohlfarth, C. (2013). *Segregation an deutschen Schulen. Ausmaß, Folgen und Handlungsempfehlungen für bessere Bildungschancen*. Zugriff am 15.06.2021 unter https://www.researchgate.net/publication/303718332_Segregation_an_deutschen_Schulen_Ausmass_Folgen_und_Handlungsempfehlungen_fur_bessere_Bildungschancen

- Muthén, B. O. & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus (Mplus web notes: No. 4)*. Zugriff am 17.02.2021 unter <https://www.statmodel.com/download/webnotes/CatMGLong.pdf>
- Muthén, L. K. & Muthén, B. O. (2012). *Mplus user's guide* (7. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Muthén, B. O., Muthén, L. K. & Asparouhov, T. (2015). *Estimator choices with ordinal outcomes*. Zugriff am 13.02.2021 unter <https://www.statmodel.com/download/EstimatorChoices.pdf>
- Nagengast, B. & Marsh, H. W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*, 104(4), 1033–1053. <https://doi.org/10.1037/a0027697>
- Nagy, G., Trautwein, U., Baumert, J., Köller, O. & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, 12(4), 323–345. <https://doi.org/10.1080/13803610600765687>
- Newsom, J. T. (2015). *Longitudinal structural equation modeling. A comprehensive introduction*. New York: Routledge.
- Nick, S., Urhahne, D., Stang, J. & Parchmann, I. (2015). Lässt sich Erfolg bei Wettbewerben vorhersagen? In S. Bernholt (Hrsg.), *Heterogenität und Diversität – Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht* (S. 70–72). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Niedersächsisches Kultusministerium. (2014). *Die Arbeit in den Schuljahrgängen 5 bis 10 der Integrierten Gesamtschule (IGS)*. RdErl. d. MK v. 1.8.2014 – 34-81071 – VORIS 22410. Zugriff am 22.06.2021 unter https://www.mk.niedersachsen.de/download/91811/Erlass_Die_Arbeit_in_den_Schuljahrgaengen_5_bis_10_der_Integrierten_Gesamtschule_IGS_vom_01.08.2014.pdf
- Niedersächsisches Kultusministerium. (2015a). *Die Arbeit in den Schuljahrgängen 5 bis 10 der Kooperativen Gesamtschule (KGS)*. RdErl. d. MK v. 3.8.2015 – 34 – 81072 – VORIS 22410. Zugriff am 22.06.2021 unter https://www.mk.niedersachsen.de/download/100080/Erlass_Die_Arbeit_in_den_Schuljahrgaengen_5_bis_10_der_Kooperativen_Gesamtschule_KGS_-_Stand_1.8.2015.pdf
- Niedersächsisches Kultusministerium. (2015b). *Die Arbeit in den Schuljahrgängen 5 bis 10 des Gymnasiums*. RdErl. d. MK v. 23.6.2015 – 33-81011 – VORIS 22410. Zugriff am 22.06.2021 unter https://www.mk.niedersachsen.de/download/98074/Erlass_Die_Arbeit_in_den_Schuljahrgaengen_5_bis_10_des_Gymnasiums_v_23.6.2015.pdf

- Niedersächsisches Kultusministerium. (2019). *Klassenbildung und Lehrkräftestundenzuweisung an den allgemein bildenden Schulen. RdErl. d. MK v. 21. 03. 2019 – 34-84001/3 – VORIS 22410*. Zugriff am 22.06.2021 unter https://www.mk.niedersachsen.de/download/149600/Runderlass_Klassenbildung_und_Lehrkraeftestundenzuweisung_an_den_allgemein_bildenden_Schulen_vom_21.03.2019.pdf
- Niedersächsisches Kultusministerium. (2020a). *Die wichtigsten Fragen und Antworten zum Übergang von der Grundschule auf die weiterführende Schule*. Zugriff am 13.07.2021 unter https://www.mk.niedersachsen.de/download/108312/Die_wichtigsten_Fragen_und_Antworten_zum_Uebergang_von_der_Grundschule_auf_eine_weiterfuehrende_Schule.pdf
- Niedersächsisches Kultusministerium. (2020b). *Die Arbeit in der Grundschule. RdErl d. MK vom 01.08.2020 – VORIS 22410*. Zugriff am 13.07.2021 unter https://www.mk.niedersachsen.de/download/140388/Die_Arbeit_in_der_Grundschule_RdErl_d_MK_vom_01.08.2020.pdf
- Nonte, S. (2013). Entwicklungen und Auswirkungen der Schulprofilierung an allgemeinbildenden Schulen in ausgewählten europäischen Ländern und Implementationsperspektiven für Deutschland. *International Review of Education*, 59(2), 243–262. <https://doi.org/10.1007/s11159-013-9357-1>
- Nonte, S. (2022). Hintergründe von Schul- und Klassenprofilierung. In T. C. Stubbe, S. Nonte, M. Haas & M. Krieg (Hrsg.), *Musik- und MINT-Profile an Niedersächsischen Gymnasien und integrierten Gesamtschulen. Ergebnisse der Studie ProBiNi*. Münster: Waxmann. Manuskript in Vorbereitung.
- Nonte, S., Haas, M. & Stubbe, T. C. (2019). Unterrichtsqualität in MINT-Klassen. Befunde aus der quasiexperimentellen Studie ProBiNi. In T. Ehmke, P. Kuhl & M. Pietsch (Hrsg.), *Lehrer. Bildung. Gestalten. Beiträge zur empirischen Forschung in der Lehrerbildung* (S. 325–335). Weinheim: Juventa.
- Nonte, S., Knigge, J. & Stubbe, T. C. (2019). Differentielle Itemfunktionen und Messinvarianz in standardisierten Musiktests – eine Frage der Testfairness. *Beiträge empirischer Musikpädagogik*, 10, 1–21.
- Nonte, S., Krieg, M. & Stubbe, T. C. (2021). Is gender role self-concept a predictor for music class attendance? Findings from secondary schools in Lower Saxony (Germany). *Psychology of Music*, 1–17. <https://doi.org/10.1177/03057356211042931>
- Nonte, S. & Naacke, S. (2010). MUKUS – Die Ergebnisse. In A. Lehmann-Wermser, S. Naacke, S. Nonte & B. Ritter (Hrsg.), *Musisch-kulturelle Bildung an Ganztagschulen. Empirische Befunde, Chancen und Perspektiven* (S. 53–210). Weinheim: Juventa.

- Nonte, S. & Schurig, M. (2019). Instrumentalunterricht am Übergang vom Kindes- zum Jugendalter. Wer bleibt am Ball? Determinanten im Kindesalter zur Vorhersage der Teilnahme am Instrumentalunterricht in der sechsten und siebten Jahrgangsstufe. In K. Schwippert, A. Lehmann-Wermser & V. Busch (Hrsg.), *Mit Musik durch die Schulzeit? Chancen des Schulprogramms JeKi – Jedem Kind ein Instrument* (S. 31–58). Münster: Waxmann.
- Nonte, S. & Schwippert, K. (2012). Musikalische und sportliche Profile an Grundschulen – Auswirkungen auf Klassenklima und Selbstkonzept. *Beiträge empirischer Musikpädagogik*, 3(1), 2–25.
- Nonte, S., Steinmayr, R. & Scholz, L. A. (2020). Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 223–262). Münster: Waxmann. <https://doi.org/10.31244/9783830993193>
- Nonte, S. & Stubbe, T. C. (2017, September). *Wer entscheidet sich für eine Profilklassen. Vertiefende Analysen zu ProBiNi*. Symposium durchgeführt auf der 82. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF), Tübingen.
- O'Mara, A. J., Green, J. & Marsh, H. W. (2006). Administering self-concept interventions in schools: No training necessary? A meta-analysis. *International Education Journal*, 7(4), 524–533.
- Otto Hahn Gymnasium Göttingen. (2021). *Die Musikklasse*. Zugriff am 19.08.2021 unter https://www.ohg-goe.net/files/theme_files/artikel/sekundarstufe1/musikzweig/MusikklassePPPneu_v21.pdf
- Pabst-Krueger, M. (2020). Klassenmusizieren. In W. Jank (Hrsg.), *Musikdidaktik. Praxishandbuch für die Sekundarstufe I und II* (8. Aufl., S. 158–168). Berlin: Cornelsen.
- Palmer, D. H. (2009). Student interest generated during an inquiry skills lesson. *Journal of Research in Science Teaching*, 46(2), 147–165. <https://doi.org/10.1002/tea.20263>
- Pawek, C. (2009). *Schülerlabore als interesselördernde außerschulische Lernumgebungen für Schülerinnen und Schüler aus der Mittel- und Oberstufe* (Dissertation, Universität Kiel). Zugriff am 10.09.2021 unter https://macau.uni-kiel.de/receive/diss_mods_00003669
- Pekrun, R. (1988). *Emotion, Motivation und Persönlichkeit*. Weinheim: Beltz.
- Penfield, R. D. & Lam, T. C. M. (2000). Assessing differential item functioning in performance assessment: Review and recommendations. *Educational Measurement: Issues and Practice*, 19(3), 5–15. <https://doi.org/10.1111/j.1745-3992.2000.tb00033.x>
- Piaget, J. (1960). *The psychology of intelligence*. Patterson: Littlefield, Adams.

- Pietsch, M., Bonsen, M. & Bos, W. (2006). Ein Index sozialer Belastung als Grundlage für die Rückmeldung ‚fairer Vergleiche‘ von Grundschulen in Hamburg. In W. Bos & M. Pietsch (Hrsg.), *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen* (S. 225–245). Münster: Waxmann.
- Potvin, P. & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, 50(1), 85–129. <https://doi.org/10.1080/03057267.2014.881626>
- Potvin, P., Hasni, A., Sy, O. & Riopel, M. (2020). Two crucial years of science and technology schooling: A longitudinal study of the major influences on and interactions between self-concept, interest and the intention to pursue S&T. *Research in Science Education*, 50, 1739–1761. <https://doi.org/10.1007/s11165-018-9751-6>
- Prenzel, M. (1988). *Die Wirkungsweise von Interesse. Ein pädagogisch-psychologisches Erklärungsmodell*. Opladen: Westdeutscher Verlag.
- Prenzel, M. & Krapp, A. (1992). Zur Aktualität der pädagogisch-psychologischen Interessenforschung. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze der pädagogisch-psychologischen Interessenforschung* (S. 1–8). Münster: Aschendorff.
- Prenzel, M., Lankes, E.-M. & Minsel, B. (2000). Interessenentwicklung in Kindergarten und Grundschule: Die ersten Jahre. In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 11–29). Münster: Waxmann.
- R Development Core Team. (2017). *R: A language and environment for statistical computing* [Computer software]. Vienna: R Foundation for Statistical Computing. Zugriff am 22.03.2021 unter <http://www.R-project.org>
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Reimann, M., Herzog, S., Parchmann, I. & Schwarzer, S. (2020). Wirksamkeit der schulischen Vor- und Nachbereitung eines Schülerlaborbesuches. *Zeitschrift für Didaktik der Naturwissenschaften*, 26, 227–240. <https://doi.org/10.1007/s40573-020-00121-y>
- Reinders, H. (2006). Kausalanalysen in der Längsschnittforschung. Das Cross-Lagged-Panel Design. *Diskurs Kindheits- und Jugendforschung*, 1(4), 569–587.
- Renninger, K. A., Hidi, S. & Krapp, A. (Hrsg.). (1992). *The role of interest in learning and development*. Hilldale: Erlbaum.

- Rheinberg, F. & Vollmeyer, R. (2000). Sachinteresse und leistungsthematische Herausforderung – zwei verschiedenartige Motivationskomponenten und ihr Zusammenwirken beim Lernen. In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 145–162). Münster: Waxmann.
- Rhemtulla, M., Brosseau-Liard, P. E. & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354–373. <https://doi.org/10.1037/a0029315>
- Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., Harring, J. R. & Roberts, B. W. (2019). The effects of getting a new teacher on the consistency of personality. *Journal of Personality*, 87(3), 485–500. <https://doi.org/10.1111/jopy.12410>
- Rippl, S. & Seipel, C. (1997). Gruppenunterschiede – Fakt oder Artefakt? Das Problem der Bedeutungsäquivalenz am Beispiel der Rechtsextremismusforschung. *Zeitschrift für Soziologie*, 26(2), 139–150.
- Rippl, S. & Seipel, C. (2015). *Methoden kulturvergleichender Sozialforschung. Eine Einführung*. Wiesbaden: Springer VS.
- Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education*, 5, 1–7. <https://doi.org/10.3389/educ.2020.589965>
- Robitzsch, A., Kiefer, T. & Wu, M. (2017). *TAM: Test analysis modules* [Computer-software]. Zugriff am 22.03.2021 unter <https://CRAN.R-project.org/package=TAM>
- Röllke, K. (2019). *Was kann ein Schülerlabor leisten? Konzeptionierung des Schülerlabors tutolab-biotechnologie als Lehr-Lern-Labor mit Angeboten zur Breiten- und Begabtenförderung von Schülerinnen und Schülern* (Dissertation, Universität Bielefeld). Zugriff am 10.09.2021 unter https://pub.uni-bielefeld.de/download/2935577/2935981/R%C3%B6llke_Thesis_Gesamt.pdf
- Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Rosenbaum, P. R. & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33–38. <https://doi.org/10.2307/2683903>
- Rosenzweig, E. O. & Wigfield, A. (2017). What if reading is easy but unimportant? How students' patterns of affirming and undermining motivation for reading information texts predict different reading outcomes. *Contemporary Educational Psychology*, 48, 133–148. <https://doi.org/10.1016/j.cedpsych.2016.09.002>

- Rost, D. H. & Sparfeldt, J. R. (2002). Facetten des schulischen Selbstkonzepts. Ein Verfahren zur Messung des differentiellen Selbstkonzepts schulischer Leistungen und Fähigkeiten (DISK-Gitter). *Diagnostica*, 48(3), 130–140. <https://doi.org/10.1026//0012-1924.48.3.130>
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581–592. <https://doi.org/10.2307/2335739>
- Rürup, M. (2019). Schulautonomie in Deutschland – Ein Dauerthema der Schulreform. In E. Rauscher, C. Wiesner, D. Paasch & P. Heißenberger (Hrsg.), *Schulautonomie – Perspektiven in Europa. Befunde aus dem EU-Projekt INNOVITAS* (S. 61–75). Münster: Waxmann. <https://doi.org/10.25656/01:18141>
- Rutkowski, L. & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, 74(1), 31–57. <https://doi.org/10.1177/0013164413498257>
- Rutkowski, L. & Svetina, D. (2017). Measurement invariance in international surveys: Categorical indicators and fit measure performance. *Applied Measurement in Education*, 30(1), 39–51. <https://doi.org/10.1080/08957347.2016.1243540>
- Ryan, R. M. & Powelson, C. L. (1991). Autonomy and relatedness as fundamental to motivation and education. *The Journal of Experimental Education*, 60(1), 49–66. <https://doi.org/10.1080/00220973.1991.10806579>
- Scherrer, V. & Preckel, F. (2019). Development of motivational variables and self-esteem during the school career: A meta-analysis of longitudinal studies. *Review of Educational Research*, 89(2), 211–258. <https://doi.org/10.3102/0034654318819127>
- Schiefele, H. (1974). *Lernmotivation und Motivlernen*. München: Ehrenwirth.
- Schiefele, H. (1986). Interesse – Neue Antworten auf ein altes Problem. *Zeitschrift für Pädagogik*, 32(2), 153–162. <https://doi.org/10.25656/01:14382>
- Schiefele, U. (1996). *Motivation und Lernen mit Texten*. Göttingen: Hogrefe.
- Schiefele, U. (2000). Befunde – Fortschritte – neue Fragen. In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 227–241). Münster: Waxmann.
- Schiefele, U. (2001). The role of interest in motivation and learning. In J. M. Collis & S. Messick (Hrsg.), *Intelligence and personality: Bridging the gap in theory and measurement* (S. 163–194). Mahwah: Erlbaum.
- Schnell, R., Hill, P. B. & Esser, E. (2018). *Methoden der empirischen Sozialforschung*. München: Oldenbourg Verlag.
- Schulte, K., Nonte, S. & Schwippert, K. (2013). Die Überprüfung von Messinvarianz in international vergleichenden Schulleistungstudien am Beispiel der Studie PIRLS. *Zeitschrift für Bildungsforschung*, 3(2), 99–118. <https://doi.org/10.1007/s35834-013-0062-8>

- Schwab, S. & Helm, C. (2015). Überprüfung von Messinvarianz mittels CFA und DIF-Analysen. *Empirische Sonderpädagogik*, 7(3), 175–193. <https://doi.org/10.25656/01:11380>
- Schwarzer, S. & Parchmann, I. (2018). Schülerlabore und Schülerforschungszentren. In K. Sommer, J. Wambach-Laicher & P. Pfeifer (Hrsg.), *Konkrete Fachdidaktik Chemie. Grundlagen für das Lernen und Lehren im Chemieunterricht* (S. 666–678). Seelze: Aulis Verlag.
- Schwippert, K., Lehmann-Wermser, A. & Busch, V. (Hrsg.). (2019). *Mit Musik durch die Schulzeit? Chancen des Schulprogramms JeKi – Jedem Kind ein Instrument*. Münster: Waxmann.
- Seaton, M., Marsh, H. W. & Craven, R. G. (2009). Earning its place as a pan-human theory: Universality of the big-fish-little-pond effect across 41 culturally and economically diverse countries. *Journal of Educational Psychology*, 101(2), 403–419. <https://doi.org/10.1037/a0013838>
- Seidel, T., Reinhold, S., Holzberger, D., Mok, S. Y., Schiepe-Tiska, A. & Reiss, K. (2016). *Wie gelingen MINT-Schulen? Anregungen aus Forschung und Praxis*. Münster: Waxmann.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46(3), 407–441. <https://doi.org/10.2307/1170010>
- Siedenburg, I. & Harbig, H. (Hrsg.). (2018a). *Kreatives Musizieren in der Bläserklasse. Wege zum Improvisieren und Komponieren zwischen Jazz und experimentellem Klang*. Münster: Waxmann.
- Siedenburg, I. & Harbig, H. (2018b). Kreatives Musizieren in der Bläserklasse – Editorial. In I. Siedenburg & G. Harbig (Hrsg.), *Kreatives Musizieren in der Bläserklasse. Wege zum Improvisieren und Komponieren zwischen Jazz und experimentellem Klang* (S. 1–7). Münster: Waxmann.
- Simon, F. (2019). *Der Einfluss von Betreuung und Betreuenden auf die Wirksamkeit von Schülerlaborbesuchen. Eine Zusammenhanganalyse von Betreuungsqualität, Betreuungsmerkmalen und Schülerlaborzielen sowie Replikationsstudie zur Wirksamkeit von Schülerlaborbesuchen* (Dissertation, Technische Universität Dresden). Zugriff am 10.09.2021 unter <https://tud.qucosa.de/api/qucosa%3A37051/attachment/ATT-0/>
- Sonnenburg, N. (2020). Schulprofilierung im Kontext internationaler Entwicklungen. Einflüsse auf die Entscheidung für das Schulprofil Europaschule. In C. I. Keßler & S. Nonte (Hrsg.), *(Neue) Formen der Differenzierung. Schul- und Klassenprofilierung im Spannungsfeld gesellschaftlicher Teilhabe und sozialer Ungleichheit* (S. 95–111). Münster: Waxmann.

- Specht, W. (2011). Restschulen und Restklassen. Ein vernachlässigtes Phänomen im Gefolge neuer Steuerungsformen. In H. Altrichter, M. Heinrich & K. Soukup-Altrichter (Hrsg.), *Schulentwicklung durch Schulprofilierung? Zur Veränderung von Koordinationsmechanismen im Schulsystem* (S. 141–160). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92825-8_4
- Spinath, B. & Spinath, F. M. (2005). Longitudinal analysis of the link between learning motivation and competence beliefs among elementary school children. *Learning and Instruction*, 15(2), 87–102. <https://doi.org/10.1016/j.learninstruc.2005.04.008>
- Spinath, B. & Steinmayr, R. (2008). Longitudinal analysis of intrinsic motivation and competence beliefs: Is there a relation over time? *Child Development*, 79(5), 1555–1569. <https://doi.org/10.1111/j.1467-8624.2008.01205.x>
- Spychiger, M. (2013). Das musikalische Selbstkonzept: Wer ich bin und was ich kann in der Musik. *Üben & Musizieren*, 6, 18–21.
- Stang, J., Urhahne, D., Nick, S. & Parchmann, I. (2014). Wer kommt weiter? Vorhersage der Qualifikation zur Internationalen Biologie- und Chemie-Olympiade auf Grundlage des Leistungsmotivations-Modells von Eccles. *Zeitschrift für Pädagogische Psychologie*, 28(3), 105–114. <https://doi.org/10.1024/1010-0652/a000127>
- Stark, S., Chernyshenko, O. S. & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology*, 91(6), 1292–1306. <https://doi.org/10.1037/0021-9010.91.6.1292>
- Stäudel, L. (2018). Naturwissenschaftlicher Anfangsunterricht. In K. Sommer, J. Wambach-Laicher & P. Pfeifer (Hrsg.), *Konkrete Fachdidaktik Chemie. Grundlagen für das Lernen und Lehren im Chemieunterricht* (S. 608–616). Seelze: Aulis Verlag.
- Steeh, A., Höffler, T., Höft, L. & Parchmann, I. (2021a). Exploring science competition participants' expectancy-value perceptions and identification: A latent profile analysis. *Learning and Instruction*, 74, 1–10. <https://doi.org/10.1016/j.learninstruc.2021.101455>
- Steeh, A., Höffler, T., Höft, L. & Parchmann, I. (2021b). First steps toward gender equity in the chemistry olympiad: Understanding the role of implicit gender-science stereotypes. *Journal of Research in Science Teaching*, 58(1), 40–68. <https://doi.org/10.1002/tea.21645>
- Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling: A Multidisciplinary Journal*, 5(4), 411–419. <https://doi.org/10.1080/10705519809540115>

- Streller, M. (2015). *The educational effects of pre and post-work in out-of-school laboratories* (Dissertation, Technische Universität Dresden). Zugriff am 10.09.2021 unter <https://d-nb.info/1083463365/34>
- Stubbe, T. C., Nonte, S., Haas, M. & Krieg, M. (Hrsg.). (2022). *Musik- und MINT-Profile an Niedersächsischen Gymnasien und integrierten Gesamtschulen. Ergebnisse der Studie ProBiNi*. Münster: Waxmann. Manuskript in Vorbereitung.
- Stubbe, T. C., Krieg, M., Haas, M. & Nonte, S. (2022). Anlage und Durchführung von ProBiNi. In T. C. Stubbe, S. Nonte, M. Haas & M. Krieg (Hrsg.), *Musik- und MINT-Profile an Niedersächsischen Gymnasien und integrierten Gesamtschulen. Ergebnisse der Studie ProBiNi*. Münster: Waxmann. Manuskript in Vorbereitung.
- Svetina, D. & Rutkowski, L. (2017). Multidimensional measurement invariance in an international context: Fit measure performance with many groups. *Journal of Cross-Cultural Psychology*, 48(7), 991–1008. <https://doi.org/10.1177/0022022117717028>
- Svetina, D., Rutkowski, L. & Rutkowski, D. (2020). Multi-group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/semTool packages. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(1), 111–130. <https://doi.org/10.1080/10705511.2019.1602776>
- Swarat, S., Ortony, A. & Revelle, W. (2012). Activity matters: Understanding interest in school science. *Journal of Research in Science Teaching*, 49(4), 515–537. <https://doi.org/10.1002/tea.21010>
- Teman, E. D. (2012). *The performance of multiple imputation and full information maximum likelihood for missing ordinal data in structural equation models*. Ann Arbor, MI: ProQuest.
- Temme, D. & Hildebrandt, L. (2009). Gruppenvergleiche bei hypothetischen Konstrukten – Die Prüfung der Übereinstimmung von Messmodellen mit der Strukturgleichungsmethodik. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 61, 138–185.
- Terhag, J. (2012). Musizieren mit Schulklassen. Argumente für die Überwindung von Praxisfeindlichkeit und Theorieverdrossenheit. In M. Pabst-Krueger & J. Terhag (Hrsg.), *Musizieren mit Schulklassen. Praxis – Konzepte – Perspektiven* (S. 13–24). Oldershausen: Lugert.
- Theorell, T., Lennartsson, A.-K., Madison, G., Mosing, M. A. & Ullén, F. (2015). Predictors of continued playing or singing – from childhood and adolescence to adult years. *Acta Paediatrica*, 104(3), 274–284. <https://doi.org/10.1111/apa.12870>
- Thiel, A. & Teubert, H. (2018). Die Bewegte Schule. In H. Barz (Hrsg.), *Handbuch Bildungsreform und Reformpädagogik* (S. 503–510). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-07491-3_47
- Todt, E. (1978). *Das Interesse. Empirische Untersuchungen zu einem Motivationskonzept*. Bern: Huber.

- Todt, E. (1985). Die Bedeutung der Schule für die Entwicklung der Interessen von Kindern und Jugendlichen. *Unterrichtswissenschaft*, 4, 362–376.
- Todt, E. (1990). Entwicklung des Interesses. In H. Hetzer & R. Arbinger (Hrsg.), *Angewandte Entwicklungspsychologie des Kindes- und Jugendalters* (S. 213–264). Heidelberg: Quelle & Meyer.
- Todt, E. & Schreiber, S. (1998). Development of interest. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Secon conference on interest and gender* (S. 25–40). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Tolman, E. C. (1932). *Purposive behavior in animals and men*. New-York: Appleton-Century-Crofts.
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O. & Baumert, J. (2006). Tracking, grading and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, 98(4), 788–806. <https://doi.org/10.1037/0022-0663.98.4.788>
- Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G. & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study. *Journal of Educational Psychology*, 104(3), 763–777. <https://doi.org/10.1037/a0027470>
- Travers, R. M. W. (1978). *Children's interests*. Kalamazoo, MI: Michigan University, College of Education.
- Tucker, L. R. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1–10.
- Unger, C. (2015). *Wettbewerbssteuerung im Primarschulbereich. Vergleichende Fallanalysen*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-07440-1>
- van Buuren, S. & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- van der Gaag, M. & Snijders, T. A. B. (2005). The resource generator: Social capital quantification with concrete items. *Social Networks*, 27(1), 1–29. <https://doi.org/10.1016/j.socnet.2004.10.001>
- van Vorst, H., Fechner, S. & Sumfleth, E. (2012). Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie. In S. Bernholt (Hrsg.), *Inquiry-based learning – Forschendes Lernen* (S. 311–313). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN).
- Varjo, J. & Kalalahti, M. (2019). The art of governing local education markets – Municipalities and school choice in Finland. *Education Inquiry*, 10(2), 151–165. <https://doi.org/10.1080/20004508.2018.1514907>
- Verband deutscher Musikschulen. (2021). *VdM Jahresbericht 2020 – Themenschwerpunkte und statistische Daten*. Bonn: VdM Verlag.

- von Hippel, P. T. (2015). New confidence intervals and bias comparisons show that maximum likelihood can beat multiple imputation in small samples. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(3), 1–16. <https://doi.org/10.1080/10705511.2015.1047931>
- Wang, J. & Wang, X. (2020). *Structural equation modeling: Applications using Mplus* (2. Aufl.). Hoboken, NJ: Wiley.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450. <https://doi.org/10.1007/BF02294627>
- Warzecha, M. (2013). Boys' perception of singing: A review of the literature. *National Association for Music Education*, 32(1), 43–51. <https://doi.org/10.1177/8755123313502341>
- Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values and task perceptions according to gender and domain in 7th- through 11th-grade Australian students. *Child Development*, 75(5), 1556–1574.
- Weber, C., Moosbrugger, R., Hasengruber, K., Altrichter, H. & Schrodtr, H. (2019). Wer unterrichtet wen? Die Zusammensetzung von Klassen und Schu- len und die Zuteilung von Lehrkräften. In S. Breit, F. Eder, K. Krainer, C. Schreiner, A. Seel & C. Spiel (Hrsg.), *Nationaler Bildungsbericht Österreich 2018. Fokussierte Analysen und Zukunftsperspektiven für das Bildungswesen* (S. 143–182). Graz: Leykam.
- Weber, T., Nonte, S. & Stubbe, T. C. (2021). Die Verbreitung von Schul- und Klassenprofilen an weiterführenden Schulen in Hamburg und im Regierungs- bezirk Braunschweig in Niedersachsen. In M. Haas & M. Krieg (Hrsg.), *Lernen gestalten. Beiträge der empirischen Forschung zu Innovation in Schule und Unterricht* (S. 29–45). Göttingen: Universitätsverlag.
- Weber, E. W., Spychiger, M. & Patry, J.-L. (1993). *Musik macht Schule: Biografie und Ergebnisse eines Schulversuchs mit erweitertem Musikunterricht*. Essen: Die Blaue Eule. Zugriff am 28.09.2021 unter <https://ewaweber.ch/musik-macht-schule>
- Weiß, M. (2001). Quasi-Märkte im Schulbereich. Eine ökonomische Analyse. *Zeitschrift für Pädagogik*, 43, 69–85. <https://doi.org/10.25656/01:7915>
- Welch, A. & Huffman, D. (2011). The effect of robotics competitions on high school students' attitudes toward science. *School Science and Mathematics*, 111(8), 416–424. <https://doi.org/10.1111/j.1949-8594.2011.00107.x>
- Widaman, K. F., Ferrer, E. & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*, 4(1), 10–18. <https://doi.org/10.1111/j.1750-8606.2009.00110.x>

- Widaman, K. F. & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle & S. G. West (Hrsg.), *The science of prevention: Methodical advances from alcohol and substances abuse research* (S. 281–324). Washington, DC: American Psychological Association.
- Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, 6(1), 49–78. <https://doi.org/10.1007/BF02209024>
- Wigfield, A. & Cambria, J. (2010). Students' achievement values, goal orientations and interest: Definitions, development and relations to achievement outcomes. *Developmental Review*, 30(1), 1–35. <https://doi.org/10.1016/j.dr.2009.12.001>
- Wigfield, A. & Eccles, J. S. (1994). Children's competence beliefs, achievement values and general self-esteem. Change across elementary and middle school. *Journal of Early Adolescence*, 14(2), 107–138. <https://doi.org/10.1177/027243169401400203>
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbretton, A. J. A., Freedman-Doan, C. & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across elementary school years: A 3-year study. *Journal of Educational Psychology*, 89(3), 451–469. <https://doi.org/10.1037/0022-0663.89.3.451>
- Wigfield, A., Rosenzweig, E. O. & Eccles, J. S. (2017). Achievement values. Interactions, interventions and future directions. In A. J. Elliot, C. S. Dweck & D. S. Yeager (Hrsg.), *Handbook of competence and motivation* (S. 116–134). New York: The Guilford Press.
- Wigfield, A., Tonks, S. M. & Klauda, S. L. (2016). Expectancy-value theory. In K. R. Wentzel & D. B. Miele (Hrsg.), *Handbook of motivation at school* (S. 55–74). New York: Routledge.
- Wilde, M., Bätz, K., Kovaleva, A. & Urhahne, D. (2009). Überprüfung einer Kurzsкала intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 31–45.
- Willems, A. S. (2018). Unterrichtsqualitätsprofile und ihr Zusammenhang zum situationalen Interesse in Mathematik. In A. Krüger, F. Radisch, A. S. Willems, T. H. Häcker & M. Walm (Hrsg.), *Empirische Bildungsforschung im Kontext von Schule und Lehrer*innenbildung* (S. 154–167). Bad Heilbrunn: Verlag Julius Klinkhardt.

- Wirth, R. J. (2008). *The effect of measurement non-invariance on parameter estimation in latent growth models* (Dissertation, Universität North Carolina). Zugriff am 01.02.2021 unter <https://core.ac.uk/download/pdf/210598885.pdf>
- Wohlkinger, F. (2014). *Die Rolle des Schülers bei der Wahl der weiterführenden Schule. Eine vergleichende Untersuchung von Grundschulern aus Bayern und Sachsen*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-06001-5>
- Wohlkinger, F. & Ditton, H. (2012). Entscheiden die Schüler mit? Der Einfluss von Eltern, Lehrern und Kindern auf den Übergang nach der Grundschule. In R. Becker & H. Solga (Hrsg.), *Soziologische Bildungsforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie* (S. 44–63). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-00120-9_2
- Wolter, I. & Hannover, B. (2016). Gender role self-concept at school start and its impact on academic self-concept and performance in mathematics and reading. *European Journal of Developmental Psychology*, 13(6), 681–703. <https://doi.org/10.1080/17405629.2016.1175343>
- Wrape, E. R., Dittloff, A. L. & Callahan, J. L. (2014). Gender and musical instrument stereotypes in middle school children: Have trends changed? *National Association for Music Education*, 34(3), 40–47. <https://doi.org/10.1177/8755123314564255>
- Wu, H. & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika*, 81(4), 1014–1045. <https://doi.org/10.1007/s11336-016-9506-0>
- Wulf, C., Haberstroh, S. & Petersen, M. (2020). Überblick. In C. Wulf, S. Haberstroh & M. Petersen (Hrsg.), *Forschendes Lernen. Theorie, Empirie, Praxis* (S. IX–XV). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-31489-7_1
- Wulf, C., Thiem, J. & Gess, C. (2020). Motivationale Faktoren im Wirkungskontext von Forschendem Lernen. In C. Wulf, S. Haberstroh & M. Petersen (Hrsg.), *Forschendes Lernen. Theorie, Empirie, Praxis* (S. 129–144). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-31489-7_10
- Yamaha. (2016). *Die Yamaha Bläserklasse*. Zugriff am 20.08.2021 unter https://de.yamaha.com/de/files/Brosch%C3%BCre%20Yamaha%20Bl%C3%A4serKlasse%202016_63d2397bed931aadca97ad8304e51330.pdf
- Yamaha. (2019). *Leitfaden Bläserklasse. Ein Konzept für das erfolgreiche Unterrichten mit Blasinstrumenten*. Zugriff am 20.08.2021 unter https://de.yamaha.com/de/files/2019_leitfaden_BK_b2978856711b703c8ef2c0fb353cbdf3.pdf

- Zunker, N. & Neumann, M. (2020). Stratifizierung durch Profilierung? Das Zusammenspiel von inhaltlichen Profilen und Merkmalen der Schülerkomposition an nichtgymnasialen weiterführenden Schulen in Berlin. In C. I. Keßler & S. Nonte (Hrsg.), *(Neue) Formen der Differenzierung. Schul- und Klassenprofilierung im Spannungsfeld gesellschaftlicher Teilhabe und sozialer Ungleichheit* (S. 35–53). Münster: 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. <https://doi.org/10.25656/01:22165>
- Zunker, N., Neumann, M. & Maaz, K. (2019). Schulwahl in Berlin – Wie hängen Schulangebot und Nachfrage zusammen? *Schulmanagement*, 50(1), 23–25. <https://doi.org/10.25656/01:17342>

11.2 Abbildungsverzeichnis

Abbildung 2.01: Rahmenmodell der Interessengenese.....	23
Abbildung 2.02: Bedingungen des Fach- und Sachinteresses.....	25
Abbildung 3.01: Multidimensionale und hierarchische Struktur des Selbstkonzepts (nach Shavelson et al., 1976).....	36
Abbildung 3.02: Struktur des schulischen Selbstkonzepts im revidierten Modell (nach Marsh et al., 1988)	38
Abbildung 4.01: Schaubild der Erwartungs- und Wertkomponente (nach Eccles & Wigfield, 2002).....	46
Abbildung 4.02: Erwartungs-Wert Modell (nach Eccles & Wigfield, 2002).....	47
Abbildung 8.01: ProBiNi-Studiendesign der Längsschnittstudie.....	102
Abbildung 8.02: Messmodell – konfirmatorische Faktorenanalyse.....	125
Abbildung 8.03: Bivariates manifestes Cross-Lagged-Panel-Modell mit zwei Messzeitpunkten.....	134
Abbildung 8.04: Bivariates latentes Cross-Lagged-Panel-Modell mit zwei Messzeitpunkten.....	135
Abbildung 9.01: Konfirmatorische Faktorenanalyse unter Annahme skalarer Messinvarianz im Längsschnitt des Konstrukts Fachinteresse (FI) in den Domänen Musik (MU) und Naturwissenschaften (NA)	143
Abbildung 9.02: Konfirmatorische Faktorenanalyse unter Annahme skalarer Messinvarianz im Längsschnitt des Konstrukts Sachinteresse (SI) in den Domänen Musik (MU) und Naturwissenschaften (NA)	145
Abbildung 9.03: Konfirmatorische Faktorenanalyse unter Annahme skalarer Messinvarianz im Längsschnitt des Konstrukts Fähigkeitsselbstkonzept (SK) in den Domänen Musik (MU) und Naturwissenschaften (NA)	147
Abbildung 9.04: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in der Gesamtstichprobe.....	160
Abbildung 9.05: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in der Gesamtstichprobe.....	161
Abbildung 9.06: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in der Gesamtstichprobe	163
Abbildung 9.07: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in der Gesamtstichprobe	164

-
- Abbildung 9.08: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in den Musik-Klassen und den Musik-Vergleichsklassen ohne Profil167
- Abbildung 9.09: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Musik (MU) in den Musik-Klassen und den Musik-Vergleichsklassen ohne Profil169
- Abbildung 9.10: Cross-Lagged-Panel-Modell – Zusammenhänge des Fachinteresses (FI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in den MINT-Klassen und den MINT-Vergleichsklassen ohne Profil171
- Abbildung 9.11: Cross-Lagged-Panel-Modell – Zusammenhänge des Sachinteresses (SI) und des Fähigkeitsselbstkonzepts (SK) in der Domäne Naturwissenschaften (NA) in den MINT-Klassen und den MINT-Vergleichsklassen ohne Profil173

11.3 Tabellenverzeichnis

Tabelle 8.01: Beschreibung der Gesamtstichprobe zum ersten Messzeitpunkt.....	104
Tabelle 8.02: Beschreibung der Analysestichprobe.....	105
Tabelle 8.03: Anteil der Mädchen und Jungen in der Analysestichprobe – absolute und relative Häufigkeiten.....	106
Tabelle 8.04: Beschreibung des sozialen Status in der Analysestichprobe.....	107
Tabelle 8.05: Kennwerte der Items zum musikalischen und naturwissenschaftlichen Fachinteresse zum ersten Messzeitpunkt.....	108
Tabelle 8.06: Kennwerte der Items zum musikalischen und naturwissenschaftlichen Sachinteresse zum ersten Messzeitpunkt.....	109
Tabelle 8.07: Kennwerte der Items zum musikalischen Fähigkeitsselbstkonzept zum ersten Messzeitpunkt.....	110
Tabelle 8.08: Kennwerte der Items zum naturwissenschaftlichen Fähigkeitsselbstkonzept zum ersten Messzeitpunkt.....	111
Tabelle 8.09: Interne Konsistenz (<i>Cronbachs a</i>) der zentralen Konstrukte zu den drei Messzeitpunkten (MZP)	112
Tabelle 9.01: Model-Fit der konfirmatorischen Faktorenanalysen der zentralen Konstrukte in den Domänen Musik und Naturwissenschaften.....	140
Tabelle 9.02: Model-Fit der Messinvarianzanalysen (Längsschnitt) des Konstrukts Fachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)	142
Tabelle 9.03: Model-Fit der Messinvarianzanalysen (Längsschnitt) des Konstrukts Sachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)	144
Tabelle 9.04: Model-Fit der Messinvarianzanalysen (Längsschnitt) des Konstrukts Fähigkeitsselbstkonzept in den Domänen Musik und Naturwissenschaften (Nawi)	146
Tabelle 9.05: Model-Fit der Messinvarianzanalysen (Gruppenvergleich) des Konstrukts Fachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)	148
Tabelle 9.06: Model-Fit der Messinvarianzanalysen (Gruppenvergleich) des Konstrukts Sachinteresse in den Domänen Musik und Naturwissenschaften (Nawi)	149
Tabelle 9.07: Model-Fit der Messinvarianzanalysen (Gruppenvergleich) des Konstrukts Fähigkeitsselbstkonzept in den Domänen Musik und Naturwissenschaften (Nawi)	150

Tabelle 9.08: Deskriptive Befunde – Entwicklung des durchschnittlichen musikalischen Fachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen).....	151
Tabelle 9.09: Deskriptive Befunde – Entwicklung des durchschnittlichen naturwissenschaftlichen Fachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen).....	152
Tabelle 9.10: Deskriptive Befunde – Entwicklung des durchschnittlichen musikalischen Sachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen).....	154
Tabelle 9.11: Deskriptive Befunde – Entwicklung des durchschnittlichen naturwissenschaftlichen Sachinteresses zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen).....	154
Tabelle 9.12: Deskriptive Befunde – Entwicklung des durchschnittlichen musikalischen Fähigkeitsselbstkonzepts zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen).....	155
Tabelle 9.13: Deskriptive Befunde – Entwicklung des durchschnittlichen naturwissenschaftlichen Fähigkeitsselbstkonzepts zwischen dem ersten und dritten Messzeitpunkt (MZP) in den Profilklassen und den jeweiligen Parallelklassen ohne Profil (Vergleichsklassen).....	157

Profilierungsmaßnahmen auf Klassenebene (sogenannte Profilklassen) stellen ein zentrales Element aktueller Schulentwicklungsprozesse dar und gelten als prominenteste Form der praktischen Umsetzung zunehmender Schulautonomie. In der Sekundarstufe I gewinnen vor allem Profilklassen im musikalischen sowie naturwissenschaftlichen Bereich (Musik- bzw. MINT-Klassen) zunehmend an Popularität. Einzelne Gestaltungsmerkmale des Profilverrichts (u.a. zusätzliche Unterrichtsstunden, innovative und praxisbezogene Lehr-Lern-Formate sowie der Einbezug außerschulischer Lernorte) werden mit vielfältigen pädagogischen Potentialen verbunden – insbesondere für das Interesse sowie das Fähigkeitsselbstkonzept der Schülerinnen und Schüler. Hinsichtlich einer differenzierten Analyse der Entwicklung des Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts der Schülerinnen und Schüler in Profilklassen der Sekundarstufe I besteht aktuell jedoch ein Forschungsdesiderat.

Die Ergebnisse der vorliegenden Arbeit geben Aufschluss über die Entwicklung des durchschnittlichen Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzepts zwischen der fünften und siebten Jahrgangsstufe in den Domänen Musik und Naturwissenschaften. Dabei wird der Effekt des Besuchs einer Musik- oder MINT-Profilklassen herausgestellt. Anhand latenter Cross-Lagged-Panel-Modelle werden darüber hinaus die Stabilität sowie die längsschnittliche Zusammenhangsstruktur des Fach- beziehungsweise Sachinteresses und des Fähigkeitsselbstkonzepts untersucht und mögliche Unterschiede zwischen Profilklassen und deren Parallelklassen ohne Profil beleuchtet. Die Befunde ermöglichen die Evaluation des Potentials von Profilklassen zur Förderung des Interesses sowie des Fähigkeitsselbstkonzepts und bieten vielfältige Anknüpfungspunkte für die pädagogische Praxis.